HUGO J (2010) 4:1–190
DOI 10.1007/s11568-010-9143-0

CONTENTS

Cover art: Dr. Paul Liam Harrison

Paul Harrison is an artist with a background and prevailing interest in print, printmaking and publishing. His practice inherently combines the use of traditional print methods and materials with new and developing technologies. He is currently a researcher at the University of Dundee, Visual Research Centre, where he co-ordinates and facilitates activity in fine print publishing.

His interests include the investigation of image processing and development currently taking place in and around the biosciences, and in particular the area of gene research. His work is generally of a collaborative nature and he is presently engaged in projects with scientists at the University of Dundee Biocentre, the Human Genetics Unit, MRC, Edinburgh and Cold Spring Harbor Laboratory (CSHL) New York. He is also artist in residence at the Human Genome Organisation (HUGO) and visiting fellow/artist in residence at the Centre for Economic and Social Aspects of Genomics (CESAGen) at the University of Cardiff. His most recent work includes a project funded by the Wellcome Trust, titled 'Designs for Life'.

An exhibition of work from the Designs for Life Project will be held at the Le Corum Conference Centre for the duration of HGM2010 where Paul will also be available for discussion. He will be conducting a series of informal interviews regarding personal experiences and views on the relationships between art and science and he would be delighted to hear from you if you would be interested in participating in such a discussion. Paul can be contacted on site or at; p.l.harrison@dundee.ac.uk

The 'Designs for Life' project can be viewed in further detail at:
www.designsforlifeproject.co.uk
A number of other projects can be viewed at:
www.paulliamharrison.co.uk or: www.vrc.dundee.ac.uk

Cover Image:
Title: 'Boulevard of Broken Dreams I'
(A collaboration between Dr. Paul Liam Harrison and Dr. Laura Trinkle-Mulcahy – produced as part of the 'Designs for Life' Project.)
Medium: Acrylic screenprint on BFK rives 280gsm paper.
Edition: Six
Size: 110 x 75 cm
Date: 2010

Welcomes you to the 14th Human Genome Meeting 2010 Montpellier, France

Edison T. Liu HUGO PRESIDENT

Mireille Claustres CHAIR, HGM2010

Human Genome Organisation (HUGO)
60 Biopolis Street
#02-01 Genome Building
Singapore 138672
Tel: +65 6808 8192
Fax: +65 6808 8297
E-mail: hugoadmin@gis.a-star.edu.sg
Website: www.hugo-international.org
www.hgm2010.org

The Human Genome Organisation (HUGO) is an international organisation of scientists involved in various aspects of human genetics. HUGO was conceived in 1988 at the first meeting on Genome Mapping and Sequencing at Cold Spring Harbor, New York. From an initial group of 42 scientists from 17 countries, HUGO has increased its membership to over 1,200 members, from 69 countries in a period of only two decades. Over the years, HUGO has played an essential role behind the scenes for the human genome project. With its mission to promote international collaborative efforts to study the human genome and the myriad issues raised by our increasing knowledge of the genome, HUGO has had noteworthy successes in some of the less glamorous, but nonetheless vital, aspects of the human genome project.

As a truly international organisation, and currently in its 22nd year, HUGO refocused its efforts towards the medical implications of genomic knowledge. Looking forward, HUGO is working to enhance the genomic capabilities of the emerging and developing countries of the world. The excitement and interest in genomic sciences in Asia, the Middle East, South America and Africa are palpable and our hope is that these technologies will help to aid national development and worldwide health.

Mission Statement

- to investigate the nature, structure, function and interaction of the genes, genomic elements and genomes of humans and relevant pathogenic and model organisms;
- to characterise the nature, distribution and evolution of genetic variation in humans and other relevant organisms;
- to study the relationship between genetic variation and the environment in the origins and characteristics of human populations and the causes, diagnoses, treatments and prevention of disease;
- to foster the interaction, coordination, and dissemination of information and technology between investigators and the global society in genomics, proteomics, bioinformatics, systems biology, and the clinical sciences by promoting quality education, comprehensive communication, and accurate, comprehensive, and accessible knowledge resources for genes, genomes and disease;
- to sponsor factually-grounded dialogues on the social, legal, and ethical issues related to genetic and genomic information and championing the regionally-appropriate, ethical utilisation of this information for the good of the individual and the society.

HGM2010
HUGO PRESIDENT
Edison T. Liu

HGM2010
INTERNATIONAL SCIENTIFIC COMMITTEE

Abecasis, Goncalo (USA)
Aitman, Tim (UK)
Antonarakis, Stylianos E. (Switzerland)
Brookes, Anthony (UK)
Bruford, Elspeth (UK)
Cann, Howard (France)
Cassiman, Jean-Jacques (France)
Chakravarti, Aravinda (USA)
Chen, Y T (Taiwan)
Cox, David (USA)
Estivill, Xavier (Spain)
Froguel, Philippe (UK)
Futreal, Andy (UK)
Gibbs, Richard (USA)
Gojobori, Takashi (Japan)
Hudson, Tom (Canada)
Jimenez-Sanchez, Gerardo (Mexico)
Khusnutdinova, Elza (Russia)
Kwok, Pui-Yan (USA)
Lancet, Doron (Israel)
Mandel, Jean-Louis (France)
McCarthy, Mark (UK)
Ropers, Hans-Hilger (Germany)
Schork, Nicholas J. (USA)
Taylor, Todd (Japan)
Van Heyningen, Vernonica (UK)
Yang, Huanming (China)

HGM2010
LOCAL ORGANISING COMMITTEE

CHAIRPERSON
Mireille Claustres (Montpellier, France)

Jean-Marie Blanchard (Montpellier, France)
Jacques Demaille (Montpellier, France)
Joel Bockaert (Montpellier, France)
Alain Bucheton (Montpellier, France)
Philippe Jeanteur (Montpellier, France)
Nathalie Cartier (Paris, France)
Marie-Paule Lefranc (Montpellier, France)
John de Vos (Montpellier, France)

IURC team (Montpellier)
Sylvie Tuffery-Giraud
Christophe Beroud
Gwenaelle Collod-Beroud
Albertina de Sario (Montpellier, France)

Arnold Munnich (Paris, France)
Marc Delpech (Paris, France)
Philippe Amouyel (Lille, France)

HGM2010
HUGO ORGANISING COMMITTEE

Madeline Shee
Diana Hon
Ashton Quek
Hanny Dwiyanti Setiaji
Elspeth Bruford
Todd Taylor
Cathy Pole
Loo Sze Zhen

Inserm

Institut national
de la santé et de la recherche médicale

A Warm Welcome to Montpellier!

We are delighted to organise, on behalf of the Human Genome Organisation (HUGO), the next HGM meeting in Montpellier in May 2010.

Through an extraordinary international effort, scientists have sequenced the genome of plants, animals, microorganisms and humans. Comparative genomics opens new windows into the evolutionary mechanisms underlying genetic variation, adaptation and speciation. New sequencing and array technologies have revealed an unanticipated level of genomic structural variation, which has led to the discovery of DNA signatures for a growing number of diseases, while genome-wide association studies have begun identifying variants predisposing to common multifactorial diseases. Impressive progress has already been made in deciphering genetic, epigenetic and environmental contributions to disease risk.

We are living through an unprecedented era of progress in bioinformatics and biotechnology, and the "genomics revolution" is transforming health research. The new tools it provides will generate an entirely different type of health care, one that is predictive, preventive, participative and personalised ("P4 medicine"). We will see new ways of diagnosing diseases and new ways of thinking about disease prevention and health promotion. In terms of genomics-based medicine, knowing who we are will allow better targeting of cures.

We aim to bring together senior and junior investigators, postdoctoral and postgraduate researchers and the most famous world experts in genomics and in genetics to share basic and translational research as well as educational and social issues. In addition, updates on research will be highlighted for epigenetic, non-coding RNAs, pharmacogenetics, and therapy for genetic disorders. HGM2010 also aims to promote communication between participants to address some of the most pressing educational, ethical, legal and social questions raised by recent advances in genetic and genomic research.

The University of Montpellier is one of the oldest in the world, founded in 1160, and the Mediterranean climates as well as the reputation of hospitality and open-mindedness of the region are attractive to students (85,000 among a global population of 500,000). Montpellier, the "scientific city", is famous for research activity (5,000 researchers), particularly in life sciences, agronomics and chemistry. It is also the 5th largest University Hospital in France and number one for Medical Conferences.

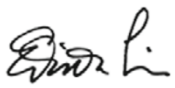
The Congress Centre (the Corum) is ideally situated in the attractive, pedestrianised town centre. Montpellier, capital of the Languedoc-Roussillon region, is a very pleasant place. Wander through its mediaeval streets lined with shops, bars and restaurants, visit one of the richest fine arts collections in Europe. The beaches of the Mediterranean are only 10 km away; you have the beautiful and famous Camargue and the stunning Millau bridge nearby and the region includes five of the most prestigious UNESCO world heritage sites.

Mireille Claustres
Chair, HGM2010

HGM2010 President's Welcome

I wish to welcome all delegates of the HGM 2010 meeting to Montpellier. The focus of this conference is on the next generation of genomic technologies and their impact on how we conceptualize medicine and how we conduct medical research. The precision and the comprehensiveness of the technologies permit discovery at orders of magnitude greater than few years ago, and we are now grappling with findings of high complexity. Thus, we move towards computational modeling and systems approaches in the analysis of genomic information. Montpellier is an intellectual center for the sciences and mathematics in France and in Europe and is the most appropriate venue to celebrate the impact of these advances.

The pace of change in genomic medicine has been breathtaking, and HUGO is also changing, reflecting both the nature of genomics sciences, and the interests of our membership. We are more global with greater representation from countries who are new entrants in genomic research, and we are more diverse in what we call our intellectual domain. Both the national origins of our key speakers and the representation of the delegates are indications of this change. In addition, we will host the **Chen Awards for Distinguished Academic Achievement in Human Genetic and Genomic Research** and **Chen New Investigator Award** . These awards were started by Professor Yuan-Tsong (Y-T) Chen and Mrs Alice Der-Shan Chen of Taiwan to recognize an Asian geneticist who has contributed to the development of biomedical sciences in the region. The two awardees are Dr. Yusuke Nakamura from the University of Tokyo for his pioneering work on human variation, and Dr. Huck Hui Ng from the Genome Institute of Singapore for his groundbreaking research in stem cell transcriptional genomics. Scientific globalization is alive and well at HGM 2010.



Edison T. Liu
HUGO President
Executive Director, Genome Institute of Singapore

SPONSORS

ACKNOWLEDGEMENTS

HGM 2010 has received financial support from the following organisations. We would like to extend our utmost appreciation to them for all of their support.

DIAMOND SPONSORS

Roche
life technologies
illumina
Affymetrix
Agilent Technologies

GOLD SPONSORS

NGFN National Genome Research Network
Springer
GenomeCanada
ImMunoGeneTics
Knome

SILVER SPONSORS

P&G
DNAgenotek
Helicos
PACIFIC BIOSCIENCES
Fluidigm

INSTITUTIONAL SPONSORS

la Région Languedoc Roussillon
Montpellier Agglomération
Inserm
aviesan
AFM
UM1
ARCHON GENOMICS XPRIZE
ENVOY Montpellier
CNRS
Genome Institute of Singapore

CORPORATE SPONSORS

Agency for Science, Technology and Research
Roche

GENERAL INFORMATION

About Montpellier

Montpellier is the 8th largest city in France and is the capital of the Languedoc-Roussillon region and the Hérault department, located in the scenic South of France. During the 19th century the city developed into an industrial centre. In the 1960s, its population grew dramatically after French settlers from Algeria were resettled in the city following Algeria's independence from France. In the 1980s and 1990s, the city underwent a number of major redevelopment projects, including the building of Le Corum (the HGM2010 venue) and especially the development of the Antigone District. Montpellier also hosts one of the oldest universities, the University of Montpellier which was founded in 1160. As a "scientific city" Montpellier is famous for its research activities and medical schools.

Currency: Euro
Location: Languedoc-Roussillon
Area: approximately 22 square miles / 57 square kilometers
Population: approximately 250,000
Language: French

Time Zone: GMT / UTC +1
Country dialing code: +33
Telephone area code: 04
Average daily Montpellier January temperature: 12 C / 54 F
Average daily Montpellier July temperature: 29 C / 84 F

Did You Know?

The city is situated on hilly ground 10 kilometres (6 mi) inland from the Mediterranean coast on the River Lez. The name of the city, which was originally *Monspessulanus*, is said to be derived from *mont pelé* (the naked hill, because the vegetation was poor), or *le mont de la colline* (the mount of the hill)

Montpellier is located 52 kilometres (32 mi) from Nîmes, 168 kilometres (104 mi) from Marseille, 248 kilometres (154 mi) from Toulouse, and 750 kilometres (466 mi) from the capital of France, Paris.

Montpellier's highest point is the Place du Peyrou, at an altitude of 57 m (187.01 ft). The city is built on two hills, Montpellier and Montpelliéret; thus some of its streets vary greatly in altitude. Some of the streets are also very old and narrow, giving an intimate feel.

Main sights

- The main focus point of the city is the Place de la Comédie.
- The Musée Fabre
- The Jardin des plantes de Montpellier – the oldest botanical garden in France, founded in 1593
- La Serre Amazonienne, an Amazonian greenhouse
- The fourteenth century Saint Pierre Cathedral
- The Porte du Peyrou, a triumphal arch
- The Saint Clément Aqueduct
- The Antigone District and other housing projects were designed by the architect Ricardo Bofill from Catalonia, Spain
- A number of *châteaux*, " follies", built by wealthy merchants surround the city

The Festival de Radio France et Montpellier is a summer festival of opera and music held in Montpellier. The music festival concentrates on classical music and jazz with about 150 events, including opera, concerts, films, and talks. Most of these events are free and are held in the historic courtyards of the city or the modern concert halls of Le Corum.

Your discovery of Montpellier wouldn't be complete without sampling the fine local cuisine. Your taste buds will help you appreciate the city from another angle.

Vibrant and active, Montpellier is a city that is simply alive and waiting to be discovered.

Source: http://en.wikipedia.org/wiki/Montpellier

GENERAL INFORMATION

HUGO Information Desk

The HUGO Information Desk is located at the entrance of the Exhibition area (Level 0 of Le Corum). Visit us for information about HUGO and its activities, initiatives (HUGO-WHO White Paper Services; HUGO-OECD Analytical Paper; HUGO Blog etc.), forthcoming meetings, and membership, as well as general information and services about HGM2010.

The HUGO Gene Nomenclature Committee (HGNC, www.genenames.org) will also be contactable at the HUGO Information Desk. Please drop by if you have any enquires about gene nomenclature.

HGM2010 Secretariat

Located in the Tiberiade Room (Level 0 of Le Corum), you can find free internet access and friendly volunteers who will assist you with on-site registration, payment and general enquiries. Speakers are welcomed to preview your presentations in the Tiberiade Room at any time, and must submit their slides here at least 12 hours before the start of their session.

On-Site Registration

Registration will take place on:

Tuesday 18 May 9.00 AM – 5.00 PM
Wednesday 19 May 8.30 AM – 5.00 PM

The conference bag containing the HGM2010 Programme/Abstract book and other essential information can be collected at the on-site registration counters - located in the Berlioz Foyer (Level 0 of Le Corum). From Thursday 20 May onwards, all registrant related matters should be directed to the HGM2010 Secretariat (Tiberiade Room – Level 0 of Le Corum)
Registration will take place:

Badges

Badges will be issued upon registration and must be worn at all times during HGM2010. Friendly Hosts/Hostesses will be manning each entry point to the scientific sessions and **ANYONE WITHOUT A BADGE WILL BE REFUSED ENTRY** . Replacement badges will incur a charge of €10 each. Your badge allows you to attend the Welcome Cocktail reception on Tuesday 18 May and the conference dinner on Thursday 20 May.

Badges are colour coded as follows:

Blue – Speakers and Session Chairs
Red – Participants
Green – Exhibitors
White – Single Day Pass Entrants
Black – Staff and Volunteers
Yellow – Students

Travel/Tourist Information

Please drop by the HUGO Information Desk and speak to our friendly representative from Voyages Kuoni SA Destination Management for information about excursions to the sea-side, nearby vineyards, the beautiful countryside, a culinary exploration or just a simple city tour.

Welcome Cocktail Reception

The Welcome Reception will be held in the Exhibition area (Berlioz Foyer – Le Corum of Level 0) in conjunction with the Inauguration of the Exhibition and poster presentations on Tuesday, 18 May. All HGM2010 registrants are welcome, so please come and view the posters and support our exhibitors and sponsors while enjoying a cocktail.

Food/Drinks

Complimentary tea, coffee, and biscuits will be provided in the morning and afternoon tea breaks in the Exhibition Area (Berlioz Foyer – Le Corum, Level 0). Delegates can purchase a range of reasonably priced drinks and snacks from the Brassarie (Le Corum, Level 1). Admittance to the Welcome Cocktail Reception (Tuesday 18 May) and Conference Dinner (Thursday 20 May) is inclusive in your registration fee. Extra Cocktail and Conference Dinner Tickets for non-registrants can be purchased from the HGM2010 Secretariat (Tiberiade Room – Level 0 of Le Corum) at €150 (exVAT 19.6%).

Poster Sessions

All posters must be put up on the poster boards by 4pm on Tuesday, 18 May. Each presenter is required to present their poster at the session specified by their poster number in the Poster Abstract listing in this book. Presenters not at their poster during the specified session will be ineligible for the Poster Prizes.

Poster Session I (W) -	Wednesday 19 May	11.35 AM – 1.05 PM
Poster Session II (T) -	Thursday 20 May	11.25 AM – 12.55 PM
Poster Session III (F) -	Friday 21 May	12.25 PM – 1.55 PM

Springer Poster Competition

HUGO is delighted that Springer, a major publisher of books and journals in the field of genomic medicine - including The HUGO Journal, the Official Journal of the Human Genome Organisation, will be sponsoring a poster competition at HGM2010. The competition will be judged by members of the HUGO Council and the HGM2010 International Scientific Committee.

1st prize: €500 plus one year's free subscription to three Springer Journals

2nd prize: €300 plus one year's free subscription to two Springer Journals

3rd prize: €100 plus one year's free subscription to The HUGO Journal

HUGO Gene Nomenclature Committee Nomenclature Poster Prize

The HUGO Gene Nomenclature Committee (HGNC) will again be awarding prizes to the authors of the two posters which contain the HIGHEST NUMBER of approved gene symbols. This will be judged by members of the HGNC team with a score of +1 for use of a correct symbol and -3 for an incorrect one. Each prize is a one year subscription to one Springer Journal of your choice and to The HUGO Journal; the HGNC would like to thank Springer for their sponsorship of this competition.

Chen Award 2010

HUGO wishes to thank Professor Yuan-Tsong Chen and Alice Der-Shan Chen of the Chens' Foundation for the honour of once more administering the Chen Award (2010). Please join us to celebrate the achievements and talents in Genomic Sciences in the Asia Pacific region at the Chen Award lectures (18 May and 20 May).

Distinguished Academic Achievement in Human Genetic and Genomic Research 2010

Professor Yusuke Nakamura (Japan)

Chen Young Investigator Award 2010

Dr. Ng Huck Hui (Singapore)

PROGRAMME SUMMARY

18 MAY 2010

TIME PROGRAMME		Venue
9.00 AM - 12.00 PM	HUGO-OECD Discussion (by invitation only)	Louisville Room
12.00 PM - 1.00 PM	Lunch Break	
1.00 PM - 2.00 PM	Press Conference	Barcelone Room
2.00 PM - 2.05 PM	Opening Address by Edison T Liu (HUGO)	Salle Pasteur
2.05 PM - 2.10 PM	Welcome Address by Mireille Claustres (INSERM)	Salle Pasteur
2.05 PM - 2.10 PM	Address by Dominique Daegelen (INSERM)	Salle Pasteur
2.15 PM - 4.15 PM	Plenary: Synthethic and Systems Genomics	Salle Pasteur
4.15 PM - 4.35 PM	Tea Break	Berlioz Foyer
4.35 PM - 6.35 PM	Plenary: Epigenetics in Development and Human Disease	Salle Pasteur
6.35 PM - 8.05 PM	Welcome Cocktail Reception / Poster Session Inauguration / Exhibition Inauguration	Berlioz Foyer

19 MAY 2010

TIME PROGRAMME		Venue
8.30 AM - 10.55 AM	Symposium: Biobanking	Salle Pasteur
	Symposium: Microbial Genomics and Metagenomics	Salle Einstein
9.30 AM - 9.55 AM	Tea Break	Berlioz Foyer
11.00 AM - 11.40 AM	Featured talk	Salle Pasteur
11.40 AM - 1.10 PM	Poster Session 1 & Exhibition	Berlioz Foyer
	Lunch Seminars	Sully 1 & 2
	HUGO Council Meeting	Louisville Room
1.10 PM – 3.10 PM	Workshop: Computational Biology, Computational Genomics	Salle Pasteur
	Workshop: Genetic and Genomic Databases	Salle Einstein
	Workshop: Genomics, Ethics, Law & Society	Sully 1
	Workshop: Genome Variation in Disease I	Sully 2
	Workshop: Genome Variation in Evolution	Sully 3
3.15 PM – 4.35 PM	Plenary: An Expanding RNA World	Salle Pasteur
4.35 PM - 4.50 PM	Tea Break	Berlioz Foyer
4.50 PM - 6.10 PM	Plenary: Gene Expression and Human Variation	Salle Pasteur

PROGRAMME SUMMARY

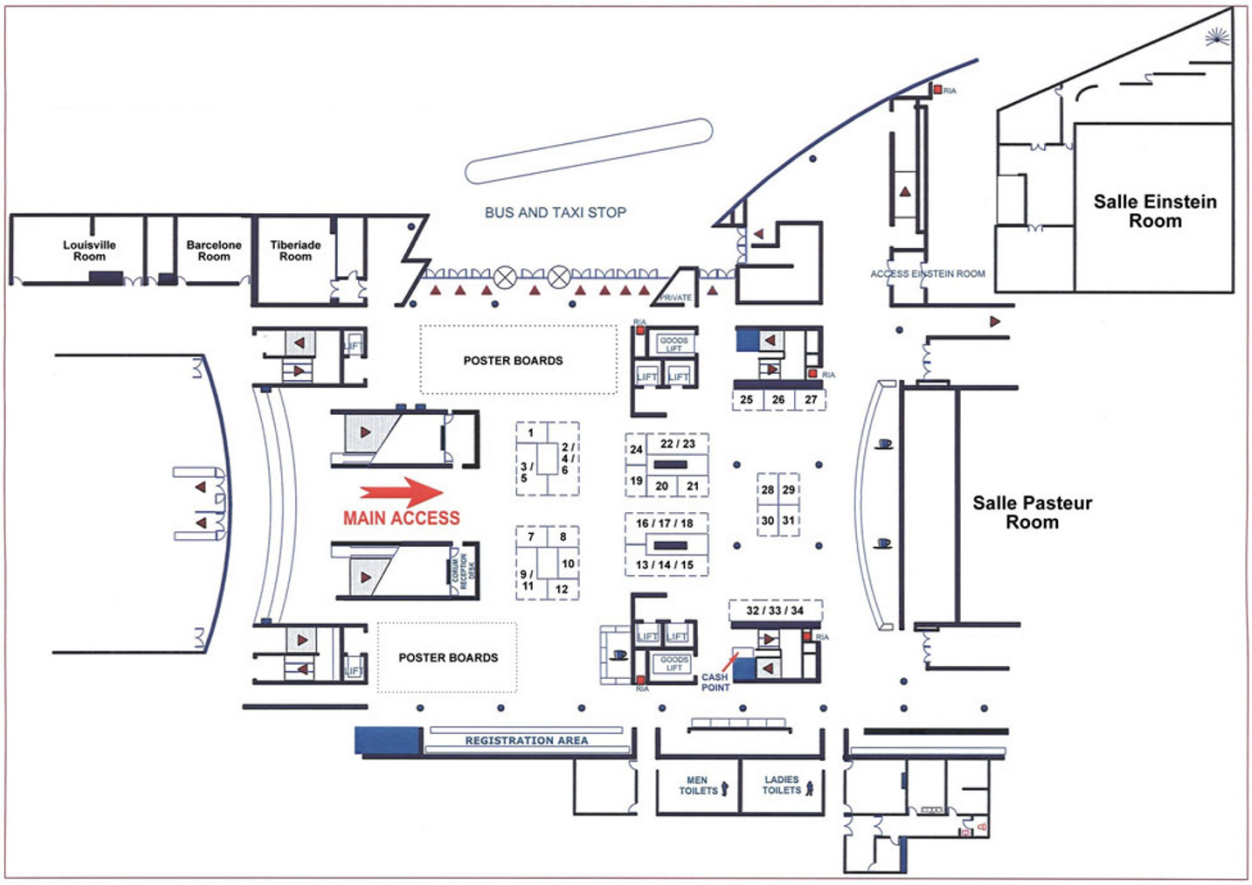
20 MAY 2010

TIME PROGRAMME		Venue
8.30 AM - 10.45 AM	Plenary: Human Genome Variation in Evolution and Disease	Salle Pasteur
9.50 AM - 10.05 AM	Tea Break	Berlioz Foyer
10.45 AM - 11.25 AM	Featured talk	Salle Pasteur
11.25 AM - 12.55 PM	Poster Session 2 & Exhibition	Berlioz Foyer
	Lunch Seminars	Sully 1 & 2 & Salle Pasteur
12.55 PM - 2.55 PM	Workshop: Cancer Genomics I	Salle Pasteur
	Workshop: Evolutionary Genomcis	Salle Einstein
	Workshop: Genetics of Health and Disorder	Sully 1
	Workshop: Genomics Medicine and Pharmacogenomics	Sully 2
	Workshop: Genetics and Genomics in Developing Countries	Sully 3
3.00 PM - 5.15 PM	Symposium: Quantitative Genomics (GWAS, QTL, Expression Cassettes)	Salle Pasteur
	Symposium: Next Generation Sequencing - 1000 Genome Project and Beyond	Salle Einstein
4.25 PM - 5.10 PM	Tea Break	Berlioz Foyer
5.20 PM - 6.00 PM	Featured talk	Salle Pasteur
6.30 PM - 7.00 PM	Bus transfer to Conference Dinner venue	
7.00 PM - 9.00 PM	Conference Dinner	

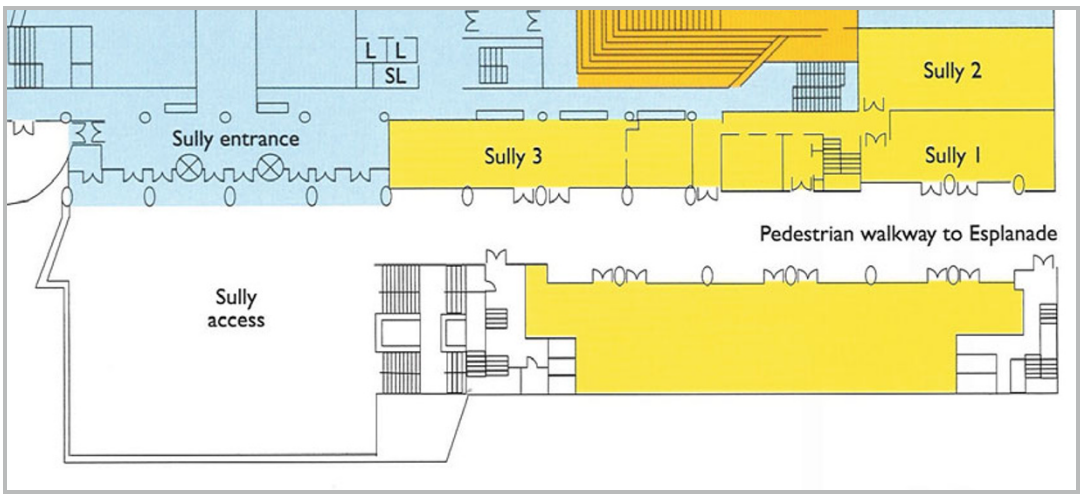
21 MAY 2010

TIME PROGRAMME		Venue
8.30 AM – 10.25 AM	Symposium: Pathways Networks and System Biology	Salle Pasteur
	Symposium: Nutrigenetics and Nutrigenomics	Salle Einstein
9.30 AM - 9.45 AM	Tea Break	Berlioz Foyer
10.30 AM - 12.00 PM	Workshop: Cancer Genomics II	Salle Pasteur
	Workshop: Genome Variation in Disease II	Salle Einstein
	Workshop: Epigenetics and microRNA-mediated gene regulation	Sully 1
	Workshop: Structural Variations and Chromosomal Aberrations	Sully 2
	Workshop: Functional Genomics	Sully 3
12.00 PM - 1.30 PM	Poster Session 3 & Exhibition	Berlioz Foyer
	Lunch Seminars	Sully 1 & 2
1.30 PM - 5.00 PM	HUGO Forum - Sequencing Futures: Sequencing Technologies Now and the Future	Salle Pasteur
3.30 PM – 4.00 PM	Tea Break	Berlioz Foyer
5.00 PM - 5.15 PM	Closing Ceremony	Salle Pasteur

VENUE FLOORPLAN (LE CORUM - MONTPELLIER, FRANCE)



Level 0 – Registration Area, Secretariat Room (Tiberaide Room), Salle Einsten and Salle Pasteur



Level 1 – Sully 1, Sully 2 and Sully 3

FULL PROGRAMME

18 May 2010

Time	Programme	
9.00 AM - 12.00 PM	HUGO-OECD Discussion (By invitation only)	
12.00 PM - 1.00 PM	*Lunch Break*	
1.00 PM - 2.00 PM	Press Conference	
2.00 PM - 2.05 PM	Opening Address by **Edison T Liu** (HUGO)	
2.05 PM - 2.10 PM	Welcome Address by **Mireille Claustres** (INSERM)	
2.10 PM - 2.15 PM	Address by **Dominique Daegelen** (INSERM)	
2.15 PM - 4.15 PM	Synthethic and Systems Genomics Chaired by **Edison T Liu** Venue: Salle Pasteur	
	Drew Endy Work towards an 8-bit engineered genetic combinatorial counter	
	Christina D Smolke Programming cellular behavior with RNA controllers	**HGM 2010 Registration**
	Yusuke Nakamura (Chen Award Lecture) DNA variations in human and medical genetics: My 25-year experience	
4.15 PM - 4.35 PM	*Tea Break*	
4.35 PM - 6.35 PM	Epigenetics in Development and Human Disease Chaired by **Veronica van Heyningen** Venue: Salle Pasteur	
	Robert Feil Epigenetic mechanisms in genomic imprinting	
	Anne C Ferguson-Smith Genomic imprinting and the epigenetic control of mammalian development and human disease	
	Stephen Baylin Evolution and implications of the cancer epigenome	
6.35 PM - 8.05 PM	**Welcome Cocktail Reception** together with **Poster Session Inauguration** and **Exhibition Inauguration**	

FULL PROGRAMME

19 May 2010

Time	Programme	Time	Programme
8.30 AM - 10.55 AM	Biobanking Chaired by **Anthony Brookes** / **Richard Cotton** Venue: Salle Pasteur	**8.30 AM - 10.55 AM**	Microbial Genomics and Metagenomics Chaired by **Todd Taylor** / **Martin Hibberd** Venue: Salle Einstein
	Bartha Maria Knoppers The Public Population Project in Genomics : An "Ethical" Proof of Concept		**Martin Hibberd** Dengue virus genomics, a window into evolution, virulence and transmission patterns
	David B Goldstein Rare and common variants in human disease		**Peer Bork** Comparative (meta) genomics: Lessons from a tiny bacterium and a complex microbial community
09.40 AM - 09.55 AM	*Tea Break*	9.30 AM - 9.45 AM	*Tea Break*
	David Cox Advancing biobanking through public-private collaboration		**Stanislav Dusko Ehrlich** Characterization of the human "other" genome by the MetaHIT Consortium
	Jennifer Harris Biobank harmonization for health		**Edward Rubin** Assembly of genes and genomes from bio-mass degrading cow rumen
10.55 AM - 11.00 AM	*Room Change*		
11.00 AM - 11.40 AM	**Charles N Rotimi** (Featured Talk) Genetic diversity and health: Opportunities and challenges presented by African genomes Venue: Salle Pasteur		
11.40 AM - 1.10 PM	*Lunch & Exhibition* | Life Technologies Venue: Sully 1 | Affymetrix Venue: Sully 2 | HUGO Council Meeting (by invitation only) Venue: Louisville | Poster Session 1 Venue: Berlioz Foyer		
1.10 PM - 3.10 PM	Computational Biology, Computational Genomics Chaired by **Doron Lancet** / **John de Vos** Venue: Salle Pasteur | Genetic and Genomic Databases Chaired **by Yoshiyuki Sakaki** / **Marie-Paule LeFranc** Venue: Salle Einstein | Genomics , Ethics, Law & Society Chaired by **Ruth Chadwick** / **Benjamin Capps** Venue: Sully 1 | Genome Variation in Disease 1 Chaired by **Christian Jorgensen** / **Elspeth Bruford** Venue: Sully 2 | Genome Variation in Evolution Chaired by **Martin Hibberd** / **Agnes Rotig** Venue: Sully 3		
3.10 PM - 3.15 PM	*Room Change*		
3.15 PM - 4.35 PM	An Expanding RNA World Chaired by **Julian Venables** / **Sumio Sugano** Venue: Salle Pasteur		
	Gil Ast Importance of alternative splicing in transcriptomic diversity and disease		
	Yoshihide Hayashizaki Transcriptome analysis — A way to illuminate the genome network		
4.35 PM - 4.50 PM	*Tea Break*		
4.50 PM - 6.10 PM	Gene Expression and Human Variation Chaired by **Edison T Liu** / **Doron Lancet** Venue: Salle Pasteur		
	Michael Snyder Transcription binding variation in eucaryotes		
	Howard McLeod Using the genome to guide therapy		
6.10 PM	End of Day		
6.30 PM - 8.00 PM	Free and Easy | HUGO Ethics Committee Meeting (by invitation only) | PASNP Phase II Planning Meeting (by invitation only)		

FULL PROGRAMME

20 May 2010

Time	Programme	Time	Programme
8.30 AM - 10.45 AM	Human Genome Variation in Evolution and Disease Chaired by **Stylianos E Antonarakis** / **Richard Cotton** Venue: Salle Pasteur		
	Mark McCarthy Type 2 diabetes: Susceptibility variants across the allele frequency spectrum		
	Stephen W Scherer Clinical context of copy number and structural variation in the human genome		
9.50 AM - 10.05 AM	*Tea Break*		
	Felix Jin Li Identification of copy number variation hotspots in human populations		
10.45 AM - 11.25 AM	**Ng Huck Hui** (Chen Award Lecture) Deciphering and reconstructing the embryonic stem cell transcriptional regulatory network Venue: Salle Pasteur		
11.25 AM - 12.55 PM	*Lunch & Exhibition* Roche Venue: Sully 1 Agilent Technologies Venue: Sully 2 President's Luncheon (by invitation only) Poster Session 2 Venue: Berlioz Foyer **Larry Kedes -** Archon Genomics XPRIZE (Featured Talk) Swifter, Lower, Stronger: the $10 Million Olympic Medal for Next Generation Sequencing Venue: Salle Pasteur		
12.55 PM - 2.55 PM	Cancer Genomics I Chaired by **Philippe Broet** / **Bernard Klein** Venue: Salle Pasteur Evolutionary Genomics Chaired by **Harris A Lewis** / **Christian Hamel** Venue: Salle Einstein Genetics of Health and Disorder Chaired by **Stephen W Scherer** / **Veronica Van Heyningen** Venue: Sully 1 Genomic Medicine and Pharmacogenomics Chaired by **Lance Miller** / **Marc Delpech** Venue: Sully 2 Genetics and Genomics in Developing Countries Chaired by **Dhavendra Kumar** / **Michele Ramsay** Venue: Sully 3		
2.55 PM - 3.00 PM	*Room Change*		
3.00 PM - 5.15 PM	Quantitative Genomics (GWAS, QTL, Expression cassettes) Chaired by **Martin Vingron** / **Yoshihide Hayashizaki** Venue: Salle Pasteur	3.00 PM - 5.15 PM	Next Generation Sequencing - 1000 Genome Project and Beyond Chaired by **Ruan Yijun** Venue: Salle Einstein
	Emmanouil Dermitzakis Cellular genetics and genomics		**Henry Yang Huanming** Human Genomics - From HGP to 1000 genomes and beyond
	Frank Johannes Complex trait dynamics following epigenomic perturbation		**Ruan Yijun** Chromatin Interaction and Transcription Regulation
	Augustine Kong Some thoughts on genetics studies in the near future		**Gilean McVean** The landscape of human genetic variation as viewed from the 1000 Genomes Project
4.25 PM - 4.40 PM	*Tea Break*	4.55 PM - 5.10 PM	*Tea Break*
	Martin Vingron Transcription factor binding site analysis and histone modifications point at two classes of promoters		
5.15 PM - 5.20 PM	*Room Change*		
5.20 PM - 6.00 PM	**Manel Esteller** (Featured Talk) Cancer Epigenetics: From DNA methylation to microRNAs Venue: Salle Pasteur		
6.00 PM	End of Day		
6.30 PM - 7.00 PM	Transfer to Dinner Venue		
7.00 PM - 9. 00PM	Conference Dinner		

FULL PROGRAMME

21 May 2010

<table>
<tr><th>Time</th><th>Programme</th><th>Time</th><th>Programme</th></tr>
<tr><td>8.30 AM - 10.25 AM</td><td>Pathways Networks and System Biology
Chaired by Gerardo Jimenez-Sanchez / Laurent Journot
Venue: Salle Pasteur</td><td>8.30 AM - 10.25 AM</td><td>Nutrigenetics and Nutrigenomics
Chaired by Jean-Louis Gueant / Samir K Brahmachari
Venue: Salle Einstein</td></tr>
<tr><td></td><td>Hiroki R Ueda
Systems Biology of Mammalian Circadian Clocks</td><td></td><td>Susan E Ozanne
Mechanisms linking poor early nutrition and later risk of type 2 diabetes</td></tr>
<tr><td></td><td>Eran Segal
Transcriptional Lego: Predictable control of gene expression by manipulating promoter building blocks</td><td></td><td>Harris A Lewin
Nutrition: Will livestock Lead the way?</td></tr>
<tr><td>9.30 AM - 9.45 AM</td><td colspan="3">Tea Break</td></tr>
<tr><td></td><td>Luis Serrano
Systems biology analysis of the EGF MAPK pathway: From structures to data integration</td><td></td><td>Jean-Christophe Glaszmann
Genomics for food improvement</td></tr>
<tr><td>10.25 AM - 10.30 AM</td><td colspan="3">Room Change</td></tr>
<tr><td>10.30 AM - 12.00 NN</td><td colspan="3"><table><tr><td>Cancer Genomics II
Chaired by Wei Chia-Lin / Albertina de Sario
Venue: Salle Pasteur</td><td>Genome Variation in Disease II
Chaired by Y T Chen / Valere Cacheux-Rataboul
Venue: Salle Einstein</td><td>Epigenetics and microRNA-mediated gene regulation
Chaired by Lim Bing / Jean-Marie Blanchard
Venue: Sully 1</td><td>Structural Variations and Chromosomal Aberrations
Chaired by Lawrence Stanton / Richard Redon
Venue: Sully 2</td><td>Functional Genomics
Chaired by Samir K Brahmachari / Alain Bucheton
Venue: Sully 3</td></tr></table></td></tr>
<tr><td>12.00 NN - 1.30 PM</td><td colspan="3"><table><tr><td>Lunch</td><td>Illumina
Venue: Sully 1</td><td>Poster Session 3
Venue: Berlioz Foyer</td><td></td></tr></table></td></tr>
<tr><td>1.30 PM - 5.00 PM</td><td colspan="3">HUGO Forum - Sequencing Futures: Sequencing Technologies Now and the Future -
Chaired by Herve Thoreau / Wei Chia-Lin
Venue: Salle Pasteur</td></tr>
<tr><td></td><td colspan="3">Radoje Drmanac
The path to affordable human genome sequencing for large-scale studies of genetic diseases</td></tr>
<tr><td></td><td colspan="3">Nathaniel Pearson
Deciphering the oracle's words: Analysis tools for human whole genomes and exomes</td></tr>
<tr><td></td><td colspan="3">Marcus Droege
Re-sequencing of the human genome using the 454 Genome Sequencer systems FLX & Junior</td></tr>
<tr><td></td><td colspan="3">Patrice M Milos
An unbiased, quantitative view of genome biology only possible with Helicos Single Molecule Sequencing</td></tr>
<tr><td></td><td colspan="3">Francisco M De La Vega
The road to 99.999% accuracy single molecule sequencing</td></tr>
<tr><td></td><td colspan="3">Stephen Turner
Applications of SMRT™ sequencing outside the performance envelope of first and second generation sequencing</td></tr>
<tr><td></td><td colspan="3">Mostafa Ronaghi
Current and future outlook of genomic technologies</td></tr>
<tr><td></td><td colspan="3">*Tea Break (3.30 PM – 4.00 PM)</td></tr>
<tr><td>5.00 PM - 5.15 PM</td><td colspan="3">Closing Ceremony (Springer Poster Awards, HGNC Nomenclature Awards and Announcement of future HGM)</td></tr>
</table>

The goal of National Research Program for Genomic Medicine (NRPGM) is to undertake basic research and technology development of the above 6 research project areas, facilitated and supported by the well-established national core facilities covering the 5 major fields. Since its start in 2002, significant achievements have been attained on identification of disease related genes, elucidation of the function and etiology; development of disease animal models; clinical applications of the biomarkers on disease prevention, diagnostics, and therapeutics, as well as drug and target discovery. NRPGM members eagerly seek opportunities to collaborate with Taiwan and worldwide partners from industries, pharmaceutical companies, and academic institutes in research, development, and applications of these areas. Please see NRPGM website (**http://nrpgm.sinica.edu.tw**) for information about the program, project applications, collaborations, achievements, patents, etc.

NRPGM Core facilites welcome Taiwan and international researchers to collaborate and to apply for use of their services. For collaborations, service items, charges, payment methods, and inquiry, please see NRPGM website (**http://nrpgm.sinica.edu.tw**).

Sponsoring Agencies: NSC DOH MOEA

PANEL OF INTERNATIONAL SPEAKERS

Plenary: Synthetic and Systems Genomics (18 May 2010)

Drew Endy

Director, BIOFAB
President, The BioBricks Foundation
Assistant Professor, Stanford Bioengineering

Title: Work towards an 8-bit engineered genetic combinatorial counter

Modest information storage systems implemented inside living cells would enable new approaches for researching and controlling biological processes such as development, cancer, and aging. Our current capacity to engineer and operate genetically encoded information storage systems is quite limited. Specific limitations include the lack of sufficient molecular components to build with, rules of composition supporting device and system integration, an understanding for how to implement reliable behavior given thermal noise at the molecular scale, and an understanding for how to engineer reliable systems that evolve. I'll introduce applications of genetic information storage systems, review past and current accomplishments from the field, introduce our experimental work on composable set/reset latches built with serine recombinases, and our theoretical work on a framework supporting the engineering of higher-order information storage systems. Given that an 8-bit counter likely requires the successful integration of at least 10-fold more components than any existing engineered genetic system, I'll also discuss the current state of, and needs regarding, foundational tools supporting genetic engineering.

Biosketch

Drew Endy grew up in Pennsylvania earning degrees in civil and environmental engineering at Lehigh. Following a summer internship at AMTRAK, he earned a PhD in biochemical engineering at Dartmouth. He then studied microbiology and genetics as a postdoc at UT Austin and UW Madison. He worked with Roger Brent and Sydney Brenner at the Molecular Sciences Institute in Berkeley. Drew joined MIT as a Fellow in Biology and Biological Engineering in 2002, later joining the faculty in the newly created Department of Biological Engineering. At MIT, Drew developed and taught 5 courses in helping to launch MIT's new undergraduate major in BE. His lab published a systematic redesign ("refactoring") of a natural organism's genome, successfully implementing over 600 simultaneous genetic changes while maintaining organism viability, a comprehensive "datasheet" for an abstracted genetic device, methods and a reference standard for measuring gene expression inside living cells, and a contrarian observation that incredibly tiny and "molecularly noisy" bacteria appear to implement reliable behavior. Drew has helped found a few companies, the iGEM competition, the synthetic biology (SB#.0) conference series, and the BioBricks Foundation, a non-profit that is working to create open technology platforms supporting the next generation of biotech. His group at Stanford Bioengineering is working to implement scaleable genetic memory systems, for storing and controlling modest amounts of information inside living systems. Esquire named Endy one of the 75 most influential people of the 21st century.

Christina D. Smolke

Assistant Professor of Bioengineering
Bioengineering Department
Stanford University

Title: Programming cellular behavior with RNA controllers

Cellular behavior is encoded and controlled by complex genetic networks. Synthetic genetic devices that interface with native pathways can be used to change natural networks to implement new forms of control and behavior. Significant recent work on the engineering of synthetic gene networks has been limited by an inability to interface with native networks and components. To overcome these limitations, we have developed RNA control devices that process and

transmit molecular signals that are received by integrated sensor domains to targeted protein level outputs, linking computation and logic to gene expression and thus cellular behavior in mammalian cells. The modularity inherent in our device design supports the rational assembly of these RNA controllers from independent components exhibiting basic functions and the extension to more sophisticated information processing schemes, highlighting the potential of synthetic biology strategies to support the rapid engineering of cellular behavior. Coupled with technologies that enable the *de novo* generation of new RNA sensor components, RNA devices allow researchers to construct various user-programmed information processing operations in living systems. The application of these molecular devices to developing new disease treatment strategies such as targeted molecular and cellular therapeutics will be discussed.

Biosketch

Christina Smolke has been an Assistant Professor in the Department of Bioengineering at Stanford University since January 2009. Prior to that she was an Assistant Professor in the Division of Chemistry and Chemical Engineering at the California Institute of Technology. She graduated with a B.S. in Chemical Engineering and a minor in Biology from the University of Southern California in 1997. She conducted her graduate training as a National Science Foundation Fellow in the Chemical Engineering Department at the University of California at Berkeley and earned her Ph.D. in 2001. Christina conducted her postdoctoral training as a National Institutes of Health Fellow in Cell Biology at UC Berkeley and began her position at Caltech in 2003. She has pioneered a research program focused on developing foundational technologies for the design and construction of engineered ligand-responsive RNA-based regulatory molecules, their integration into cellular computation and signal integration strategies, and their reliable implementation into diverse cellular engineering applications. These technologies are resulting in scaleable platforms for the construction of molecules that allow regulation of targeted gene expression levels in response to user-specified endogenous and exogenous molecular ligands. Christina's innovative research program has been recognized with the receipt of a National Science Foundation CAREER Award, a Beckman Young Investigator Award, an Alfred P. Sloan Research Fellowship, World Technology Network Award in Biotechnology, and the listing of Christina as one of Technology Review's Top 100 Young Innovators in the World.

Chen Award Lecture (18 May 2010)

Yusuke Nakamura

Director and Professor
The University of Tokyo

Title: DNA variations in human and medical genetics: My 25-year experience

DNA variations have contributed enormously to the fields of human and medical science in the last two-three decades, especially for identification of genes responsible or susceptible to various diseases, those involved in cancer and those associated with efficacy or adverse reactions of various drugs. The types of genetic variation used in these studies have changed in the past 25 years and can be classified into five major classes: RFLP (restriction fragment length polymorphism), VNTR (variable number of tandem repeat), STR (short tandem repeat or microsatellite), SNP (single-nucleotide polymorphism) and CNV (copy-number variation). Genetic linkage analysis using these tools mapped and discovered genes responsible for hundreds of hereditary diseases. Furthermore, construction of the international SNP database and recent development of high-throughput SNP typing platforms enabled us to perform genome-wide association studies, which have identified genes (or genetic variations) susceptible to common diseases or those associated with drug responses. Genome-wide sequencing of individual DNAs is gaining immense scope. I would like to summarize the history of polymorphic DNA markers and their contribution to the genetic analysis of both rare hereditary diseases and common diseases, as well as recent advances in pharmacogenetics, including our contribution to these areas.

Biosketch

Prof. Nakamura has been contributing to genomic medicine and also cancer research fields for more than two decades. He was one of pioneers in developing and applying genetic polymorphic markers (VNTR and SNP) in the medical genomics field. DNA polymorphic markers developed and mapped by his groups have contributed to map and clone genes responsible to hereditary diseases, those susceptible to common diseases, those related to drug response, and those involved in cancer. His groups in Cancer Institute, University of Tokyo, Osaka University and RIKEN SNP Research Center (present name is RIKEN Center for Genomic Medicine) in Japan have isolated dozens of genes of medical importance by genomics approach. In

addition, RIKEN SNP Research Center led by him participated in the International HapMap Consortium and made the largest contribution in the Phase 1 HapMap project published in Nature (2005). His contribution in human genetics and cancer research fields can be measured by his publication of more than 1,000 articles, including 32 articles in American Journal of Human Genetics, 107 in Cancer Research,15 in Nature, 36 in Nature Genetics, and 11 in Science, that have been cited for more than 65,000 times in total.

Plenary: Epigenetics in Development and Human Disease (18 May 2010)

Robert Feil

Senior Scientist, CNRS
Group Leader, The Institute of Molecular Genetics

Title: Epigenetic mechanisms in genomic imprinting

In contrast to other groups of animals, in placental mammals the maternally and paternally inherited genomes are functionally non-equivalent. This functional asymmetry between the parental genomes is a consequence of differential marking of the genomic DNA in the egg versus the sperm. These DNA methylation 'imprints' on the parental genomes persist in the developing embryo, and convey the allelic expression of genes from either their maternal or their paternal copy. Some hundred genes are controlled by this epigenetic phenomenon called 'genomic imprinting'. Many of these play important roles in foetal development and growth, particularly of the placenta, others influence behaviour. Pathological perturbation of genomic imprinting gives rise to growth-related and behavioural diseases in humans. Environmental stress can readily perturb imprints as well, and this may have long-lasting phenotypic consequences. Different ART technologies, for instance, seem to increase the frequency of the foetal overgrowth syndrome Beckwith-Wiedemann Syndrome (BWS) and other imprinting-related disorders.

I will discuss our current understanding of how DNA methylation imprints become established during spermatogenesis and oogenesis, and how they are maintained during development. Imprints are targeted to specific regulatory sequences that control imprinted gene expression, the so-called 'imprinting control regions' (ICRs). After fertilisation, these imprints are maintained in all the somatic cells and tissues. Perturbation of this maintenance process gives rise to altered DNA methylation at ICRs, and consequently, imprinting-related diseases. Our group is particularly interested in the organisation of chromatin at ICRs and to which extent histone modifications contribute to imprint establishment and maintenance. We also explore the importance of histone methylation in the tissue-specific regulation of imprinted gene expression. I will present several examples of how lysine methylation on histone H3 controls imprinted genes in the placenta, and in brain, and will describe the enzymatic machineries involved.

Biosketch

We are interested in the regulation of genomic imprinting in mammals, and in how this epigenetic mechanism influences embryonic and extra-embryonic development. Our projects aim at unravelling the role of chromatin modifications and DNA methylation, and investigate the pathological deregulation of imprinting in human diseases and as a consequence of in vitro manipulation.

Anne C Ferguson-Smith

Professor of Developmental Genetics
Department of Physiology Development and Neuroscience
University of Cambridge

Title: Genomic imprinting and the epigenetic control of mammalian development and human disease

Epigenetic modifications play key roles in chromosome architecture and integrity, chromatin structure and gene activity and repression. One of the key challenges in contemporary genome biology is to understand the relationship between the epigenome and the underlying DNA sequence with which it interacts. Genomic imprinting is a normal epigenetic regulatory process causing genes to be expressed from only one of the two parental chromosome homologues according to their parental origin. Imprinted genes function prenatally and postnatally in a range of developmental pathways controlling normal organogenesis, growth and metabolism. Perturbations of genomic imprinting are responsible for several disorders exhibiting parental-origin effects in their patterns of inheritance. Abnormal imprinting has also been implicated in a wide range of different cancers. More recently, genome-wide association studies that have considered the parent-of-origin of variants have identified novel associations between imprinted domains and complex diseases.

The mouse is a useful model for the analysis of defective imprinting mechanisms and the phenotypic contributions of genes whose dosage is perturbed in imprinting disorders. Recent advances have included the identification of mutations in trans-acting factors in mouse and human that are required for the stable maintenance of epigenetic states at multiple imprinted domains, and whose further analysis is likely to provide novel insights into molecules that integrate genomic with epigenomic information. In recent years, data have emerged identifying discordance of imprinting between mouse and man, polymorphic imprinting between different individuals and tissue and cell-specific imprinting within individuals. This suggests that epigenetic control modulating mono-allelic and biallelic expression at imprinted loci might be a dynamic process with the potential to act as a regulatory mechanism controlling gene dosage in different developmental contexts. This potential plasticity of epigenetic control at imprinted domains may contribute to their vulnerability, and help contribute to our understanding of the mechanisms and consequences of epimutation of the genome in a wider context.

Biosketch

The Ferguson-Smith team is interested in the molecular events governing mammalian development, and in understanding situations where normal developmental processes have been disturbed. In particular, their integrated programme of research is directed towards investigating the function, regulation and evolution of imprinted genes and also, in using genome imprinting as a model to understand the relationship between genomic organization and the epigenetic control of genome function. Current research focuses on three main themes: (1) Stem cells and the epigenetic programme in vitro and in vivo, (2) Functional epigenomics including the roles of small and large non-coding RNAs, repetitive sequences and cis-acting regulatory features, (3) Developmental and physiological functions of imprinted genes in health and disease and the extent to which this is influenced by the environmental modulation of epigenetic states. Genomic imprinting is a remarkable normal process that causes some genes to be expressed from maternally inherited chromosomes and others from paternally inherited chromosomes. This means that the egg and sperm contribute unequal functions to the developing conceptus through the parental-origin specific expression of imprinted genes. In mouse and man, disorders can arise when the dosage of imprinted genes is altered through imbalances in the parental-origin of particular chromosomes, by mutations in the single active allele or by mutations affecting the epigenetic process controlling imprinting. The team is investigating the developmental consequences of altering the dosage of imprinted genes on the epigenetic regulation of prenatal growth and development. Currently they are particularly interested in exploring of the consequences altering the dosage of imprinted genes in postnatal processes - notably in the regulation of neurogenesis and metabolism. This functional analysis combined with comparative analysis in different mammalian species, is contributing to our understanding of the evolution of genomic imprinting.

Genome function is regulated by epigenetic modifications that can affect a number of processes including chromosome architecture and integrity, chromatin structure and gene expression. During multiple stages in the lifetime of the organism these modifications, including DNA methylation and histone modifications, provide a dynamic, heritable and reversible method to affect genome function without changing the DNA sequence. Large and small non-coding RNAs are also involved. It is well-established that all these epigenetic modifications regulate the parental-origin specific expression of imprinted genes. Most imprinted genes identified to date are located in clusters of maternal and paternally expressed alleles controlled by both short and long range cis-acting epigenetic regulatory features. Imprinted domains are therefore an excellent in vivo model system to investigate the role of epigenetic modifications in genome function. Anne Ferguson-Smith obtained her PhD from Yale University. She commenced her studies on genomic imprinting during postdoctoral work first at The Babraham Institute and then at the Wellcome/CRC Institute of Cancer and Developmental Biology in Cambridge. She contributed to the identification of the first imprinted genes and determined the value of the mouse as a model for studying imprinted disorders in human. During this time, she also identified a role for DNA methylation in imprinting control. Anne subsequently obtained a faculty position at the University of Cambridge where her team identified and characterized a novel large imprinted domain (the Dlk1-Dio3 cluster) implicated in a wide range of disease processes and provided novel insights into our understanding of imprinting function, regulation and evolution. The team continues to pioneer the application of imprinting in a wider context as a model for understanding the epigenetic control of mammalian genome function. She is currently Professor of Developmental Genetics in the Department of Physiology Development and Neuroscience at the University of Cambridge and an A*-star Adjunct Research Director at the Singapore Institute for Clinical Sciences.

Stephen Baylin

Professor of Oncology
The Johns Hopkins University School of Medicine
McKusick-Nathans Institute of Genetic Medicine

Title: Evolution and implications of the cancer epigenome

We now increasingly realize from work over the past decade, especially, that cancer is a disease of epigenetic as well as genetic, abnormalities. Thus, changes in heritable gene expression patterns dictated by other than alterations in the primary base sequence of DNA, are more frequent than genetic changes in cancer, and critically important. For example, every individual patient's tumor, for all cancer types studied to date, harbors hundreds of genes which are transcriptionally silenced by the abnormal addition of an epigenetic regulatory modification, DNA methylation, to the proximal promoter region. From such observations we may now talk about not only the cancer genome but the cancer epigenome and efforts are now expanding to jointly study both.

One of the most important roles of epigenetic abnormalities in cancer may be a response to chronic cell stress that helps to underpin abnormal cell expansion and survival in the earliest phases of tumorigenesis. These events provide the cell population "substrate" for enhancing the likelihood that subsequent genetic and epigenetic events can foster subsequent progression events towards invasive disease and metastases. Concrete examples of this may be found in the pre-invasive promoter DNA methylation, and associated transcriptional silencing and loss of gene function, for genes controlling virtually every key signaling pathway fundamental for tumorigenesis including cell cycle and control of stem/progenitor cell regulation (p16, p15), the Wnt pathway (APC, SOX17), DNA repair (MLH1), cell adhesion and invasion (E-cadherin), and apoptosis (DAP-kinase). The keys to the molecular events which may drive the above abnormalities may be emerging from links between epigenetic control of gene expression in embryogenesis and the above abnormal expansion of progenitor cell populations. This may involve a molecular progression which abnormally locks in a normally "plastic" chromatin regulation of gene expression which maintains the state of embryonic stem (ES)/progenitor cells. In turn, this molecular progression may promote abnormal progenitor cell compartment survival and impede normal lineage commitment and maturation of involved cells and result in a primitive state of abnormally expanding cells in early cancer evolution.

The translational implications of understanding the cancer epigenome are potentially profound. For example, promising use of abnormal gene promoter DNA methylation as a biomarker system for cancer risk assessment, early cancer diagnosis, molecular staging of tumors, and monitoring of drug sensitivities is distinctly emerging. The possibility of"epigenetic therapy" for cancer is increasingly being pursued and particularly the concept of reversal of epigenetic gene silencing is being studied as a molecular target strategy for cancer therapeutics.

Biosketch

Scientific Summary: Our research has contributed heavily to the concept that epigenetically mediated loss of gene function is a major player in the progression of human cancer. This process, for which aberrant gene promoter hypermethylation is a signature and a component of aberrant loss of transcription for involved genes, is now known to be an alternative to coding region mutations for loss of function of over half the classic tumor suppressor genes and for a growing list of candidate tumor suppressor genes in virtually every type of human cancer. We are attempting to understand the abnormalities of chromatin and methylation assembly that may account for the appearance of these epigenetic abnormalities during tumor development and how they mediate the transcriptional repression. We are learning, in this regard, that an interaction between the DNA methylation, histone de-acetylase (HDAC), and histone methylating enzymes mediates the transcriptional silencing. We have also discovered that the enzymes that catalyze DNA methylation, the DNA methyltransferases (DNMT's), are more complex than previously thought and can both inhibit transcription, and interact with HDAC's, independent of mediating the methylation. In collaboration with the Vogelstein-Kinzler lab, we have identified that an interaction between DNMT's is required, in colon cancer cells, to maintain the abnormal promoter methylation and silencing of important tumor suppressor genes. All of these studies are giving us a much more complete picture of the machinery that mediates aberrant promoter methylation in cancer. They, also, are contributing to the translational goal of targeting reversal of abnormal gene silencing as a cancer prevention and/or therapy strategy.

General Summary: The Baylin laboratory, in close collaboration with the Herman laboratory, seeks to uncover the molecular basis for a change in the DNA of cancer cells which causes impaired function of important regulatory genes which are necessary to prevent the formation of tumors. Many such losses of gene function in cancer are caused by gene mutations, which cause the cell to produce defective proteins from these genes. These genetic changes are permanent in the DNA and cannot be reversed. The change studied by the Baylin and Herman labs is an alternative way in which cancer cells inactivate genes by adding methyl groups to the start regions of genes and this is associated with a "silencing" of the gene. This process prevents the gene from making any protein and, over the past several years, is being recognized as a very frequent abnormality in all types of human cancer. Unlike DNA mutations, DNA methylation abnormalities are reversible by drugs in a laboratory setting and this reversal allows cancer cells to reactivate the silenced genes and produce normal protein. Understanding how the abnormal DNA methylation arises in cancer cells, and how this change leads to silencing of genes, is extremely important for enriching the possibility of reversing this process as a strategy to prevent and/or treat cancer. Clinical trials are already underway at Hopkins and other institutions with some of the drugs that can potentially accomplish these goals. The Baylin and Herman labs work closely with clinicians to monitor these trials and to improve the approaches for ultimate success.

Symposium: Biobanking (19 May 2010)

Bartha-Maria Knoppers

Professor
Director, Centre of Genomics and Policy, Faculty of Medicine
Department of Human Genetics, McGill University, Canada
Former Chairperson, HUGO Ethics Committee

Title: The Public Population Project in Genomics : An " Ethical" Proof of Concept

It is not the number or size of biobanks that makes them a valuable asset for research, but rather their quality and interoperability. It was with these two goals in mind that the Public Project in Genomics (P3G) was launched in 2007.

What are the results of this effort? Where is it going? More importantly, what were the challenges it faced and the lessons learned that can serve to guide the next generation of genomic research?

Biosketch

Dr. Knoppers (PhD., O.C.) is Director of the Centre of Genomics and Policy at the Department of Human Genetics, Faculty of Medicine, McGillUniversity. She is former Chair of the International Ethics Committee of the Human Genome Organization (1996-2004) and currently, Chair of the Ethics Working Party of the International Stem Cell Forum (U.K.). In 2003, she founded the international Public Population Project in Genomics (P3G) and became the Principal Investigator of CARTaGENE. CARTaGENE is a resource of samples and data on the genetic diversity of the population of Quebec

David B Goldstein

Director, Duke Institute for Genome Sciences and Policy
Center for Human Genome Variation

Title: Rare and common variants in human disease

There are now several confirmed common variants that influence common diseases, responses to infection, and responses to drugs. For most common diseases however all common variants identified explain only a few percent of the known heritability, and many of the signals emerging from genome wide association studies have yet to be tracked to single common variants, raising the possibility that in some cases the signals emerge from associated sets of rare variants. Here I argue that progress in identifying the so called 'missing heritability' for many human traits will be facilitated using an extreme phenotype whole-genome sequencing paradigm. I illustrate the basic structure and motivation for this approach using examples from our work on host determinants of control of HIV-1.

Biosketch

Dr. David Goldstein is Professor of Molecular Genetics & Microbiology, Department of Biology and Director of Duke Institute for Genome Science & Policy, Center for Human Genome Variation since June 2005. Dr. Goldstein received his PhD in Biological Sciences from Stanford University in 1994, and from 1999 to 2005 was Wolfson Professor of Genetics at University College London. In April 2007, he was appointed Honorary Professor, Institute of Neurology, University College London, UK.

Dr. Goldstein is the author of over 100 scholarly publications in the areas of population and medical genetics. His principal interests include human genetic diversity, the genetics of neurological disease, and pharmacogenetics. As part of the Center for HIV-AIDS Vaccine Immunology (CHAVI), his laboratory also has a particular focus on the host genetics of HIV infection. He is the recipient of one of the first seven nationally awarded Royal Society / Wolfson research merit awards in the UK for his work in human population genetics.Dr. Goldstein was awarded the Triangle Business Journal Health Care Heroes Award:Innovator/Researcher Award, March 2008.

David Cox

Senior Vice President and Chief Scientific Officer
Applied Quantitative Genotherapeutics
Pfizer Inc.

Title: Advancing biobanking through public-private collaboration

In order to make drugs that matter, the pharmaceutical needs to engage the broader scientific and healthcare community in a more collaborative fashion than is presently the case. Historically, there has been poor alignment of molecular understanding with clinical need. What is required to improve this alignment is long-term clinical epidemiological data associated with critical samples not presently available to industry. While academia often looks to industry as simply a source of cash, private sector contributions of knowledge and data, added to public sector data and samples, could potentially be much more valuable than cash alone. But how does one reconcile altruism, protection of stakeholders and open collaboration, primary principles of public biobanking, with intellectual property and profit, which are often viewed as synonymous with industry involvement? One solution is collaboration at the "precompetitive stage", such that all knowledge gained from an industry-academic collaboration would be freely available to all in the public domain, with no intellectual property restrictions. Such public domain biological knowledge would then be the basis for development of novel therapies by either the public or the private sector. Scientists in industry and academic could lead by example, setting up such pre-competitive collaborations and demonstrating to the world that they can work. One example of such an approach is the "expert centre" concept now being developed as part of the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) in Europe.

Biosketch

David serves as Chief Scientific Officer for the Applied Quantitative Genotherapeutics Unit of Pfizer's Biotherapeutics Research & Development. This new unit brings together human genetics, systems biology, and cell biology, combining internal capabilities with outside collaborations, to focus on increasing preclinical target validation with the aim of significantly improving clinical survival. David is a co-founder of Perlegen, and was most recently Chief Scientific Officer of the company since its formation in 2000. David was Professor of Genetics and Pediatrics at the Stanford University School of Medicine as well as the co-director of the Stanford Genome Center. He obtained his A.B. and M.S. degrees from Brown University in Rhode Island and his M.D. and Ph.D. degrees from the University of Washington, Seattle. He completed a Pediatric Residency at the Yale-New Haven Hospital in New Haven, Connecticut and was a Fellow in both genetics and pediatrics at the University of California, San Francisco. David is certified by the American Board of Pediatrics and the American Board of Medical Genetics. He was an active participant in the large scale mapping and sequencing efforts of the Human Genome Project while carrying out research involving the molecular basis of human genetic disease. David has been a member of several commissions and boards, including the National Bioethics Advisory Commission (NBAC) and the Health Sciences Policy Board of the Institute of Medicine. He has also served on a number of international committees, including the Council of the Human Genome Organization (HUGO). He has authored over 100 peer-reviewed scientific publications and has served on numerous editorial boards. Dr. Cox's honors include election to the Institute of Medicine of the National Academy of Sciences.

Jennifer Harris

Senior Researcher
Norwegian Institute of Public Health (NIPH)
Oslo, Norway

Title: Biobank harmonization for health

Unravelling the causes of complex diseases and translating these efforts into public health and clinical practice are primary goals of biomedicine. It has become well recognized that conducting the high-throughput and interdisciplinary science underlying this agenda is heavily reliant on access to large-scale biobanks of wide geographic span which curate a rich array of data including genotypes, other biomarkers, clinical measures, environmental exposures, life-style and social factors plus biospecimens. The FP6 project entitled 'Promoting Harmonisation of Epidemiological Biobanks in Europe' (PHOEBE) (http://www.phoebe-eu.org) was designed to promote the harmonization of European population-based biobanks to optimise their ability of biobanks to communicate with one another, share ideas, information and data, and collaborate effectively in a complex world where laws and ethical guidelines often differ between nations and over time. To fulfil its mission PHOEBE worked closely with other international harmonization initiatives to help bring coherence to these issues and to forge a common way forward. This presentation will describe the challenges encountered and solutions proposed to help harness the full potential of biobanking in the generation and translation of research findings to improve health.

Biosketch

Dr. Jennifer R. Harris is a senior researcher in genetic epidemiology at the Norwegian Institute of Public Health (NIPH) in Oslo. She led the FP6 project *Promoting Harmonization of Epidemiological Biobanks in Europe* (PHOEBE), she is a scientific partner in FP7 project *European Network of Genetic and Genomic Epidemiology* (ENGAGE) and has been extensively involved in several other international biobanking initiatives such as P3G, BBMRI and GenomEUtwin. She is also a special expert at the National Institute on Aging, NIH, USA where she develops research directions integrating genetics and genomics with behavioural and social research. She is the founder of the NIPH twin panel and program of research, which was established under her research interests to integrate developmental approaches with genetic research into complex phenotypes. Her current research uses the discordant twin design to study epigenetic and CNV influences on autoimmune disorders and she is also studying genetic and environmental influences on body composition.

She obtained her PhD in Human Development and Genetics from the PennsylvaniaStateUniversity and was awarded a post-doctoral research fellowship at the Karolinska Institutet in Stockholm through the John D. and Catherine T. MacArthur Foundation. She has broad commitment to the wider scientific community and serves on several expert panels, boards, steering groups, scientific advisory committees and editorial boards.

Symposium: Microbial Genomics and Metagenomics (19 May 2010)

Martin Hibberd

Senior Group Leader
Assoc Director, Infectious Diseases
Genome Institute of Singapore

Title: Dengue virus genomics, a window into evolution, virulence and transmission patterns

Dengue infection results in a wide clinical spectrum, ranging from asymptomatic, through fever to life threatening complications. However, the mechanisms that lead to these outcomes remain largely unknown. To investigate if there are early processes that can perhaps direct the host response in a way that can lead to a specific disease, we have used whole genome expression profiling from the first few days to convalescence and correlated this with disease progression. We have also noted that these host-responses can be linked to clades of viruses and are now exploring their relationship to viral sequence. Recent developments in technology has enabled us to sequence these viruses deeply, to recognize minority genomes that may have an impact on disease outcome and may also potentially be used to follow viral evolution occurring in one single transmission. Little is known about the pattern and dynamics of dengue virus (DENV) transmissions within outbreak situations. By exploiting genomic data from an intensively studied, we are able to describe the molecular epidemiology of DENV at a uniquely fine-scaled temporal and spatial resolution. We also investigated host response differences between these strains and found that they exhibited a marked difference in inducing type I IFN response. Investigating further strains suggested that this difference is strain-dependent but not serotype-specific. Our report indicates the existence of a strain-specific virulence factor that may impact on disease severity.

This genomics approach has identified new understandings of the causes of different outcomes; and the identified genes in humans and virus may be suitable targets for the development of much needed drugs and diagnostics against dengue.

Biosketch

Dr Martin Hibberd BSc(Hons) PhD; Senior Group Leader, Infectious Diseases at the Genome Institute of Singapore and has adjunct positions at the National University of Singapore and Imperial College (London, UK). Graduated with Honors from Brunel University in 1985 (West London, UK) and received his Doctorate, on the immune-genetics of the human T-cell antigen receptor, from King's College, London. Has a broad scientific background spanning both microbial and human determinants of infectious and inflammatory diseases. Previous posts include WHO-funded Senior Microbiologist at the UK's central Public Health Laboratories, and for seven years prior to his current appointment he was Lecturer and Senior Lecturer in Pediatric Infectious Diseases at the Imperial College School of Medicine, one of the very top-ranking British universities. Current research interests cover both pathogen and host aspects of infectious disease, understanding how microbial agents causes the observed disease (including pathogen identification and sequence characterization) and why specific individuals are susceptible to the disease (using host genetics on a genomic scale). Approaching infectious disease from these two directions also allows specific host pathogen responses to be investigated (utilizing RNA micro and low density arrays). This work aims to identify key host responses to specific pathogens that could be targeted by new therapies.

Peer Bork

Senior group leader (bioinformatics) at the EMBL (Heidelberg)
Joint coordinator of EMBL Structural and Computational Biology unit
Visiting group leader at the MDC (Berlin-Buch)

Title: Comparative (meta) genomics: Lessons from a tiny bacterium and a complex microbial community

Environmental sequencing is uncovering genomic parts of various microbial communities, but our understanding of their functioning remains limited. I will briefly introduce the powers and pitfalls of metagenomics (Raes and Bork, 2008) and will zoom in on the human gut where recent technological advances (Qin et al., 2010) have led to a wealth of data and where first concepts arise on how the landscape of gut microbiomes around the globe might look like. As in the future, single cell sequencing will enable the study of the species in a community one by one and as one not only needs a broad overview, but also a detailed understanding of the roles of individual species, I will also introduce a pilot study of one of the smallest bacteria, *Mycoplasma pneumoniae*. Data on the transcriptome, metabolome and proteome have been consistently generated and integrated to reveal a wealth of information about the biology of a genome-reduced bacterium that was found to be remarkably complex (Kuehner et al, 2009, Yus et al., 2009, Guell et al., 2009).

Raes J and Bork B, Nat Rev Microbiol. 2008 Sep;6(9):693-9

Qin J et al., Nature. 2010 Mar 4;464(7285):59-65

Kuehner S et al., Science. 2009 Nov 27;326(5957):1235-40

Yus E et al., Science. 2009 Nov 27;326(5957):1263-8

Guell M et al., Science. 2009 Nov 27;326(5957):1268-71

Biosketch

Peer Bork, PhD, is senior group leader and joint head of the Structural and Computational Biology unit at EMBL, a European research organization with headquarters in Heidelberg. He also holds an appointment at the Max-Delbrueck-Center for Molecular Medicine in Berlin. Dr Bork received his PhD in Biochemistry (1990) and his Habilitation in Theoretical Biophysics (1995). He works in various areas of computational biology and systems analysis with a focus on function prediction, comparative analysis and data integration. He has published more than 400 research articles in international, peer-reviewed journals, among them more than 45 in Nature, Science and Cell. According to ISI (analyzing the last 10 years), Dr. Bork is currently the most cited European researcher in Molecular Biology and Genetics. He is on the editorial board of a number of journals including Science and PloS Biology, and functions as senior editor of the journal Molecular Systems Biology. Dr Bork co-founded four biotech companies, two of which went public. More than 25 of his former associates now hold professorships or other group leader positions in prominent institutions all over the world. He received the "Nature award for creative mentoring" for his achievements in nurturing and stimulating young scientists. He was also the recipient of the prestigious "Royal Society and Academie des Science Microsoft award" for the advancement of science using computational methods.

Stanislav Dusko Ehrlich

Research Director
Microbial Genetics Research Unit, Department of Microbiology and the
Food Chain, INRA Centre of Jouy-en-Josas
National Institute of Agronomic Research (INRA), France

Title: Characterization of the human "other" genome by the MetaHIT Consortium

Humans live in association with a large number of microorganisms, which affect our health and well-being. Most reside in the intestine, where their numbers exceed 10-fold the number of our cells. The intestinal micro-organisms are thought to have an important role in health and disease, but are still poorly characterized. To understand and exploit the impact of the gut microbes on human health and wellbeing it is necessary to decipher the content, diversity and functioning of the gut microbial community.

The European project on the metagenomics of the human intestinal tract, MetaHIT, (http://www.metahit.eu) has as a first objective creation of a catalog of the microbial genes from our intestinal tract, thus laying foundations for characterization of the gut microbial community. Next, it aims to explore associations between microbial genes and human phenotypes. For that, it develops, on the one hand, molecular tools for profiling of the intestinal microbial genes that are harbored by any individual and on the other a bio-informatics resource to organize and interpret heterogeneous information, including sequencing data and clinical metadata. Finally, MetaHIT develops approaches to detect and analyze functional interactions of microbes and the human host, focusing on the role of target genes in the microbial cell and the effect of gene products on the human host.

We have established an extensive catalog of the gut microbial genes, based on the Illumina metagenomic sequencing. The catalog contains 3.3 million nonredundant microbial genes, derived from 576.7 Gb sequence of the DNA prepared from faecal samples of 124 individuals of the European origin. The sequencing and the assembly was carried out BGI Shen Zhen, one of the MetaHIT partners. The gene set is more than 150 times larger than the human gene complement, contains an overwhelming majority of the prevalent microbial genes present in the cohort and likely includes a large proportion of the prevalent human intestinal microbial genes. The gene pool is largely shared among individuals of the cohort. Over 99% of the genes are bacterial, suggesting that the entire cohort harbors between 1000 and 1150 prevalent bacterial species and each individual at least 160 such species, which are also largely shared. We define and describe the minimal gut metagenome and the minimal gut bacterial genome in terms of functions encoded by the gene set.

Qin, J. et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464, 59-65 (2110).

Biosketch

Dr. S. Dusko Ehrlich received his Ph.D. from Université Paris VII in 1973 and was a research associate of Prof. Joshua Lederberg, Nobel prize winner, at Stanford University Medical school, from 1973 till 1977. He joined French National Institute for Medical Research (INSERM) as a Research Director in 1977, founded and led a research group at the National Center for Scientific Research (CNRS) Institute Jacques Monod at the University Paris VII until 1986. He joined National Institute for Agricultural Research (INRA) in 1986, founded the Microbial Genetics Unit and directed it till 2008. He founded the Microbiology Department of INRA in 1991 and chaired it till 2002. He has authored over 300 research papers and holds an H-index of 58. He is a member of the European Molecular Biology Organisation, the Croatian Academy of Arts and Sciences and the French Academy of Agriculture and currently coordinates the European Commission FP7 project MetaHIT (Metagenomics of the Human Intestinal Tract). His research interests are mainly in bacterial genetics, genomics and metagenomics.

Edward Rubin

Director of the U.S. Department of Energy Joint Genome Institute (JGI)
Director of the Genomics Division at Lawrence Berkeley National Laboratory (LBNL)

Title: Assembly of genes and genomes from bio-mass degrading cow rumen

The paucity of enzymes able to efficiently deconstruct plant material represents an important bottleneck in industrial scale production of cellulosic biofuels. The cow rumen harbors many uncultured microbes that have evolved optimized molecular machineries for the efficient conversion of biomass into monomeric sugars. In this study we generated and analyzed 100 gb of metagenomic sequence from the microbial community adherent to switchgrass incubated in cow rumen to characterize the genes from and the genomes of organisms involved in its deconstruction. We identified more that 300 novel full length cellulolytic genes and demonstrated approximately a third of the 100 expressed and tested for cellulose degradation activity as positive. Near complete draft coverage of 2 novel microbial genomes was generated. Our deep metagenomic sequencing of biomass adherent microbes followed by computational and functional analyses has

dramatically expanded the number and diversity of cellulolytic genes able to degrade biomass as well as revealed the draft genomes for several completely novel uncultured rumen microbes.

Biosketch

Internationally recognized geneticist Eddy Rubin has served since 2002 as Director of the U.S. Department of Energy Joint Genome Institute (JGI), and Director of the Genomics Division at Lawrence Berkeley National Laboratory in Berkeley, CA. As Director of the JGI, one of the world's largest sequencing centers, he oversaw the sequencing and analysis of 13% of the human genome (human chromosomes 5, 13 and 16), parts of the Neanderthal genome and the genomes of more than 200 animals, plants and microbes. Following the completion of the Human Genome Project, Dr. Rubin reoriented the JGI towards the application of genomics to studies related to bioenergy and climate change.

With more than 200 peer-reviewed publications, over 30 of them in the leading journals Science and Nature, his research focuses on the development of computational and biological approaches for studying genomes. His early scientific work centered on the functional exploration of the human genome, harnessing sequence comparisons between species for the discovery of genes and non-coding sequences of pivotal evolutionary and biomedical importance. More recently, he spearheaded the new science of metagenomics, deriving important insights from his investigations of microbial communities inhabiting environments ranging from gutless ocean-dwelling worms to cow rumen.

In addition to his research and the training of more than 50 scientists who have passed through his laboratory, Dr. Rubin has been an influential leader in the field of genomics through his community activities. He sits on the editorial boards of several leading journals, including the Board of Reviewing Editors for the journal Science, has organized and chaired major genomics meetings and is a member of multiple scientific advisory boards. Through government advisory committee membership, Dr. Rubin has actively influenced the directions of genomic research at the U.S. Department of Energy and U.S. National Institutes of Health.

Featured Talk (19 May 2010)

Charles N. Rotimi

Director, Center for Research on Genomics and Global Health (CRGGH)

Title: Genetic diversity and health: Opportunities and challenges presented by African genomes

Africa is the common birthplace of all human populations. As a result, modern humans have lived longer on the Africa continent than in any other geographical regions of the world. This long evolutionary history has led to several important cultural and genetic characteristics of the African people. For example, greater genetic variation is seen in present-day Africa populations, resulting in more haplotypes, lower levels of linkage disequilibrium (LD), more divergent patterns of LD and more complex patterns of population substructure. Africa has the highest linguistic diversity in the world with an estimated 2,000 languages; notably, Nigeria as a nation has over 500 listed languages. Also, due to vast cultural practices and uneven economic developments, African populations display considerable phenotypic variation. Given these important observations and the critical role of Africa in human evolutionary history, it is surprising that there has not been a global effort to systematically sample African populations for a comprehensive documentation of the scope and extent of genetic variation on the continent. In fact, an argument can be made that Africa has been, for most part, left out of the genomic revolution to date. A stark demonstration of this point is the fact that, despite the well documented success stories of genome-wide association studies (GWAS) in identifying susceptibility variants for several human diseases (www.genome.gov/gwastudies), it is remarkable that only one GWAS, the MalariaGen Project, has been published in a large collection of African individuals to date. This observation is particularly troubling given that Africa is faced with the heaviest global burden of disease, the least equipped health care system and perhaps, the least technologically developed to take advantage of the growing health and economic success stories of the genomic revolution.

The inclusion of three African populations in the international HapMap and several more African populations in the 1000 genome project and the recent whole genome sequencing of a Yoruba and Southern African individuals are beginning to address this inequity in genomic research. This presentation will explore our current understanding of the scope of genetic diversity in Africa and its implication for health and human history. A case will be made that, by not fully engaging African populations in genomic research, scientists are likely to compromise the novel insights that could be gleaned from

studying the different genetic architecture of Africa people. For example, reduced LD in African populations is likely to provide a better chance of identifying causal variants for human diseases and the long evolutionary history of the African people is also likely to advance our understanding of human adaptive history.

Biosketch

Charles Rotimi, PhD, a genetic epidemiologist and a biochemist, is a senior investigator in the Inherited Disease Branch of the NHGRI intramural program. He is the Director of the Center for Research on Genomics and Global Health (CRGGH). The mission of this new trans-NIH center is to advance research into the role of culture, lifestyle, genetics and genomics in disease etiology and health disparities. Dr. Rotimi develops genetic epidemiology models and conducts population genetics research that explores the patterns and determinants of common complex diseases in human populations with particular emphasis on populations of the African Diaspora. As a senior investigator and director of the CRGGH, Dr. Rotimi leads a team of researchers across multiple disciplines (Medicine, genetics/genomics, epidemiology, statistics and informatics) to understand the complex interactions between inherited characteristics and environmental factors in disease susceptibility, variable drug response and health disparities. For example, his team conducted the first GWAS of an African American cohort that identified several genetic variants including PMS1, SLC24A4, YWHA7, IPO7, and CACANA1H underlying susceptibility to Hypertension and blood pressure control. His lab also conducted the first and only genome-wide linkage study of Type 2 Diabetes in West Africans (the AADM Study). Dr. Rotimi's lab continues to contribute to the global understanding of human genetic variation and its implication for differential disease distribution, variable drug response and human migration history. He is the current president of the African Society of Human Genetics (www.afshg.org).

Plenary: An Expanding RNA World (19 May 2010)

Gil Ast

Professor of Molecular Biology and Bioinformatics
Department of Human Molecular Genetics and Biochemistry
Tel Aviv University Medical School

Title: Importance of alternative splicing in transcriptomic diversity and disease

Alternative splicing is a major mechanism for generation of transcriptomic and proteomic diversity, especially in humans. Mutations that affect splicing are correlated with many human diseases and cancer. However, despite decades of research we still do not understand how short exonic sequences recognized within the vast intronic oceans in which they are embedded are selected by the splicing machinery. For many years, studies aiming to shed light on this process were focused at the level of RNA. However, we increasingly understand that splicing is not an isolated process; rather it occurs and regulated co-transcriptionally. Studies by our group and others show that DNA structure in terms of nucleosome positioning and specific histone modifications may also play a role in splicing. We will first discuss evidence regarding the coupling between transcription and splicing and between chromatin structure and splicing. Finally, we will show how GC content led to the appearance of two types of exon/intron structures – termed exon and intron definition – which the splicing machinery handles differently. Splicing-disrupting mutations in these two systems have different outcome on the exon selection process. Overall, we will show how the genomic environment surrounding mutations affect the exon selection process.

Biosketch

Our group has made seminal contributions to our understanding of the acquisition of complexity in the expression of genomic information through alternative mRNA splicing, and the involvement of this process in genetic disorders and cancer. The broad focus of our laboratory is on regulation of splicing of pre-mRNA and the importance of alternative splicing in the generating transcriptomic diversity unique to our species. We study mechanisms of alternative splicing regulation using a combination of computational and experimental methods. We also study the potential link between DNA packaging by histones and epigenetic modifications and how these effect processing of pre-mRNA.

Yoshihide Hayashizaki

Director
Omics Science Center
RIKEN Genomic Sciences Center

Title: Transcriptome analysis — A way to illuminate the genome network

The first step on the road to knowledge was taken in 1995 when we first glimpsed the RNA complexity. We realized the need of new technologies to construct the full-length cDNA library (Cap trapper method, normalization, subtraction, new cDNA vector and thermoprotection technology of reverse transcriptase using trehalose), and high throughput sequencing (Riken integrated sequence analysis system; RISA) was essential to reach the goal of large-scale data collection. In order to process the data we started to collect people, and the International FANTOM consortium was founded in 2000. The combination of talented people and data have given us millions of full-length cDNA sequences, CAGE tags and ditags, collected in the *de facto* world standard database (5 visits/second)

The efforts of the FANTOM consortium have revealed non-coding RNA (ncRNA) as the major product of the genome (53%), and that the actual number of genes is diminished through the history of life science due to gene fusion, but increased by the number of ncRNA. We also learned that 72% of all genes have sense/antisense transcripts.

All the discoveries we have made regarding RNA have given us a new view of the central dogma, collapsing the old DNA to protein concept. The double stranded RNA (dsRNA) for RNAi cannot only be created by virus and miRNA and S/AS is a new source of dsRNA. RNA is flexible and regulates gene expression through RNAi and other mechanisms. Now we have found a group of transcription factors that act like a switch between RNAi and regular gene expression by interacting with the RNAi machinery directly. Altogether these new findings dissolve the central dogma into a new picture with new interaction cascade and the unexpected complexity of combined omics.

Biosketch

Yoshihide Hayashizaki received his M.D. and Ph.D. from Osaka University Medical School in 1982 and 1986, respectively.In 1992, he joined RIKEN, and was appointed Project Director for the RIKEN Genome Project in 1995. Since then he has been taken a transversal data-driven approach to analyze transcriptomes by developing unique technologies including a series of full-length cDNA technologies. With this approach, he has established large amount of full-length cDNA clone bank. This activity was followed by organization of FANTOM (Functional Annotation of Mammalian), an international consortium, originally to annotate a large number of cDNA and subsequently expanded to transcriptome and network analysis. FANTOM activities revealed that more than 63% of the genome — not just the ~1.5% fraction that are protein-coding exons — is transcribed as RNA. In 2008, he appointed to the Director of the Omics Science Center. The center was established to link a variety of omics subdisciplines to molecular networks and pathways in order to advance our understanding of biological phenomena as systems at the molecular level.

Plenary: Gene Expression & Human Variation (19 May 2010)

Michael Snyder

Professor and Chair of Genetics
Director, Stanford Center for Genomics and Personalized Medicine

Title: Transcription binding variation in eukaryotes

Variation in transcriptional regulation is thought to be a major cause of phenotypic diversity. Although widespread differences in gene expression among individuals of a species have been observed, few studies have examined the variability of transcription factor (TF) binding, and thus the extent and underlying genetic basis of TF binding diversity is

largely unknown. We mapped differences in transcription binding among individuals and elucidated the genetic basis of such variation on a genome-wide scale for both yeast and humans. For humans we mapped the binding sites of RNA Polymerase II (PolII) and a key regulator of immune responses, NFκB (p65), in ten lymphoblastoid cell lines and found that 25% and 7.5% of the respective binding regions differed between individuals. Binding differences were frequently associated with SNPs and genomic structural variants (SVs) and were often correlated with differences in gene expression, suggesting functional consequences of binding variation.

To further understand the genetic basis of transcription factor binding variation, we mapped the binding sites of Ste12 in pheromone-treated cells of 43 segregants of a cross between two highly diverged yeast strains and their parental lines. We found that the majority of TF binding variation is *cis*-linked and that many variations are associated with polymorphisms residing in the binding motifs of Ste12 as well as those of several proposed Ste12 cofactors. We also identified trans factors that modulate Ste12 binding to specific promoters. Ste12 binding strongly correlates with gene expression indicating that binding variation is functional. Overall these different studies identified genetic regulators of molecular diversity among individuals and provide novel insights into variation in eucaryotes and mechanisms of gene regulation.

Biosketch

Michael Snyder is the Stanford Ascherman Professor and Chair of Genetics and the Director of the Center of Genomics and Personalized Medicine. Dr. Snyder received his Ph.D. training at the California Institute of Technology and carried out postdoctoral training at Stanford University. He is a leader in the field of functional genomics and proteomics. His laboratory study was the first to perform a large-scale functional genomics project in any organism, and currently carries out a variety of projects in the areas of genomics and proteomics both in yeast and humans. These include the large-scale analysis of proteins using protein microarrays and the global mapping of the binding sites of chromosomal proteins. His laboratory built the first proteome chip for any organism and the first high resolution tiling array for the entire human genome.

Howard McLeod

Director, UNC Institute for Pharmacogenomics and Individualized Therapy

Title: Using the genome to guide therapy

The field of pharmacogenomics has seen some exciting advances in the recent past. The Human Genome Project and International HapMap projects have uncovered a wealth of information for researchers. This has lead to the discovery of clinically predictive germline genotypes (e.g. UGT1A1*28-irinotecan, TYMS TSER-fluoropyrimidines, CYP2D6-tamoxifen, CYP2C19-clopidogrel), germline haplotypes (e.g. VKORC1 Haplotype A-warfarin) and somatic mutations (e.g. epidermal growth factor receptor-gefitinib/erlotinib, KRas-cetuximab/panitumumab). The introduction of FDA approved pharmacogenetic tests (UGT1A1*28, CYP2C9/VKORC1, CYP2D6/CYP2C19) and the initiation of genotype-guided clinical trials have provided the first steps towards the integration of pharmacogenomics into clinical practice. It is also clear that there are many non-scientific barriers to clinical application. These include integration of new tests into health systems, changing old habits to allow application of new data, and the reality that the cost of both testing and the therapeutic options are a key driver in health care. As the scientific evidence matures, A team research approach must be more aggressively applied in order to overcome the many obstacles to delivering more careful selection of drug therapy.

Biosketch

Dr Howard McLeod is Fred N. Eshelman Distinguished Professor and Director, UNC Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina, Chapel Hill. Dr McLeod holds appointments in the Schools of Pharmacy and Medicine and the Lineberger Cancer Center. Dr McLeod is also the Principal Investigator for CREATE, a member of the NIH funded Pharmacogenetics Research Network and is a member of the FDA committee on Clinical Pharmacology. He also directs the Pharmacogenetics for Every Nation Initiative, which aims to help developing countries use genetic information to improve National Drug Formulary decisions. Howard has published over 350 peer reviewed papers on pharmacogenomics, applied therapeutics, or clinical pharmacology and continues to work to integrate genetics principles into clinical practice to advance individualized medicine.

Plenary: Human Genome Variation in Evolution and Disease (20 May 2010)

Mark McCarthy

Robert Turner Professor of Diabetes
Oxford Centre for Diabetes, Endocrinology and Metabolism

Title: Type 2 diabetes: Susceptibility variants across the allele frequency spectrum

Ongoing efforts to define additional common variants influencing traits of biomedical importance are motivated by the biological insights that follow the identification of robust causal relationships between genome sequence variation in a given location and the phenotype of interest. Once the “low hanging fruit” have been picked, further success is increasingly dependent on improvements in power that flow from combined analysis of multiple GWA and replication data sets. A growing number of global consortia have emerged to manage such efforts.

For type 2 diabetes (T2D), the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) consortium has concentrated current efforts on T2D signals in European-descent populations. In recent months, the efforts of the consortium have contributed to expansion of the number of confirmed T2D common variant signals to almost 40. The newly-identified loci include several that appear to influence insulin action (though most are associated with deficits in beta-cell function), several that map to imprinted loci and have been shown to have parent-of-origin effects, and the first located on the sex-chromosomes. One of the striking features of these data is the frequency with which the same regions harbour independent signals influencing other, non-diabetes-related, traits: this speaks to tissue-specific subtleties in regulatory control around the genes concerned.

The principal challenges going forwards are: (a) to move from these association signals to understanding of biological mechanisms through identification of the causal variants driving these effects and the genes and pathways through which they operate; (b) to obtain a more complete enumeration of alleles contributing to genetic risk of T2D through extending the spectrum of variants evaluated; and (c) to translate these discoveries into advances in clinical management of diabetes. The growing capacity of high-throughput sequencing technologies is likely to have a transformative effect on all three, and I will describe some of the approaches we are taking to address these challenges.

Biosketch

Mark McCarthy is Robert Turner Professor of Diabetes at Oxford University, based at the Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM) and at the Wellcome Trust Centre for Human Genetics. His work on the genetics and genomic basis of type 2 diabetes seeks to define the relationship between genetic variation, environment, intermediate phenotypes, and clinical disease, in the belief that such information will translate into a significant impact on clinical care for people with diabetes, through more rational and effective deployment of available preventative and therapeutic modalities.

He obtained his medical degree at the University of Cambridge and postgraduate clinical training in general medicine, diabetes and endocrinology at the London Hospital. Following research training with – amongst others – Newton Morton and Eric Lander, he took up an appointment as Senior Lecturer (subsequently Professor in Genomic Medicine) at Imperial College in 1995, moving to Oxford as the Robert Turner Professor of Diabetes in 2002. In the past decade, his group has become established as one of the leading international teams working on the genetics of T2D, with expertise extending from physiological characterisation through to high-throughput genomic analysis and related issues in computational biology. Prof McCarthy currently leads the T2D component of the Wellcome Trust Case Control Consortium, and has been intimately involved in many of the recent discoveries generated by the application of genome wide association approaches to diabetes, obesity and related traits. His group is increasingly involved in efforts to understand how such information can be translated into advances in clinical management, and in extending these studies beyond the European context. He has a strong commitment to the wider research community, having served as a member and chair of the Diabetes UK Research Committee, and on Wellcome and MRC panels. He serves on the Editorial Board of Diabetes, Diabetologia, PLoS Medicine, Human Molecular Genetics and several other journals.

Stephen W Scherer

Director, McLaughlin Centre and the Toronto Centre for Applied Genomics
Professor, Hospital for Sick Children and University of Toronto

Title: Clinical context of copy number and structural variation in the human genome

During the past five-years, copy number variation (CNV) and structural variation (SV) have emerged as highly prevalent forms of genomic variation, bridging the interval between long-recognised microscopic chromosomal alterations and single-nucleotide changes. These genomic differences among humans reflect the dynamic nature of genomes, and account for both natural variations among us and variations that predispose to conditions of medical consequence. I will discuss CNV-SV in their historical and medical contexts, focusing on how these variations can be recognised, documented, characterised and interpreted. I will also discuss how they can cause disease or influence adaptation to an environment. Our recent work on autism spectrum and related neuropsychiatric disorders will also be described to illustrate salient characteristics and residual enigmas of this complex form of genetic variation.

Biosketch

Dr. Scherer holds the GlaxoSmithKline-CIHR Endowed Chair in Genetics and Genomics at The Hospital for Sick Children and University of Toronto. He has made numerous contributions to medical genetics including mapping sequencing and disease gene studies of human chromosome 7. He collaborated with the J. Craig Venter Institute to generate the first genome sequence of an individual ushering in an era of genomic medicine. In 2004, his team co-discovered global gene copy number variation (CNV) and has since shown that CNV is the most abundant type of variation of human DNA. His group has also discovered CNV to contribute to the etiology of autism and the Database of Genomic Variants he founded facilitates hundreds of thousands of diagnoses each year. He sits on the Scientific Advisory Board of Combimatrix Diagnostics and Autism Speaks, and he is on the Board of Trustees of Genome Canada and the Human Genome Organization (HUGO). Dr. Scherer has won numerous honors including the 2004 Steacie Prize, an International Howard Hughes Medical Institute Scholarship, and the 2008 Premier Summit Award for Medical Research.

Felix Jin Li

Professor and Vice President of Fudan University
Professor at the National Human Genome Center in Shanghai
Co-Director of CAS-MPG Partner Institute for Computational Biology at Chinese Academy of Sciences

Title: Identification of copy number variation hotspots in human populations

Copy number variants (CNVs) in the human genome contribute to both Mendelian and complex traits as well as genomic plasticity in evolution. Much progress was made to unravel the CNV formation mechanisms; however, the evaluation of CNV mutation rate at genome level poses an insurmountable practical challenge. We showed that an approximate estimation of CNV mutation rate could be achieved using the phylogeny information of flanking SNPs. This allows a genome-wide comparison of mutation rates between CNVs for identifying mutational hotspots using HapMap genotyping data. The mutation rates for the majority of 4,330 CNVs investigated in this study are at the order of 10-5 per generation, consistent with experimental observations at individual loci. Notably, the mutation rates of 132 (3.0%) CNVs are at the order of 10-3 per generation, therefore, identified as hotspots. Further analyses revealed that genome architecture has a potential role in inciting mutational hotspots in the human genome.

Biosketch

Li Jin is a Chinese geneticist who led the research that concluded that all East Asians, including the Chinese, originated from Africa, adding support to the recent single-origin hypothesis of which he is considered a leading proponent. His team analyzed the Y chromosomes of males around China and compared this group with those of Southeast Asians and Africans. Results of the analysis suggested that Southeast Asia was the first destination of the

migration from Africa to Asia which began approximately 60,000 years ago; from there, migrants moved into Southern China, then crossing the Yangtze River to Northern China. The 1998 study, which used genetic markers called microsatellites to compare Chinese populations, did not support an independent origin of Homo sapiens in China. The findings contradict the hypothesis that Peking Man (Homo erectus) was the ancestor of the Chinese people.

Jin is a professor and Dean of School of Life Sciences of Fudan University, as well as a professor at the National Human Genome Center in Shanghai. Jin is the principal investigator of East Asian populations for the Genographic Project which collects DNA samples to map historical human migration patterns around the world. His research interests are in human and medical genetics, population genetics, genomics & bioinformatics, anthropology.

Chen Award Lecture (20 May 2010)

Ng Huck Hui

Senior Group Leader, Genome Institute of Singapore
Associate Professor (Adjunct), National University of Singapore and
Nanyang Technological University

Title: Deciphering and reconstructing the embryonic stem cell transcriptional regulatory network

Embryonic stem (ES) cells are characterized by their ability to self-renew and remain pluripotent. Transcription factors have critical roles in the maintenance of ES cells through specifying an ES-cell-specific gene expression program. Deciphering the transcriptional regulatory network that describes the specific interactions of these transcription factors with the genomic template is crucial for understanding the design and key components of this network. To gain insights into the transcriptional regulatory networks in ES cells, we use chromatin immunoprecipitation coupled to ultra-high-throughput DNA sequencing (ChIP-seq) to map the locations of sequence specific transcription factors. These factors are known to play different roles in ES cell biology. Our study provides new insights into the integration of these regulators to the ES cell-specific transcription circuitries. Collectively, the mapping of transcription factor binding sites identifies new features of the transcriptional regulatory networks that define ES cell identity. Using this knowledge, we investigate nodes in the network which when activated, will jump-start the ES cell-specific expression program in somatic cells.

Biosketch

Huck-Hui NG is currently a Senior Group Leader of the Genome Institute of Singapore (GIS) and an Associate Professor (Adjunct) with the National University of Singapore and the Nanyang Technological University.

Huck-Hui NG graduated from the National University of Singapore and obtained his PhD from the University of Edinburgh. He spent the next few years working at the Harvard Medical School as a Damon Runyon-Walter Winchell research fellow.

His lab is studying gene regulation in stem cells. Specifically, his group is using genome wide approaches to dissect the transcriptional regulatory networks in embryonic stem cells and to identify key nodes in this network. More recently, his lab has begun to investigate the reprogramming code behind the induction of pluripotency in somatic cells. His research work has earned him several prestigious national accolades including the Singapore Youth Award 2005 and the National Science Award 2007.

Symposium: Quantitative Genomics (GWAS, QTL, Expression Cassettes) (20 May 2010)

Emmanouil Dermitzakis

Department of Genetic Medicine and Development
University of Geneva Medical School

Title: Cellular genetics and genomics

Gene expression is an important phenotype that informs about genetic and environmental effects on cellular state. The elucidation of the genetics of gene expression (also know as eQTLs) is highly informative of the impact of genetic variants in the cell and the subsequent consequences in the organism. In this talk I will discuss recent advances in three key areas of the analysis of gene expression in human populations: (i) analysis of gene expression in multiple populations and tissues and the degree of overlap off genetic effects among populations and tissues; (ii) interrogation of gene expression in population samples using next generation sequencing (RNAseq) and the impact of the increased resolution on understanding the fine structure of genetic effects; (iii) integration of eQTLs with GWAS signals to understand the causal regulatory effects on complex traits. Perspectives on cellular phenotypic variation and the elucidation of their genetic effects will also be discussed.

Biosketch

Emmanouil Dermitzakis is currently a Louis-Jeantet Professor of Genetics in the Department of Genetic Medicine and Development of the University of Geneva Medical School. He obtained his B.Sc. in 1995 and M.Sc. in 1997 in Biology from the University of Crete (Greece) and his PhD in 2001 from the PennsylvaniaStateUniversity in the US, studying the evolutionary biology and population genetics of regulatory DNA in mammals and *Drosophila*. His post-doctoral work was at the University of Geneva Medical School, focusing on comparative genome analysis and the functional characterization of conserved non-genic elements. He previously was an Investigator and Senior Investigator at the Wellcome Trust Sanger Institute since April 2004. His current research focuses on the genetic basis of regulatory variation and gene expression variation in the human genome (see also http://www.sanger.ac.uk/humgen/genevar), the processes that govern non-coding DNA evolution. He has authored and co-authored more than 50 papers in peer-reviewed journals with more that 20 of them in journals such as Nature, Science and Nature Genetics. His papers have attracted more than 4000 citations and his H-factor is 27. He has been invited to give talks and keynote lectures in the most prestigious genetics meeting in the world and is the organizer of multiple training courses including the Wellcome Trust HapMap course and co-organizer of the Wellcome Trust School of Human Genomics with Leena Peltonen. He has served as an analysis co-chair in the pilot phase of the ENCODE (ENCyclopedia Of Dna Elements) consortium and member of the analysis group of the Mouse Genome Sequencing Consortium and the International HapMap project. He now has a leading analysis role in the extension of the HapMap (aka HapMap3 project) and is a member of the analysis group of the 1000 genomes project. He currently serves in the Board of Reviewing Editors of *Science*, and he is a Senior Editor in *PLoS Genetics*.

Frank Johannes

Faculty of Mathematics and Natural Sciences
Department:Bioinformatics
Gron Institute Biomolecular Sciences & Biotechnology
Groningen Bioinformatics Centre
University of Groningen

Title: Complex trait dynamics following epigenomic perturbation

Understanding the role of epigenetic variation in complex trait inheritance is one of the major challenges at the horizon of quantitative genetics. To address this challenge experimentally two groups have recently constructed so-called Epigenetic Recombinant Inbred Lines (EpiRILs). These populations were derived from two parents with identical DNA sequences but drastically divergent methylation profiles as a result of differential epigenomic perturbation. This system therefore provides a unique opportunity to study patterns of epigenetic inheritance against an isogenic background. Estimates suggest that epigenetic variation in these populations can account for up to 30% of the variation observed in complex traits, even eight generations later.

However, many of the segregating epialleles display intriguing time-dependent characteristics. These non-Mendelian properties hint at a dynamic rather than a static epigenetic architecture. Here we incorporate these phenomena into a quantitative genetics framework. We explicitly model the time-dependent behavior of epigenetic variation over generation time, and trace its phenotypic consequences at the level of the population. Theoretical arguments show that the simultaneous action of allele-specific expression, recombination and epiallelic stability are key parameters in this system. Our results have direct implications for the dissection of complex traits in humans.

Biosketch

The basic motivation of my research is to understand the heritable basis of complex traits in experimental populations through statistical genetic and bioinformatic approaches. Here are some topics that I am currently interested in:

- QTL mapping methods with a particular focus on the characterization of time-dependent effects
- Conceptual and methodological approaches to the epigenetic analysis of complex traits
- Methodological issues in mapping chromatin differences between organisms or tissue types using ChIP-chip or ChIP-seq
- Integrative analysis of multilevel genetical-genomics data

Augustine Kong

Vice President of Statistics
DeCOde, Iceland

Title: Some thoughts on genetics studies in the near future

With the cost of full genome sequencing coming down rapidly, it is expected to have a profound impact on how human genetics studies are performed in the next few years, particularly in the area of disease variants identification. However there is much information in the existing SNP data that remains to be exploited. Realizing their full potential would not only allow us to do more now, but, most importantly, would put us in the best position to take advantage of the sequencing data when they arrive.

Biosketch

Augustine Kong has served as Vice President of Statistics at Decode genetics since 2001. Dr. Kong received his Ph.D. in statistics from Harvard University in 1986. He was a tenured professor at University of Chicago, in statistics and then in human genetics, from 1994 to 2000.

Martin Vingron

Director, Max Planck Institute for Molecular Genetics, Germany
Head of Computational Molecular Biology Department

Title: Transcription factor binding site analysis and histone modifications point at two classes of promoters

This talk will summarize our work on prediction of transcription factor binding site prediction and identification of common binding sites among co-expressed genes. Our analyses demonstrates that promoters with high and low contents in CpG dinucleotides, respectively, show distinctly different features. This division is also visible in an analysis of the relationship between histone modifications and gene expression levels.

Biosketch

Martin Vingron is a mathematician by education who has done his PhD in computational biology at EMBL in 1991. At the time and for a number of years of postdoctoral training his research has focused on the analysis of protein sequences, sequence analysis, sequence comparison, and molecular evolution. Methods of discrete optimisation were used for the design of comparison algorithms and probability theory was applied to answer questions of significance of computational results. Later, as a department head at the German Cancer Research Center, his focus shifted towards the processing and mathematical analysis of DNA microarrays. Accordingly, the methods largely drew on statistical data analysis techniques. During the last years his research interest lies in utilizing gene expression data as well as evolutionary data for the elucidation of gene regulatory mechanisms. He also acts as a chair of the *RECOMB* <http://recomb.org/> (Research in Computational Molecular Biology) steering committee.

Symposium: Next Generation Sequencing – 1000 Genome Project and Beyond (20 May 2010)

Henry Yang Huanming

President & Professor
Beijing Genomics Insitute (BGI)-Shenzhen, China

Title: Human Genomics - From HGP to 1000 genomes and beyond

The impacts of the International Human Genome Project (HGP) on life sciences and medicine have been acknowledged by most of us since its completion in 2003. The International HapMap Project, with its goals to provide powerful tools to link the differences in our genome and differences in our health, has generated millions of SNPs which have made GWAS (Genome-Wide Association Studies) a successful strategy for identification of genomic variations related to human common diseases.

The International 1000 Genomes Project was proposed by BGI in China, Sanger Institute in UK, and the 3G centers in USA in 2006, and later joined by groups in Germany and 3 major sequencer-producers. The project is designed to identify SNPs with lower MAF (Minor Allele Frequency) than those by HapMap, and other genomic variations such as CNVs (Copy Number Variations), by sequencing hundreds of samples from the 3 major populations. Since its initiation in 2008, this project has generated much more sequence data than the total amount of data ever existing. The project has just extended to its Phase II by sequencing much more samples from more selected populations to much higher depth. All the data, both preliminary and processed, will be freely available to the scientific community in the world.

Meanwhile, genomes of a dozen of individuals have been completely sequenced to various depth. The concept of human pan-genomes has been proposed based on the non-or-all of genomics regions existing in various populations. The human metagenomes projects have generated more than 3 millions of microbial ORFs, approximately 150 times that of the estimated number of human genes. The coverage and assemble of the human ancient genome have demonstrated the power of improving sequencing and innovative bioinformatics tools for the next-generation sequencing technology. Perhaps it could be concluded that now we have really entered the genomics era.

Biosketch

Dr. Yang received his Ph.D. from University of Copenhagen, Denmark, in 1988, and his postdoc training in France and USA afterwards. He returned to China in 1994 and co-founded BGI (formerly Beijing Genomics Institute) in 1999.

As the co-founders of BGI, he and his collaborators have made a significant contribution to the HGP and HapMap projects, as well as to sequencing and analysing genomes of rice, chicken, silkworm, giant panda, cucumber, and many microorganisms. BGI published the first Asian's genome, human pan-genome, human ancient genome, and human gut metagenoms by means of new-generation sequencing technoloty and innovative bioinformatic tools recently. BGI now has becomes one of the major genomics centers in the world.

Dr. Yang has received many awards and honors, including Research Leader of the Year by *Scientific American* in 2002 and Award in Biology by the *Third World Academy of Sciences*（TWAS）in 2006. He was elected as a foreign member of EMBO in 2006, and an Academician of Chinese Academy of Sciences in 2007, and a fellow of TWAS in 2008.

Ruan Yijun

Senior Group Leader
Assoc Director, Genomic Technologies

Title: Chromatin Interaction and Transcription Regulation

Genomes are known to be organized in 3D structures *in vivo* through interactions with protein factors for nuclear process such as transcription, and DNA elements separated by long genomic distances are known to functionally interact. This view has been further emphasized by recent observations that many transcription factors bind remotely to gene promoters. However, it is still largely unknown to us how and to what extent chromatin interactions are involved in transcription regulations on a whole genome scale. To study these questions, we have developed the Chromatin Interaction Analysis using Paired-End-Tag sequencing (ChIA-PET) strategy for *de novo* detection of genome-wide chromatin interactions, and demonstrated this approach through the comprehensive mapping of chromatin interactions involved in transcription regulations mediated by estrogen receptor α (ERα) in a human genome (Nature 2009 462: 58-64). In order to map all chromatin interactions involved in all transcription regulation networks in the human genome, we have applied the ChIA-PET strategy to active transcriptional marks such as RNA polymerase II (RNAPII) and trimethylation of lysine 4 on histone H3 (H3K4me3) as analysis targets in a number of human cells. Our results have shown that both RNAPII and H3K4me3 are excellent targets for ChIA-PET experiments to detect long-range chromatin interactions between gene promoters and distal regulatory elements, as well as to identify co-localization of remote genes (intra-chromosome and inter-chromosome) in close proximity of nuclear space. Through comprehensive mapping of chromatin interactions and transcriptional activities, we have revealed that a large proportion of actively transcribed genes are involved in extensive chromatin interaction looping structures. The most abundant gene-centric chromatin interactions are appeared to be within local range of megabase genomic span, and nearby genes such as gene family members are organized to share common transcription factories. In addition, we have identified many hot spots of interaction hubs, in which clusters of genes crossing large megabase distance and different chromosomes are co-localized in close proximity. Collectively, our data suggests that long-range chromatin interaction is a primary mechanism for transcription regulation in human genomes. Further analyses of the chromatin interaction and transcription maps will provide deep insights to advance our understanding of transcription regulatory networks and the human genome biology.

Biosketch

My primary interest is to elucidate the structures and dynamics of all functional DNA elements in complex genomes through transcriptome characterizations and genome interrogation. To facilitate such understanding we have developed Paired-End-Tag (PET) sequencing and mapping methodologies. We are applying these sequencing-based technologies to address complex biological questions such as how cancer cells progress and how stem cells maintain their unique properties. Another major interest of mine is to discover previously uncharacterized microbial genes and genomes that are relevant to human health.

Gilean McVean

Professor of Statistical Genetics
Department of Statistics
Oxford University

Title: The landscape of human genetic variation as viewed from the 1000 Genomes Project

The 1000 Genomes Project pilot generated over 5Tbp sequence in three separate pilot experiments. These assess the effectiveness of low coverage sequencing (approximately 4x) to find shared variants, a high coverage trio design, and targeted sequencing of 1000 genes in 700 samples. These data provide a comprehensive and detailed view of the

landscape of common variation of all types, from SNPs to small indels to structural variation, although much work remains in integrating all sources of information.

I will review what we have learned about the landscape of naturally-occurring variation from the project to date and what the full project expects to achieve. I will focus on characterising the nature and scale of variation across the genome, the distribution of mutations of likely functional consequence, differences between populations and the implications for our understanding of natural selection. I will also discuss some of the different approaches that are being used to interrogate the data, from mapping approaches to de novo assembly.

Biosketch

Research interests:
Population genetics, Coalescent modelling, Statistical genetics, Pathogen evolution and variation

My research covers several areas in evolutionary biology and population genetics, combining both theoretical work and empirical analyses. Of particular interest is the analysis of recombination from population genetic data, the relationship between linkage disequilibrium and properties of the underlying genealogy, and methods for inferring genealogical history from DNA sequence data. I am a co-chair of the Analysis Group in the 1000 Genomes Project and am involved in genome-sequencing projects that aim to answer question in medical genetics, the epidemiology of infectious disease and evolutionary biology.

Featured Talk (20 May 2010)

Manel Esteller

Director, Cancer Epigenetics and Biology Program (PEBC)
Leader of the Cancer Epigenetics Group

Title: Cancer Epigenetics: From DNA methylation to microRNAs

An altered pattern of epigenetic modifications is central to many common human diseases, including cancer. Many studies have explored the mosaic patterns of DNA methylation and histone modifications in cancer cells on a gene-by-gene basis, among them the seminal finding of transcriptional silencing of tumor suppressor genes by CpG island promoter hypermethylation. Epigenetic gene inactivation in transformed cells involves many "belts of silencing". We are in the process of completing the molecular dissection of the entire epigenetic machinery involved in methylation-associated silencing, such as DNA methyltransferases, methyl-CpG binding domain proteins, histone deacetylases, histone methyltransferases, histone demethylases and Polycomb proteins. The first indications are also starting to emerge about how the combination of cellular selection and targeted pathways leads to abnormal DNA methylation. In addition to classical tumor-suppressor and DNA repair genes, epigenetic gene silencing includes genes involved in premature aging and microRNAs with growth inhibitory functions. Recent technological advances are now enabling cancer epigenetics to be studied genome-wide. It is time to "upgrade" cancer epigenetics research and put together an ambitious plan to tackle the many unanswered questions in this field using genomics approaches to unravel the epigenome.

Biosketch

Manel Esteller (Sant Boi de Llobregat, Barcelona, Catalonia, Spain, 1968) graduated in Medicine with Honours from the Universidad de Barcelona in 1992, where he also obtained his Ph.D. degree specialising in molecular genetics of endometrial carcinoma, in 1996. He was an Invited Researcher at the School of Biological and Medical Sciences at the University of St. Andrews, (Scotland, UK) during which time his research interests focused on the molecular genetics of inherited breast cancer. From 1997 to 2001, Esteller was a Postdoctoral Fellow and a Research Associate at the Johns Hopkins University and School of Medicine, (Baltimore, USA) where he studied DNA methylation and human cancer. His work was decisive in establishing promoter hypermethylation of tumour suppressor genes as a common hallmark of all human tumours.

From October 2001 to September 2008 Manel Esteller was the Leader of the CNIO Cancer Epigenetics Laboratory, where his principal area of research were the alterations in DNA methylation, histone modifications and chromatin in human cancer. Since October 2008, Dr Esteller is the Director of the

Cancer Epigenetics and Biology Program of the Bellvitge Institute for Biomedical Research (IDIBELL) in Barcelona and leader of the Cancer Epigenetics Group. His current research is devoted to the establishment of the epigenome maps of normal and transformed cells, the study of the interactions between epigenetic modifications and non-coding RNAs, and the development of new epigenetic drugs for cancer therapy.

Author of more than two hundred-twenty original peer-reviewed manuscripts in biomedical sciences, he is also a Member of numerous international scientific societies, Editorial Boards and reviewer for many journals and funding agencies. Dr Esteller is also Associate Editor for Cancer Research, The Lancet Oncology and Carcinogenesis, Editor-in-Chief of Epigenetics and Advisor of the Human Epigenome Project, Associate Member of the Epigenome Network of Excellence and President of the Epigenetics Society. His numerous awards include: Best Young Cancer Researcher Award bestowed by the European School of Medical Oncology (1999), First Prize in Basic Research at the Johns Hopkins University and Medical Institution (1999), Best Young Investigator Award from the European Association for Cancer Research (2000), Young Investigator Award from the American Association for Cancer Research-AFLAC (2001), Carcinogenesis Award (2005), Beckman-Coulter Award (2006), Francisco Cobos Biomedical Research Award (2006), Fondazione Piemontese per la Ricerca sul Cancro (FPRC) Award (2006), Swiss Bridge Award (2006). National Research Award in Oncology "Maria Julia Castillo" (2007), "Dr Josep Trueta" Award by the Academy of Medical Sciences of Catalonia (2007), Innovation Award from the Commonwealth of Massachussets (2007), Human Frontier Science Program Award (2007) "Dr. Jacint Vilardell" Foundation Award (2008), DEbiopharm-EPFL Award (2009), Dr. Josef Steiner Cancer Research Award (2009), Lilly Foundation Preclinical Biomedical Research Award (2009), Fundación Esteve Award (2009), Fundación AECC Award for Children Cancer Research (2009), Carmen y Severo Ochoa Foundation, Molecular Biology Research Award(2009).

Dr Manel Esteller is the Director of the Cancer Epigenetics and Biology Program of the Bellvitge Institute for Biomedical Research (IDIBELL), Leader of the Cancer Epigenetics Group, Professor of Genetics in the School of Medicine of the University of Barcelona, and an ICREA Research Professor.

Symposium: Pathways Networks and System Biology (21 May 2010).

Hiroki R. Ueda

Project Leader, Laboratory for Systems Biology
Manager, Functional Genomics Unit
RIKEN Center for Developmental Biology, Riken, Kobe

Title: Systems Biology of Mammalian Circadian Clocks

Mammalian circadian clock system is a complex and dynamic system consisting of complicatedly integrated regulatory loops and displaying the various dynamic behaviors including **i**) endogenous oscillation with about 24-hour period, **ii**) entrainment to the external environmental changes (temperature and light cycle), and **iii**) temperature compensation over the wide range of temperature.

The logic of biological networks such as circadian clocks is difficult to elucidate without (1) comprehensive identification of network structure[1-3,6], (2) prediction and validation based on quantitative measurement and perturbation of network behavior[4,7], and (3) design and implementation of artificial networks of identified structure and observed dynamics[6]. In this symposium, I will report on the current progress of the analysis and synthesis of mammalian circadian clocks. I will also introduce some current attempts towards understanding sptio-temporal gene expression in the adult mouse brain.

References

1. Ueda, H.R. et al, *Nature* 418, 534-539 (2002).
2. Ueda, H.R. et al, *Nat. Genet.* 37, 187-92 (2005).
3. Sato T. K. et al, *Nat Genet.* 38, 312-9 (2006).
4. Ukai H. et al, *Nat Cell Biol.* 9, 1327-34 (2007).
5. Ukai-Tadenuma M. et al, *Nat Cell Biol.* 10, 1154-63 (2008).
6. Minami Y. et al *PNAS* 106, 9890-5 (2009).
7. Isojima Y. et al, *PNAS* 106, 15744-49 (2009).

Biosketch

Dr. Hiroki R. Ueda was born in Fukuoka, Japan, in 1975. He graduated from the Faculty of Medicine, the University of Tokyo, in 2000, and obtained his Ph.D in 2004 from the same university. While an undergraduate student, he worked as a research assistant on a biological simulation system project at

Sony Computer Science Laboratories. While a graduate student, he next went on to work as a researcher from 2000 and then group leader from 2002 at Yamanouchi Pharmaceutical Co.

He was appointed as a team leader of laboratory for systems biology at RIKEN Center for Developmental Biology (CDB) from April, 2003 and promoted to be a project leader at CDB from September, 2009. He was also appointed as a manager of Functional Genomics Unit at the CDB from October, 2004. He also became a visiting professor in Tohoku University in April, 2005-March, 2006, Tokushima University from April, 2005 and National Institute of Genetics from April, 2010, and an invited professor of biology in Osaka University from April, 2006, and an invited professor of mathematics in Kyoto University from April, 2009.

He identified system-level network structure of mammalian circadian clocks. For this discovery, He received Tokyo Techno Forum 21, Gold Medal (Tokyo Techno Forum 21, 2005), Young Investigator Awards (MEXT, 2006) and IBM Science Award (IBM, 2009). He recently solved the fundamental problem in chronobiology on the underling mechanism of singularity behavior of circadian clock that had been unsolved for more than 30 years. For this discovery, he received a Young Investigator Promotion Awards (Japanese Society for Chronobiology, 2007). He also recently discovered temperature-insensitive reaction in mammalian circadian clock, which will lead to a long-standing problem, "temperature-compensation" of circadian clock. He invented diagnostic method of body time and rhythm disorders, which opens up the new possibility of chronotherapy. For this invention, he received Japan Innovator Awards (Nikkei Business Publications Inc. 2004). His current research interests include system-level understanding of biological time, space and information, and systems-based medicine on human disease. His current research interests include system-level understanding of biological time, space and information, and systems-based medicine on human disease.

Eran Segal

Principal Investigator
Department of Computer Science And Applied Mathematics
Weizmann Institute of Science, Israel

Title: Transcriptional Lego: Predictable control of gene expression by manipulating promoter building blocks

The ability to control the timing and levels at which genes are expressed is key to most biological processes. Although we know the sequence preferences of key players in this process, we are still far from understanding how these elements combine within regulatory sequences to encode the transcriptional outcome. Based on our theoretical analyses, we devised hypotheses regarding the effect that different nucleosome disfavoring sequences that vary in length, composition, and distance from transcription factor binding sites, will have on the transcriptional outcome. To systematically test these hypotheses, we designed and synthesized ~80 promoter sequences, and fused each promoter to a fluorescent reporter, resulting in the largest library of designed promoter variants to date. Our results show that by manipulating either or both transcription factor binding sites and nucleosome disfavoring sequences in the vicinity of the site, we can tune expression levels in a predictable manner. Importantly, sequence changes that only alter nucleosome disfavoring sequences result in effects on expression comparable in magnitude to those that result from changes to transcription factor binding sites. In fact, compared to binding site changes, alterations of nucleosome disfavoring sequences likely yield more gradual changes in expression levels, and thus offer means to fine-tune gene expression with high resolution. These results have intriguing implications for evolution of gene expression, suggesting that sequence changes that alter the DNA-encoded nucleosome organization may provide an efficient genetic mechanism by which genomes may evolve and fine-tune gene expression. Overall, our results bring us a step closer towards understanding the role of various promoter elements and their combined effects on transcription, and suggest that directed design of promoter sequences that yield pre-specified expression patterns may be within reach.

Biosketch

My research interests are in **Computational Biology**, and more specifically in developing quantitative models for all levels of gene regulation, including transcription, chromatin, and translation.

Symposium: Nutrigenetics and Nutrigenomics (21 May 2010)

Susan E Ozanne

Principal Research Associate
British Heart Foundation Senior Fellow
Department of Clinical Biochemistry
University of Cambridge Metabolic Research Laboratories

Title: Mechanisms linking poor early nutrition and later risk of type 2 diabetes

A large number of epidemiological studies have revealed that there is a relationship between poor early growth and the development of cardiovascular disease, insulin resistance and type 2 diabetes in later life. The mechanistic basis of these relationships is not known. However, compelling evidence has emerged over the last fifteen years to suggest that early environmental factors and in particular early nutrition play an important role in mediating these relationships. Studies of individuals who were *in utero* during a period of famine, the Dutch Hunger Winter, have shown a direct relationship between maternal nutrition and glucose tolerance in the offspring in adult life. Further support for the importance of the fetal environment has come from studies of monozygotic twins who were discordant for type 2 diabetes. These revealed that the diabetic twins had significantly lower birth weights than their non-diabetic co-twins. A number of animal models have been developed to investigate the mechanisms by which the early environment determines future susceptibility to disease. These include models of maternal calorie restriction, intrauterine ligation and maternal dexamethasone treatment. The most extensively studied is the maternal low protein model where rats are fed a low (8 %) protein diet during pregnancy and lactation. The offspring have a low birth weight and undergo an age-dependent loss of glucose tolerance. This is associated with both pancreatic ß cell dysfunction and insulin resistance. The ß-cell dysfunction is accompanied by reduced expression of the transcription factor HNF4a. In vitro analysis has revealed that the insulin resistance is associated with changes in expression of key insulin-signalling proteins (including protein kinase C zeta and the p110ß catalytic subunit of PI 3-kinase) in muscle and adipocytes. Strikigly similar changes in insulin signalling protein expression are observed in muscle and adipose tissue biopsies from young men with a low birth weight. These differences in protein expression occur prior to the development of insulin resistance and type 2 diabetes, thus may be molecular markers of early growth restriction and thus disease risk. If extended to a clinically accessible tissue in humans this would make targeted intervention strategies a realistic possibility. The molecular mechanisms by which a phenomenon that occurs *in utero* has a phenotypic consequence many years later are only just starting to emerge and are thought to involve epigenetic modifications.

Biosketch

Early Programming of Appetite, Type 2 Diabetes, Breast Cancer and Ageing: The major focus of our research is to understand the mechanistic basis of the relationships between poor early growth and subsequent increased risk of type 2 diabetes, obesity, breast cancer and premature death. There are a large number of epidemiological studies suggesting that such relationships exist, however the molecular mechanisms mediating such phenomena are not understood.

Poor Early Growth and Diabetes: We carry out global and candidate expression studies at both the protein and RNA level using both rodent models and human tissues. Our aim is to identify the mechanisms by which poor early growth is linked to increased risk of type 2 diabetes and insulin resistance. In particular we are investigating the role played by the early environment.

Human studies: Tissue samples from low birth weight and control humans are used to establish insulin-signaling defects that may provide early indications of metabolic disease. Our insulin signaling protein expression studies in muscle and adipose tissue have already shown early defects in adult subjects who had a low birth weight. Ongoing studies in placenta will relate the expression of insulin signaling molecules to the nutritional status of both mother and baby.

Animal models: We are studying a rodent model of early nutritional growth restriction to identify molecular markers for prediction of risk of type 2 diabetes in later life. Nutritionally early growth restricted rats have been shown to develop impaired glucose tolerance in old age. At the molecular level, studies have shown defects in the pancreas, muscle, liver and adipose tissue of growth restricted rats. We are now extending these studies to determine the molecular mechanisms underlying these changes such as the role of epigenetic alterations.

Programming of Appetite: We have shown that appetite can be programmed by maternal nutrition during lactation and that the down regulation of appetite secondary to poor maternal nutrition is so powerful that it prevents diet-induced obesity in mice. We are currently involved in establishing the mechanisms underlying the early programming.

Animal Models: To determine the basis of appetite programming we have set up a model for poor fetal nutrition and then catch up growth in rats and mice using cross fostering techniques and altered maternal diet. In these animals, samples are taken at various time points and the expression of molecules, such as leptin and other adipocytokines, known to be involved with appetite regulation is examined. We are also determining the effect of the model on the neurodevelopment of appetite circuits using in situ hybridisation and tracing techniques.

Poor Early Growth and Breast Cancer: Epidemiological studies suggest that both low and high extremes of birthweight are associated with increased breast cancer risk. Some of the factors thought to mediate this risk are obesity and type-2 diabetes. It is also thought that an increased estrogen exposure mediates this risk in the high birth weight group. In animal studies, excessive in-utero estrogens have been shown to induce a higher mammary tumour risk and incidence. Our low-protein model is characterised by age-dependent loss of glucose tolerance, insulin resistance and type-2 diabetes. We have recently shown that maternal plasma estradiol levels are also 35% higher than controls in the last week of gestation. In the offspring, we have observed a period of retarded mammary development followed by rapid catch-up growth mainly of undifferentiated stem cells i.e structures such as terminal end buds and luminal epithelial cells. We are therefore investigating the hypothesis that fetal growth restriction followed by rapid catch-up growth increase an offspring's susceptibility to breast cancer in later life.

Oxidative Stress, Senescence and Ageing: For the past decade we have studied the long-term consequences of poor early growth using our rodent models and one of our most striking observations has been that life span can be increased or decreased by restricting their growth either during suckling or during fetal life respectively. These differences in lifespan are associated with differences in kidney telomere length. We have hypothesized that the rate of early growth may affect degrees of oxidative damage which in turn affect organ function leading to altered longevity. To test this we first investigated the effects of oxidative stress on regulation of stress response proteins, DNA replication and induction of cellular senescence using human fibroblasts. In parallel with the in vitro cell system we are actively examining the telomere length and expression of stress response proteins such as p53, p21 and DNA damage checkpoint proteins such as gama-H2AX and 53BP1 as well as senescence marker, SA-beta-gal in organs of our model animals in order to understand the molecular mechanisms underlying the ageing process.

Harris A Lewin

Professor of Immunogenetics
Gutgsell Endowed Chair
Department of Animal Sciences
Director, Institute for Genomic Biology
University Illinois, USA

Title: Nutrition: Will livestock Lead the way?

The sequencing, annotation and functional characterization of the genomes of cattle and pigs have been rapidly translated into new products and services for the livestock industry. The animal breeding industry adopted marker-assisted breeding in the late 1990s and is now implementing genomic selection for complex traits, such as milk production and carcass composition. In addition, the mapping and positional cloning of genes responsible for a large number of monogenic and polygenic traits has resulted in the creation of new models for human diseases and has informed our understanding of the genetics of complex traits. One of the most important areas of interface between human and livestock genomics is in the area of nutrition. In human and animal health, there is a dynamic interplay between host genotype and diet. However, for many diseases, the specific genes, pathways and networks that are influenced by diet are not well understood. In cattle and other livestock species it has been common practice for the past 50 years to develop special diet formulations based on breed and/or environmental conditions. Recently, we have discovered what may be the first example of a diet-regulated QTL (*DGAT1*) in dairy cattle that influences the development of metabolic disorders associated with the onset of lactation. Our data suggest that controlling energy intake during the prepartal period in *DGAT1* genotyped animals can be used to optimize milk production performance while minimizing the incidence of postpartal metabolic disorders, such as fatty liver disease and ketosis. For human diseases, understanding of diet-genotype interactions may be critical to promoting wellness in individuals genetically predisposed to diabetes, obesity, heart disease, gastrointestinal and immunological disorders and cancer. By implementing dietary restrictions in genetically-defined individuals, it may be possible to reduce the incidence and severity of such diseases. Thus, “personalized nutrition” could take on a significant role as a first line of defense against the “diseases of affluence” currently plaguing Western and developing nations. Research using non-inbred cattle and pigs as models may contribute critical knowledge that will pave the way toward the implementation of genetics-driven nutritional management of human diseases.

Biosketch

Professor Lewin's current research interests are in the area of mammalian comparative and functional genomics. His research has resulted in the development of high-density comparative maps for the cattle, human and other mammalian genomes, and novel software for in silico gene mapping

using the human genome as a template. In addition, his group produced the first large-scale cDNA microarray and oligoarray for functional genomics of ruminants, which have been applied to a systems level analysis of embryonic development and diet-genotype interactions.

Jean-Christophe Glaszmann

Director
Department of Biological Systems
CIRAD

Title: Genomics for food improvement

Genome analysis in crop species is progressing dramatically. Even crops referred to as 'orphan' until recently have their genome being sequenced. New sequencing techniques make it possible to undertake massive re-sequencing among representative crop genetic resources in order to access and exploit genetic diversity. The Centres of the Consultative Group on International Agricultural Research (CGIAR) together host more than 600,000 accessions of some twenty food crop species (cereals, pulses, roots and tubers, banana and plantain) including traditional varieties developed through many generations of selection by farmers, as well as wild species, breeding lines and improved varieties. Besides breeding activities aimed at productivity increase for global food security, international programmes are aimed at improving tolerance to abiotic stresses as well as nutritional content, principally for Zinc, Iron and B-carotene. Integrated projects for exploiting the just-available or soon-to-come genome sequence are also developing for diverse crops such as grapevine, tomato, cocoa or coffee, with their respective specific foci on product quality. Examples selected among the above initiatives will be briefly described.

Biosketch

Jean-Christophe Glaszmann is Director of the new Department of Biological Systems created in CIRAD *(Centre de coopération internationale en recherche agronomique pour le développement),* Montpellier, France, in January 2007. Agronomist and PhD in genetics, Jean-Christophe is an expert in genetic resources, comparative genomics and plant breeding. Employed by CIRAD since 1979, he worked in the Philippines in IRRI (International Rice Research Institute) for 6 years, and moved back to Montpellier in 1987. Jean-Christophe has been on the Board of Directors of Genoplante, and Director of the CIRAD/INRA/Supagro joint research unit "Polymorphisms of interest in agriculture". He is currently leader of the subprogramme "Genetic Diversity of Global Genetic Resources" of the Generation Challenge Programme from CGIAR, the Consultative Group on International Agricultural Research.

Luis Serrano

Head, Department Structural & Computational Biology, EMBL
(Germany)
ICREA Professor and Group Leader
Center for Genomic Regulation, Barcelona (Spain)

Title: Systems biology analysis of the EGF MAPK pathway: From structures to data integration

Signal transduction in living systems is essential to be able to interact and respond to the changes in the environment. The number of players and the different interactions they made result in very complex networks (specially in eukaryotes) which cannot be understood in a quantitative manner without the help of computer modeling and simulations. In order to model adequately protein interaction networks we need to incorporate time, space and structural information and move away from a node-edge flat network. Here we have used structural information, quantified all proteins in three different cell lines and localize them in a dynamic fashion to be able to produce a realistic model of the EGF-MAPK pathway that considers around 160 proteins. On top of it we have mapped human SNPs associated to disease on the structural map and analyzed them in terms of their effect on stability, activity and interaction with other proteins.

Biosketch

In our group we are aiming at a quantitative understanding of biological systems to an extent that one is able to predict systemic features and with the hope to rational design and modify their behaviour. This applies to any system comprising biological components that is more than the mere sum of its components, or, in other words, the addition of the individual components results in systemic properties that could not be predicted by considering the components individually. By achieving this objective we are aiming at new global understanding and treatment of human diseases in which the target will not be a single molecule but a network. For this purpose in our group we develop on one hand new software and theoretical approximations to understand complex systems and on the other we do experiments to validate our predictions.

Featured Talk (20 May 2010)

Larry Kedes

Scientific Director and Senor Advisor, X PRIZE Foundation
Visiting Professor, Geffen School of Medicine, University of California Los Angeles
Weston Visiting Professor, Weizmann Institute for Science

Title: Swifter, Lower, Stronger: the $10 Million Olympic Medal for Next Generation Sequencing

Dr. Laurence Kedes (BS, 1961; MD, 1962; BA (Hon), 2009) is the William M. Keck Emeritus Professor of Biochemistry & Molecular Biology and of Medicine at the University of Southern California, Keck School of Medicine. He was the founding Director of the Institute for Genetic Medicine (IGM). Prior to joining USC, he had a 20-year career on the Stanford faculty of medicine where he pioneered molecular genetics and eukaryotic gene expression, while also serving as a clinical hematologist.

Dr. Kedes's research interests have used molecular biology and molecular genetic technologies to study the differentiation of organisms and cells. He has made numerous contributions to understanding the role of gene expression in generating cellular phenotypes. He led the effort to successfully clone the first animal cell genes coding for a protein and to provide the first DNA sequence for animal protein genes. His interests in DNA sequences led him to formulate the need for bioinformatics methods to deal with such information and to collaborate in production of one of the first sets of computer programs (MOLGEN) to handle DNA sequence and other molecular genetic information. These efforts led to the creation by Dr. Kedes and others of BioNet, the first national computer network (pre-internet) funded by the NIH to archive and make available for all scientists DNA sequence information. The operation was the direct forerunner of GenBank, the international repository of all genetic sequences and helped form the foundation for the field of genomic bioinformatics.

In 2005, Dr. Kedes became the Scientific Director and Senior Advisor of the non-profit X PRIZE Foundation and helped lead the effort to create the $10 million Archon Genomics X PRIZE (http://genomics.xprize.org/). The purpose of the Archon Genomics X PRIZE competition is to develop radically new technology that will dramatically reduce the time and cost of sequencing genomes, and accelerate a new era of predictive and personalized medicine. The X PRIZE Foundation aims to enable the development of low-cost diagnostic sequencing of human genomes.

Biosketch

Dr. Laurence Kedes (BS, 1961; MD, 1962; BA (Hon), 2009), is the William M. Keck Emeritus Professor of Biochemistry & Molecular Biology and of Medicine at the University of Southern California Keck School of Medicine. He was the founding Director of the Institute for Genetic Medicine (IGM). Prior to joining USC, he had a 20-year career on the Stanford faculty of medicine where he pioneered molecular genetics and eukaryotic gene expression while also serving as a clinical hematologist.

Dr. Kedes's research interests have used molecular biology and molecular genetic technologies to study the differentiation of organisms and cells. He has made numerous contributions to understanding the role of gene expression in generating cellular phenotypes. He led the effort to successfully clone the first animal cell genes coding for a protein and to provide the first DNA sequence for animal protein genes. His interests in DNA sequences led him to formulate the need for bioinformatics methods to deal with such information and to collaborate in production of one of the first sets of computer programs (MOLGEN) to handle DNA sequence and other molecular genetic information. These efforts led to the creation by Dr. Kedes and others of BioNet, the first national computer network (pre-internet) funded by the NIH to archive and make available for all scientists DNA sequence information. The operation was the direct forerunner of GenBank, the international repository of all genetic sequences and helped form the foundation for the field of genomic bioinformatics.

In 2005 he became the Scientific Director and Senior Advisor of the non-profit X-Prize Foundation and helped lead the effort to create the $10 million Archon Genomic X-Prize for rapid, inexpensive genome sequencing.

HUGO Forum: Sequencing Futures: Sequencing Technologies Now and the Future (21 May 2010)

Radoje Drmanac

Chief Scientific Officer and Co-founder
Complete Genomics, Inc.

Title: The path to affordable human genome sequencing for large-scale studies of genetic diseases

Complete Genomics has developed a novel, highly accurate, high-throughput, inexpensive genome sequencing technology based on submicron DNA nanoball™ arrays (prepared by clonal DNA amplification in solution) and unchained base reading using combinatorial probe-anchor ligation (cPAL™) chemistry. These advances reduce both reagent consumption and imaging time and enable Complete Genomics to offer a sequencing service at its own commercial-scale genome center at a significantly lower cost than current methods for large-scale sequencing projects. The service includes reporting a list of sequence variants discovered using 120Gb of mapped reads per genome. We will present some recent results and plans to produce further high quality human genome sequences at a medically-relevant cost and scale.

Biosketch

Dr. Radoje (Rade) Drmanac, chief scientific officer and co-founder of Complete Genomics since 2006, is one of the leading research scientists and inventors in the field of DNA sequencing-by-hybridization (SBH) and genomic microarrays. In 1994, he co-founded Hyseq (later Nuvelo) where, as chief scientific officer, he led the effort to discover and patent thousands of genes which formed the basis of Nuvelo's drug development pipeline. Prior to Hyseq, Rade was a group leader at Argonne National Labs from 1991 to 1994 as part of the Department of Energy's Human Genome Project. He completed his postdoctoral studies in 1990 in Hans Lehrach's group at the Imperial Cancer Research Fund in London. He earned his Ph.D. in molecular biology for the conception and pioneering development of SBH technology from Belgrade University, where he also received B.S. and M.S. degrees in molecular biology.

Nathaniel Pearson

Bioinformatics Scientist, Knome Inc

Title: Deciphering the oracle's words: Analysis tools for human whole genomes and exomes

As sequencing costs plummet, and a tide of data rises, the thorough analysis of individual human genomes is emerging as a key bottleneck in the path to biological discovery. Making sense of such data requires comprehensive, well curated reference and phenotype data; robust parsing and interchange of 'moving target' file formats; reliable and efficient query matching; smart approaches to summary and comparative sequence analysis; innovative data visualization methods; and sound algorithms for predicting the functional significance of novel alleles, in order to pick leading candidates for empirical validation. By developing integrated toolkits to meet these challenges, small companies focused on genome analysis can usefully complement major sequencing platform providers, and serve as important partners to academic researchers in their quest to glean insights from ever finer data on human genetic variation.

Biosketch

Nathaniel Pearson is a geneticist at Knome, Inc., a human genome sequence analysis firm based in Cambridge, MA (USA), where he and his colleagues build and collaboratively deploy new analytic tools, in order to help biomedical researchers generate useful knowledge from the rising tide of human whole-genome and exome sequence data. Through laboratory, computational, and field work, he has sought to cast light on how our genomes offer ever deeper insights into human health, and reflect key aspects of our evolutionary history. Fascinated by human genetic variation since working as a young student in the laboratory of Luca Cavalli-Sforza (Stanford), he traveled with Spencer Wells (now at Cornell) to collect DNA samples from people throughout central and southwest Asia, to resolve ancestral demography and migration in those regions. In doctoral work with Bruce Lahn and Marty Kreitman (Univ. of Chicago), he surveyed and interpreted patterns of sequence variation on the great ape sex chromosomes, spotlighting the roles of mutation, recombination, and natural selection in driving long-term X-Y divergence. Later, in post-doctoral work on mammalian sensory genes with Jianzhi Zhang (Univ. of Michigan), he first delved into the emerging world of high-throughput sequencing, grasping its power to address important open questions in biology.

Marcus Droege

Global Marketing Director
Genome Sequencing
Roche Applied Science

Title: Re-sequencing of the human genome using the 454 Genome Sequencer systems FLX & Junior

The Genome Sequencer FLX Systems (GS FLX and GS Junior), powered by 454 Sequencing, are based on a next-generation DNA sequencing technology featuring a unique mix of long reads, exceptional data accuracy and shortest run time. It has been proven to be the most versatile of all currently available next-generation sequencing technologies, supporting more than 700 high profile studies in over 7 applications categories. 454 technology users have pursued innovative research in *de novo* sequencing, re-sequencing of whole genomes and target DNA regions, metagenomics, and RNA analysis. 454 Sequencing is a powerful tool for human genetics research, including re-sequencing the complete human exome or detection of low frequency somatic mutations linked to cancer.

This presentation will provide a short overview about the 454 Sequencing technology, and will focus on applications possible to address with the Genome Sequencer systems, with special emphasis on human re-sequencing (data presentation). It will also provide information on the newest product developments, including the 2010 launch of kits allowing to generate read lengths of >700 - 1000 or the first next generation benchtop instrument, the GS Junior.

Biosketch

1990 - 1995: Study of molecular biology at the University of Bielefeld

1995 - 1999: PhD study at the University of Bielefeld: Sequencing and bioinformatic analysis of antibiotic resistance plasmids

2000 - 2002: International Product Manager for genome sequencing and bioinformatics, MWG Biotech AG, Germany

2002 - 2004: Business Unit Director, sequencing department, MWG Biotech AG, Germany

2004 - 2005: Head of Marketing and Sales, MWG Biotech AG, Germany

2005 - Today: Global Marketing Director Genome Sequencing, Roche Applied Science

Patrice M Milos

Senior Vice President and Chief Scientific Officer
Helicos BioSciences Corporation

Title: An unbiased, quantitative view of genome biology only possible with Helicos Single Molecule Sequencing

Helicos Single Molecule Sequencing provides a unique view of genome biology through direct sequencing of cellular nucleic acid in an unbiased manner providing both quantitation and sequence information. The simple sample preparation involves no ligation or PCR amplification, allowing direct sequencing of DNA or RNA molecules. With this simplicity comes the opportunity to measure genome biology in ways not previously possible. We will discuss the diversity of applications possible on the HeliScope Genetic Analysis system and more importantly discuss why issues of amplification and sequence bias matter in your everyday research efforts. Further we will demonstrate applications unique to the Helicos single molecule sequencing technology providing clear differentiation from current amplification based technologies including the direct sequencing of RNA, the unique ability to sequence ancient and degraded nucleic acids and absolute quantitation of miRNAs, DNA and RNA molecules.

Biosketch

Dr. Milos joined Helicos in June 2007. Her expertise and extensive knowledge in the life sciences advances the company's efforts to develop innovative and breakthrough technology. She previously served as Executive Director at Pfizer Global Research and Development, where she was responsible for leadership, strategic alignment and execution of the scientific disciplines of pharmacogenomics, proteomics, metabonomics and RNA profiling across the Pfizer portfolio from early discovery into the marketplace. She joined Pfizer in 1993 and held numerous research positions of increasing responsibility focusing on Cardiovascular and Metabolic Disease, Pharmacogenomics, DNA Sequencing, Biomarkers and Molecular Sciences.

Dr. Milos serves on the National Advisory Council for Human Genome Research and was pivotal in the establishment and oversight of key Pfizer strategic investments in the genomics area, most notably, the Genetic Association Information Network. She also sits on several editorial boards for journals and has published and presented extensively in the genomics area. Dr. Milos conducted post-doctoral fellowships at Brown University and Harvard University. She earned her MS and PhD degrees at Rensselear Polytechnic Institute in Troy, NY and received her BA from The College of Saint Rose in Albany, NY.

Francisco M. De La Vega

Principal Research Fellow,
Computational Genomics Research,
Genetics Systems R&D,
Life Technologies

Title: The road to 99.999% accuracy single molecule sequencing

Life Technologies continues to improve the SOLiD™ System, a massively parallel second-generation sequencing system based on oligonucleotide probe ligation and clonally amplified DNA fragments attached to beads. The use of beads have provided increasing sequencing throughputs reaching 100 Gb/run today, and soon approaching 300 Gb/run by increasing packing and reducing bead dimension coupled to signal to noise improvements. Recent chemistry enhancements allow longer read lengths, providing 25-125 bp reads from either paired-end or mate-pair libraries with insert sizes ranging from ~100bp to ~20 kb. This ever-increasing throughput has reduced the cost of sequencing a human genome to US$6,000 today, to US$3,000 by the end of 2010, and continues to come down. A cornerstone of the SOLiD technology is its error detection and correction ability. Currently in SOLiD 4, error detection and reference-assisted correction codes provides an average 99.95% read accuracy, the highest in second-generation sequencing. We are working in new error corrections codes for SOLiD that can provide up to 99.999% read accuracy without the need of reference sequence, which would be essential in applications such as cancer somatic mutation detection, sequencing pooled or complex samples, and low pass GWAS by sequencing, among others. In addition, Life Technologies is actively developing a novel third generation

single-molecule sequencing platform based on real-time monitoring of FRET on the surface of a quantum-dot (Qdot™) monocrystal. This technology provides a portable nanometer-sized sequencing engine that enables continuous long read lengths using active-sequencer exchange and tunable high accuracy using recursive sequencing. The ability to rapidly and inexpensively produce very long reads can enable new applications that rely on understanding the long-range structure of nucleic acids (e.g. haplotyping) and, in combination with highly accurate "short-read" SOLiD data, would provide flexible and comprehensive genome analysis solutions.

Biosketch

Francisco M. De La Vega is Principal Research Fellow in Computational Genetics at Life Technologies in Foster City, California. He earned his Doctor of Science degree in Genetics and Molecular Biology at CINVESTAV (Mexico City), where he later was appointed assistant professor and headed the Department of Genetics Bioinformatics Unit. He joined Applied Biosystems in 1997 to lead the development of bioinformatics probe design pipelines for gene expression microarray and real time PCR genomic assays products for gene expression and genotyping, and later created the SNPbrowser™ Software, a tool to select SNPs and assays for genetic association studies. Francisco led the design and analysis of a pioneering project that genotyped over 200,000 SNPs in four human populations to develop validated genotyping assays and survey the patterns of genetic variation along the genome. In collaboration with the University of Kiel, participated in the discovery of a gene of Crohn disease by a novel approach, for which he was co-recipient of the 2008 Bio-IT World Best Practices Award in Basic Research. More recently, he managed the development of bioinformatics analysis tools for the SOLiD™ System, a second-generation massively parallel sequencing platform, and currently is working on the applications of high-throughput sequencing technologies in human genetics, cancer research, and genomic medicine. Francisco represents Life Technologies at the 1000 Genomes Project steering committee.

Stephen Turner

Founder & Chief Technology Officer
Pacific Biosciences

Title: Applications of SMRT™ sequencing outside the performance envelope of first and second generation sequencing

Despite their nearly universal use in DNA sequencing for several decades, DNA polymerases possess performance characteristics far beyond what first and second generation sequencing technologies have achieved. Through the use of phospholinked nucleotides, which can be incorporated with near-native kinetics by DNA polymerases, and zero-mode waveguides, which allow single-molecule detection at the high nucleotide concentrations required by native DNA replication, the natural power of these enzymes has been harnessed by Pacific Biosciences for single-molecule, real-time (SMRT™) DNA sequencing, which exhibits long read lengths, high speed and thus fast time to results. The high processivity of certain phage polymerases is translated into long readlength and applied to the shotgun whole genome assembly of a hydrogen-producing bacterium, *Rhodopseudomonas palustris (R. pal)*. To demonstrate the effectiveness of long reads and strobed reads in resolving complex genome structure, we employed a hybrid assembly strategy using Pacific Biosciences' long reads and strobe reads in conjunction with reads from the Illumina sequencing platform. The resulting assembly produced significantly fewer contigs and longer contig lengths than the starting Illumina assembly, with no misassembles when compared to the finished genome sequence. We show the role these assemblies play in the network analysis of hydrogen production by *R. pal*. Because of the kinetic information that accompanies the primary sequence extracted with every read, it is possible to detect the presence of modified nucleobases in DNA strands. In addition to allowing direct detection of epigenetic marks such as methylcytosine, it also provides a means of performing hypothesis-free investigation of structural modifications to DNA bases at the whole-genome level. The fast time to result is ideal for analysis of viral genomes in the context of outbreaks of infectious disease. To demonstrate the agility of the system we apply SMRT sequencing to raw samples taken from influenza patients to analyze the strains present at the single molecule level.

Biosketch

Dr. Turner founded Pacific Biosciences (formerly Nanofluidics) and secured its Series A funding in 2004. He was awarded a Ph.D. in Physics by Cornell University in 2000, where he worked with Prof. Harold Craighead to study the behavior of biomolecules in nano-fabricated structures. His work

contributed to the establishment of the Nanotechnology Center at Cornell. He was a member of the project team at Cornell which developed the technology now employed by Pacific Biosciences and was co-author of the cover story in *Science* magazine (January 31, 2003) that introduced the technology to the scientific community. Dr. Turner's undergraduate work was at the University of Wisconsin, Madison, where he received a Bachelor of Science in Applied Mathematics, Electrical Engineering and Physics. He is the author of over 30 scientific papers in fields ranging from nanofluidics, genetics, cell attachment to chemically- and topographically- modified surfaces, x-ray lithography and process modeling. He is listed as the inventor on nine U.S. patents and more than 20 published patent applications. Dr. Turner was recipient of the *MIT Technology Review* "TR100" Award in 2003 and the University of Wisconsin Madison Distinguished Young Alumnus Award in 2008. He is a sitting member of the National Institutes of Health grant review study section on new technologies. He oversees the scientific and technical direction of Pacific Biosciences and is a member of its Board of Directors.

Mostafa Ronaghi

Senior Vice President & Chief Technology Officer
Illumina Inc

Title: Current and future outlook of genomic technologies

Recent advancements in genomic technologies are changing the scientific horizon, dramatically accelerating biomedical research. For wide implementation of these technologies, their accuracy, throughput, cost, and workflow need to be addressed. In the past seven years, the cost of full human genome sequencing has been reduced by 4-5 orders of magnitude, and reduction will continue by another 10-fold in the next couple of years. At that cost level, major biomarker discoveries will fuel adoption of the latest genomic technologies for medicine, agriculture, consumer products, and forensics. In this talk, we will discuss Illumina's efforts and provide our perspectives on the future of genomics.

Biosketch

Mostafa Ronaghi, Ph.D., joined Illumina in August 2008 and is responsible for leading internal research programs and evaluating new technologies for the Company.

Mostafa is an experienced entrepreneur and was involved in the start-up of four life sciences companies. In 2007, Mostafa co-founded Avantome, a privately-held sequencing company. Before this, he co-founded NextBio, a search engine for life science data. In 2001, Mostafa co-founded ParAllele Bioscience, which was eventually acquired by Affymetrix, Inc., and was involved in the development and commercialization of highly multiplexed technology for genetic testing. In 1997, he co-founded Pyrosequencing AB, which was renamed to Biotage in 2003, and led the company to a successful initial public offering in June 2000 on the Stockholm Stock Exchange. Mostafa was a principal investigator at Stanford University from 2002–2008 and focused on the development of novel tools for molecular diagnostic applications. He serves on the board of directors of Microchip Biotechnologies, NextBio, and Aurora Biofuels.

Mostafa earned his Ph.D. from the Royal Institute of Technology in Sweden. Mostafa holds more than 20 pending and issued patents and has written more than 50 peer-reviewed publications in journals and books.

Exhibitor Showcase

Booth 2, 4, 6

ILLUMINA UNITED KINGDOM
Chesterford Research Park
Little Chesterford
Nr Saffron Walden
Essex
CB10 1XL, UK
Contact Person: Colette King
Direct: +44(0)1799 534 172
Tel: +44(0)1799.532.300
Fax: +44(0)1799.532.301
CKing@illumina.com
www.illumina.com

At Illumina, our goal is to apply innovative technologies and revolutionary assays to the analysis of genetic variation and function, making studies possible that were not even imaginable just a few years ago. With such rapid advances in technology taking place, it is mission critical to have solutions that are not only innovative, but flexible, scalable, and complete with industry-leading support and service. Our offering includes leading-edge solutions for:

- DNA sequencing
- SNP genotyping
- Copy number variation
- DNA methylation
- Gene expression
- Low-multiplex analysis of DNA, RNA, and protein

Booth 3, 5

BGI (Beijing Genomics Institute at Shenzhen)
Bei Shan Industrial Zone,
Yantian District, Shenzhen
518083, China
Contact Person: Wang Wei Wei
Tel: (+86) 755 25273291
wangweiwei@genomics.org.cn
http://www.genomics.cn/en/index.php

Carrying out lots of projects, BGI has established its own technical platforms based on large-scale genome sequencing, efficient bioinformatics analyses, and innovative genetic health-care initiatives. These distinguished achievements have made a great contribution to the development of genomics in the world, and have established BGI as a world-class research institution.

Booth 7

DNA Vision S.A
Avenue George Lemaitre 25 - B-6041
Charleroi – Belgium
Contact Person: Elise Debel
Direct: (+32) 071347884
Tel: (+32) 071347880
Fax: (+32) 071347879
EliseDebel@dnavision.be
www.dnavision.be/

DNAVision, specialized in molecular biology, provides a wide range of genetic services and products to the pharmaceutical, food, and biotechnology industries.Our scientists use theoretical knowledge and practical know-how to discover and develop genetic markers meeting biological and medical needs. DNAVision is a research and development company with expertise in molecular biology, specifically in genetic analysis. Using state-of-the-art instrumentation and methods validated in a quality environment, DNAVision provides analytical services designed for the health and agro-alimentary industries.

Booth 8

GenVault Corporation
6190 Corte Del Cedro, Carlsbad,
CA 92011-1515
USA
Contact Person: Mimi Tran
Direct: 17602685247
Tel: 17602685200
mtran@genvault.com
http://www.genvault.com/index.html

GenVault is the global leader in providing biosample workflow, transport, and storage solutions for genomic medicine, discovery, and identification. The company's dry-state platform enables the extraction, preservation, recovery, and distribution of DNA and RNA at ambient temperature. With GenTegra, GenPlates, and the Dynamic Archive, GenVault is innovating best practices for sample management and preservation. Learn more at www.genvault.com.

Booth 9, 11

AGILENT TECHNOLOGIES
1 rue Galvani
91745 Massy cédex
France
Tel: (+33)478778442
redia_traore@non.agilent.com
www.agilent.com

Agilent Technologies is a leading supplier of life science research systems that enable scientists to study complex biological processes and disease mechanisms. Engineered for sensitivity, reproducibility and workflow productivity, Agilent's solutions include instrumentation, software, consumables and services for genomics, proteomics and metabolomics applications.

Agilent's genomics portfolio features:

- Stratagene molecular biology reagents
- 60-mer Oligo catalog and custom microarrays for array-based Comparative Genomic Hybridization (aCGH), GeneExpression, miRNA, methylation and ChIP-on-Chip analysis
- SureSelect Target Enrichment products for next-generation sequencing
- a microarray scanner
- a Quantitative PCR platform
- the 2100 Bioanalyzer, leading microfluidic platform for DNA, RNA and protein quality control
- Liquid handling platforms
- Genespring software for microarray and qPCR data analysis

From sample preparation to data analysis, Agilent provide complete solutions optimizing workflow productivity and performance through automation.

To learn more about Agilent's genomics products and/or find the contact details of your local sales representative, please visit http://www.genomics.agilent.com.

Booth 10

Luminex B.V.
Krombraak 13-15
4906 CR OOSTERHOUT
THE NETHERLANDS
Tel: +31 162 408 333
Fax. +31 162 408 337
info@luminexcorp.com
www.luminexcorp.com

Luminex Corporation develops manufactures and markets proprietary biological testing technologies with applications throughout the diagnostics and life sciences industries. The Company's xMAP® multiplex solutions include an open architecture, multi-analyte technology platform that delivers fast, accurate and cost-effective bioassay results to markets as diverse as pharmaceutical drug discovery, clinical diagnostics and biomedical research, including the genomics and proteomics markets. The Company's xMAP technology is sold worldwide and is already is use in leading clinical laboratories

as well as major pharmaceutical, diagnostic and biotechnology companies. Further information on Luminex Corporations or xMAP can be obtained at www.luminexcorp.com

Booth 12
LGC Genomics (AGOWA)
TGS Haus 8
Ostendstr. 25 D-12459
Berlin Germany
Contact Person: Franziska Hanke
Tel: (+49) 03053042204
Fax: (+49) 03053042201
hanke@agowa.com
http://www.agowa.de/

LGC Genomics (www.lgcgenomics.com) is the Berlin-based division of LGC with specific expertise in the delivery of a wide range of genomic services and products to a variety of customers. These include pharmaceutical and biotechnical companies, private and public research organisations and medical institutions.

LGC Genomics delivers Next Generation Sequencing services, DNA sequencing, genomics services, nucleic acid extraction services and DNA extraction products.

LGC is an international science-based company and market leader in analytical, forensic and diagnostic services and reference standards.

Booth 13, 14. 15

The Centre for Arab Genomic Studies (CAGS)
Centre for Arab Genomic Studies
P.O. Box 22252, Dubai
United Arab Emirates
Contact Person: Moncy Philip
Tel: (+971)43986777
Fax: (+971)43980999
cags@emirates.net.ae
www.cags.org.ae

Since its establishment in 2003 by H.H. Sheikh Hamdan Bin Rashid Al Maktoum, Deputy Ruler of Dubai and UAE Minister of Finance, with an aim to alleviate human suffering from genetic diseases in the Arab World, the Centre For Arab Genomic Studies has dedicated itself to the mission of improving human health by characterizing and preventing genetic disorders in Arab countries based on the recent advances in human genetics.

One of the major projects undertaken by CAGS is the Catalogue of Transmission Genetics in Arabs (CTGA) database, which is a continuously updated, open-access compendium of bibliographic material and observations on human gene variants and inherited, or heritable, genetic diseases in Arabs (www.cags.org.ae). Data collections on genetic disorders have so far been completed for the UAE, Bahrain, Oman and Qatar, work is ongoing for other Arab countries. Currently, the database hosts entries for nearly 940 genetic disorders and 370 related genes.

CAGS regularly comes out with many open access publications, aimed at both the scientific community as well as the public. CAGS conducts the Pan Arab Human Genetics Conference (PAHGC) is a biennial event organized every alternate year. CAGS is active in conducting cutting edge research in the field of genetic disorders seen in the Arab World.

The support of H.H. Sheikh Hamdan Bin Rashid Al Maktoum has been instrumental in achieving most of the Centre's objectives. This support has been further reinforced with the establishment of a CAGS genetics award to honor individuals and institutions working in the field with an aim to encourage genetic research in the Arab World.

Booth 16. 17, 18

Roche
Swisslab GmbH
Pascalstr. 10
D - 10587 Berlin
Deutschland
Contact Person: Claudia Moeller
Tel: (+49)6217591317
claudia.moeller@roche.com
http://www.roche.com/index.htm

Roche provides state-of-the-art instruments and reagents for research. Discover novel variants in complex genomes with unprecedented sensitivity and specificity using the new GS Junior Benchtop System – sequencing for everyone and everyday – and GS FLX. MagNA Pure 96s purify nucleic acids from 96 samples in <1 hour, LightCycler® 1536s revolutionize real-time qPCR, and NimbleGen arrays find your gene of interest. For cell analysis, RTCA DP Instruments measure cell invasion and migration in up to three 16-well CIM-Plates and Innovatis Instruments automate cell counting using validated processes.

LIGHTCYCLER, GS FLX, GS JUNIOR, MAGNA PURE, XCELLIGENCE and INNOVATIS are trademarks of Roche.

Booth 19

DNA Genotec Inc
2 Beaverbrook Road
Kanata, Ontario, Canada
K2K 1L1
Contact Person: Danielle Hodgson
Tel: 613 723 5757
danielle.hodgson@dnaenotek.com
http://www.dnagenotek.com/index.html

DNA Genotek is focused on improving nucleic acid sample collection. The company's Oragene® product line offers researchers a non-invasive, all-in-one system for the collection, stabilization, transportation and purification of high quality DNA or RNA from saliva. Oragenes' reliability and ease-of-use have resulted in rapid worldwide adoption by top-tier health institutions.

Booth 20

Caliper Life Sciences
Central Parc
1 Allée de l'Epervier
93420 Villepinte
Contact person: Abderazzak Bensaid
Tel: +33 1 48 63 71 35
Fax: +33 1 48 63 71 33
Abderazzak.bensaid@caliperls.com
www.caliperLS.com

Caliper Life Sciences is a premier provider of cutting-edge technologies enabling researchers in the life sciences industry to create life-saving and enhancing medicines and diagnostic tests more quickly and efficiently. Caliper is aggressively innovating new technology to bridge the gap between *in vitro* assays and *in vivo* results and then translating those results into cures for human disease. Caliper's portfolio of offerings includes state-of-the-art microfluidics, lab automation and liquid handling, optical imaging technologies, and discovery and development outsourcing solutions.

Caliper has a broad expertise in automating the processes involved in the core areas of genomics research including: extraction and purification, PCR reaction setup, automation of RT-PCR as well as the separation and direct detection of DNA and RNA. Caliper's new LabChip XT has the ability to quickly, and with high resolution isolate nucleic acid fragments of specific size. The LabChip XT shares technology with Caliper's current separation instruments, but in addition to analysis it can actually collect sample for processing. The applications for this are numerous, including microRNA sequencing, transcriptome analysis, SNPs, and de novo sequencing

For more information please visit www.caliperLS.com

Booth 21
Affymetrix
Wycombe, High Wycombe,

Buckinghamshire
HP10 OHH
UK
Contact Person: Vicky Bannes
Tel: 44 (0) 1628 552 556
Fax: 44 (0) 1628 552 580
vicky_barnes@affymetrix.com
http://www.affymetrix.com/estore/

Affymetrix is evolving into a provider of scalable, innovative genomic analysis tools and reagents for discovery, exploration, validation, and genetic testing. The acquisitions of Panomics and USB bring high-throughput, multi- to single-gene assays and premium-value molecular biology reagents to customers, enabling complete solutions for genome-wide and targeted studies and a range of products for cellular and protein analysis.

Today, Affymetrix technologies enable researchers to better understand the role that genes play in disease, the effectiveness and safety of therapies, and many other biological factors that affect human well-being. Our mission is to revolutionize how the world benefits from genetic information.

Booth 22, 23

Springer
Springer-Verlag GmbH
Tiergartenstraße 17
69121 Heidelberg
Germany
Contact Person: Brigit Jung
Tel: +49 062214878913
Fax: +49 0622148768913
birgit.jung@springer.com / Nicola.Klupsch@springer.com
http://www.springer.com

Springer is a major publisher of books and journals in genomic medicine, including The HUGO Journal, the Official Journal of the Human Genome Organisation. Please stop by our booth to order books at a special conference discount and take a closer look at sample issues of journals. Staff will be on hand to answer any questions you might have about publishing with Springer

Booth 24

Helicos Biosciences
One Kendall Square
Ste. 7301
Cambridge, MA 02139
USA
Contact person: Sharon
Tel: 617 264 1857
Fax: 617 264 1757
sbleakney@helicosbio.com
http://www.helicosbio.com/

Helicos BioSciences (NASDAQ: HLCS) is a life science company focused on innovative genetic analysis technologies for the research, drug discovery, and diagnostic markets. Helicos's proprietary True Single Molecule Sequencing (tSMS)™, technology allows direct measurement of billions of strands of DNA enabling scientists to perform experiments and ask questions never before possible. Helicos is a recipient of the $1,000 genome grant and committed to providing scientists the tools to unlock the era of genomic medicine. The company is currently planning the introduction of its first product, the Helicos™ Genetic Analysis System.Based on Helicos's tSMS technology, the Helicos Genetic Analysis System enables ultra-highthroughput genetic analysis by directly sequencing single molecules of nucleic acids. The tSMS workflow is based on a simple, cost effective sample preparation process that can easily be scaled and replicated to meet the requirements of large, complex experiments, overcoming laboratory workflow bottlenecks.

Booth 25

IMGT
IMGT, Laboratoire d'ImmunoGénétique Moléculaire,
LIGM Institut de Génétique Humaine, IGH, UPR CNRS 1142
141 rue de la Cardonille 34396 MONTPELLIER Cedex 5, France
Contact Person: Marie-Paule Lefranc
Tel: +33 499619965
Fax: +33 499619901
Marie-Paule.Lefranc@igh.cnrs.fr
http://imgt.cines.fr/

IMGT®, the international ImMunoGeneTics information system® (http://www.imgt.org), created in 1989 by Professor Marie-Paule LEFRANC at Montpellier, France (Montpellier 2 University and CNRS), is the global reference in immunogenetics and immunoinformatics. Based on IMGT-ONTOLOGY, IMGT® is a high-quality integrated knowledge resource specialized in the immunoglobulins (IG) or antibodies, T cell receptors (TR), major histocompatibility complex (MHC), of human and other vertebrate species, proteins of the immunoglobulin superfamily (IgSF) and MHC superfamily (MhcSF), and related proteins of the immune system (RPI) of any species. IMGT® comprises six databases (IMGT/LIGM-DB, IMGT/GENE-DB, IMGT/3Dstructure-DB, IMGT/mAb-DB, etc.), fifteen interactive on-line tools for sequences, genome and three-dimensional (3D) structure analysis, and more than 10 000 web pages of synthesis and knowledge. IMGT® is extensively used by scientists in medical research (autoimmune diseases, infectious diseases, AIDS, leukemias, lymphomas, myelomas and other cancers), veterinary research, biotechnology related to antibody engineering and

humanization (phage displays, combinatorial libraries, chimeric, humanized and human antibodies), diagnostics (clonalities, detection and follow-up of residual diseases) and therapeutical approaches (graft, immunotherapy, vaccinology). IMGT® is available at http://www.imgt.org, freely for the academic research, and by contracts and licences for pharmaceutical companies.

Booth 26

Knome
Knome, Inc.
1 Main Street, Suite 530
Cambridge, MA 02142
USA
Contact Person: Ari Kiirikki
Tel: 978 7714876
Fax: 617 250 8610
akiirikki@knome.com
http://www.knome.com/service/knomediscovery.html

Saving you from making significant platform commitments, Knome provides rapid and economical access to the latest cutting-edge sequencing platforms, robust analytical capabilities and experienced bioinformatic professionals. Whether your need is sequencing one or one hundred samples, Knome can make your next whole genome or whole exome sequencing project fast, informative, and affordable.

Booth 27

BIOKÉ
Plesmanlaan 1d
2333 BZ Leiden
The Netherlands
Contact Person: Linsay Chung
Tel: +31 715681000
Fax: +31 715681010
l.Chung@bioke.com
http://www.bioke.com/

BIOKÉ is a key provider of products and services in life sciences. Many experts in academic institutions, hospitals and research companies as well as pharmaceutical and biotechnology companies are using our products to speed up their research and diagnostics. It is our aim to accelerate your progress in R&D and diagnostics by sharing knowledge.

We provide innovative products and services in the field of genomics, proteomics and molecular diagnostics. In these field we collaborate with experts from the University of Leiden, and other universities or research institutions. The mission of BIOKÉ is defined by four keywords: acceleration, co-operation, innovation and knowledge.

Booth 28

Fluidigm Europe B.V.
Parnassustoren, Locatellikade 1
1076 AZ Amsterdam
The Netherlands
Contact person: Eve Marquis
Tel : +31 20 578 88 53
Fax: +31 20 203 1111
eve.marquis@fluidigm.com
www.fluidigm.com

Bridging the gap between discovery and clinically useful biomarkers requires next-generation technology—tools that deliver radically new efficiencies for the analysis of nucleic acids. The BioMark™ line of life-science products from Fluidigm is meeting this challenge through unique microfluidic chips known as Dynamic and Digital Arrays. Using the recognized gold standard in PCR, these innovative products utilize TaqMan® or similar assays to deliver significant advancements in throughput and time to data while reducing reagent costs and the tyranny of pipetting. In addition, Fluidigm's exclusive Digital Arrays enable exquisite quantification of template libraries and detection of rare mutations as described in recent breakthrough publications. Discover the *New* 48.48 AccessArray™ for target enrichment and sample multiplexing.

Fluidigm develops, manufactures and markets proprietary Integrated Fluidic Circuit (IFC) systems that significantly improve productivity in life science research. Fluidigm's IFCs enable the simultaneous performance of thousands of sophisticated biochemical measurements in extremely minute volumes. These "integrated circuits for biology" are made possible by miniaturizing and integrating liquid handling components on a single microfluidic device

(IFC). Fluidigm's systems, consisting of instrumentation, software and single-use chips, increase throughput, decrease costs and enhance sensitivity compared to conventional laboratory systems.

Booth 29

GATC Biotech SARL
Parc scientifique et technologique de Luminy
Bat CCIMP Case 922
13288 Marseille cedex 09,
France
Contact Person: Marita Strohm
Tel: +33 04918284
Fax: +33 0442011200
m.strohm@gatc-biotech.com
http://www.gatc-biotech.com/

GATC Biotech is Europe's leading service provider of DNA sequencing and bioinformatics for industry and academic research.

For nearly two decades we have been continuously developing new and comprehensive solutions for DNA sequencing and bioinformatics sequence evaluation – for individual DNA samples through to complete genome projects.

All our solutions are precisely matched to the specific requirements of the customer. We can do this because we use all leading sequencing technologies on the market, such as Applied Biosystems ABI 3730xlTM for Sanger sequencing and Illumina's Genome Analyzer IITM and Roche Diagnostics GS FLXTM for Next Generation sequencing.

With a total of eleven sequencing instruments – four for Next Generation sequencing - GATC has the largest commercially available sequencing capacity in Europe with two terabases per year.

We serve over 10,000 customers in 40 countries around the globe and have subsidiary companies in Great Britain, France and Sweden – the fact that GATC Biotech is one of the top companies in this sector is not only down to technical or scientific reasons, however:

Booth 30

Malaysian Genomics Resource Center(MGRC)
27-9, Level 9, Signature Office,
Mid-Valley City,Kuala Lumpur
59200, Malaysia
Contact Person: Mazlina Binti Mat Ali
Tel: +6 03 2283 3401
Fax: +6 03 2282 8102
mazlina@mgrc.com.my
www. mgrc.com.my

We aspire to be the leading provider of bioinformatics services, offering high-end and rapid turnaround services that are valued by our customers. We hope that our services will expedite scientific breakthroughs and IP generation in the medical, agricultural and industrial sectors, and that these discoveries benefit mankind. We aim to stay ahead in this exciting field by continuously expanding, updating, and improving our services, resources and expertise.

We want to help create awareness of the enormous potential bioinformatics holds, and to help strengthen the understanding of its underlying principles and applications among Malaysians. Through this, we are committed to promoting Malaysia as a leader on the global bioinformatics and genomics stage.

Booth 31

Eurofines MWG GmbH
Anzingerstr. 7a, Ebersberg
85560, Germany
Tel: +49 80928289959
Fax: +49 809221084
npichler@eurofinsdna.com
http://www.eurofinsdna.com

Eurofins MWG Operon, founded in 1990 and member of the Eurofins Group, is an international provider of genomic services established around the core business lines DNA sequencing, oligonucleotides, siRNA and gene synthesis. The company's main mission is focussed on customer convenience and high quality services in industrial scale for the life science industries and academic research institutions around the world.

Booth 32, 33, 34

Life Technologies
25 Avenue de La Baltique BP 96
91943, Courtaboeuf Villebon sur Yvette,
France
Tel: +33 0 169598539
Fax: +33 0 169598500
corinne.malphettes@lifetech.com
www.lifetechnologies.com/home.html

Applied Biosystems is now a part of Life Technologies Corporation (NASDAQ:LIFE), a global biotechnology tools company dedicated to improving the human condition. Applied Biosystems systems, consumables and services enable researchers to accelerate scientific exploration, driving to discoveries and developments that make life even better. Life Technologies customers do their work across the biological spectrum, working to advance personalized medicine, regenerative science, molecular diagnostics, agricultural and environmental research, and 21st century forensics. Life Technologies had sales of more than $3 billion in 2008, employs approximately 9,500 people, has a presence in more than 100 countries, and possesses a rapidly growing intellectual property estate of approximately 3,600 patents and exclusive licenses. Life Technologies was created by the combination of Invitrogen Corporation and Applied Biosystems Inc. For more information on how we are making a difference please visit our website: www.lifetechnologies.com.

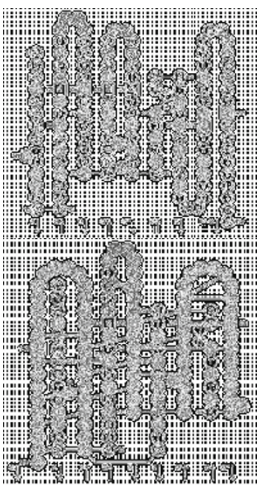
IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org

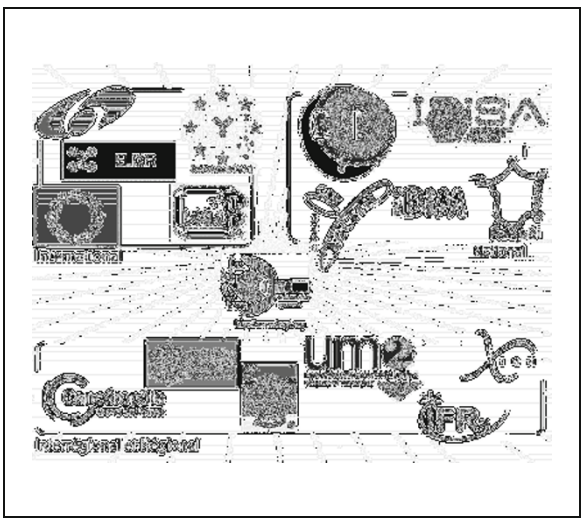
IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org, is a high-quality integrated knowledge resource specialized in the immunoglobulins (IG), T cell receptors (TR), major histocompatibility complex (MHC), immunoglobulin superfamily (IgSF), major histocompatibility complex superfamily (MhcSF) and related proteins of the immune system (RPI) of human and other vertebrate species, created in 1989 by Marie-Paule Lefranc (Université Montpellier 2, CNRS).

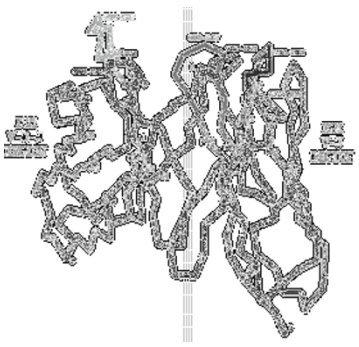
IMGT® consists of

- 6 databases (sequences, genes, structures, monoclonal antibodies)
- 15 interactive online tools
- 10 000 pages of web resources.

The IMGT Home page http://www.imgt.org (Montpellier, France) provides a common access to all Immunogenetics data.







INTERNATIONAL PANEL OF CHAIRS

Stylianos E Antonarakis
Professor and Chairman
University of Geneva Medical School
Department of Genetic Medicine and Development

Stylianos Antonarakis is the Director of the Division of Medical Genetics at the University of Geneva in Switzerland. Before moving to Geneva he was a professor at Johns Hopkins University in the USA. His lab participates in many projects involving the genetics of monogenic and polygenic disorders as well as the annotation of the human genome and particularly of human chromosome 21. He is on the editorial board of many high profile journals such as Genome Research and Genomics and has recently served as president of the European Society for Human Genetics. He has published more than 470 papers in peer-reviewed journals since 1982.

Samir K Brahmachari
Secretary to Govt. of India,
Department of Scientific and Industrial Research and Director General, CSIR

Prof. Samir K. Brahmachari pioneered functional genomics initiative in India and led the Indian Genome Variation Consortium project as the Director, Institute of Genomics and Integrative Biology, CSIR. His core expertise is in structural and computational biology. Since November 2007, he has assumed the office of the Director General of Council of Scientific and Industrial Research (CSIR), which is the largest publicly-funded organization involved in scientific and industrial research, with 37 constitutional laboratories across India. He is a member of the HUGO Council and the Advisory Board of the X Prize in Genomics. As a member of the expert group on Human Rights and Biotechnology Commission of United Nations, he has addressed issues of unethical exploitation of genetic resources of the Third World and has championed the concept of rights of patients in benefit sharing in the development of genomic medicine. He has successfully led the Indian team in the Pan Asia SNP Project. He is the recipient of a large number of National and International awards and elected member of all four National Academies of Sciences and Engineering in India. In 2009, the *Week* magazine has identified him as one of the twenty-five most valuable Indians. Currently he is leading the Open Source Drug Discovery (OSDD) project and Pharmacogenomics for affordable health programme in India, with global participation.

Jean-Marie Blanchard
Principal Investigator, French National Center for Scientific Research (CNRS)
Director, Intitute of Molecular Genetics of Montpellier
Director, Montpellier Institute of Biology

Dr Blanchard (PhD) is Director of the Institute of Molecular Genetics in Montpellier (IGMM, www.igmm.cnrs.fr) and of the Montpellier Institute of Biology (IFR122), a cluster of laboratories belonging to the department of Biology and health of Montpellier University. He is Vice-President of the Cancéropôle GSO Administrative Council, Board member of the Scientific Committee of the National League Against Cancer (LNCC) and of the Montpellier Cancer Initiative.

Philippe Broet
Group Leader, Genome Institute of Singapore
Director of the JE 2492 University Team (University Paris-Sud)

Our team develops new statistical models and methods devoted to genomic-oriented data analyses and transfers the results of this methodological research to bioclinical research. Three topics are investigated by the team that relate to class comparison, class prediction and class discovery. In particular, the team focuses on clinical cancer research and investigates the prognostic interest of combining markers from genome, transcriptome, and proteome for predicting clinical outcome. More precisely, we recently proposed an integrative genomic signature for refining the assessment of recurrence risk in early stage of non-small cell lung cancer. Our team is also working on the genetic susceptibility of HIV progression.

Anthony Brookes
Chair
Bioinformatics and Genomics
Department of Genetics
University of Leicester

Professor Brookes is an internationally recognized leader in genomics and bioinformatics, having made significant contributions to method development, disease and population genetics, DNA variation analysis, and database systems for organizing gene-disease information. His publications include over 130 peer-reviewed articles and reviews, and he has filed 6 patent applications on cloning, genotyping, and DNA analysis. Professor Brookes began his research career in the UK by achieving a first class degree in Medicine and Medical Biochemistry at Manchester University, and then a molecular genetics PhD at the Imperial Cancer Research Fund and the University of London. He subsequently established a research group in the MRC Human Genetics Unit in Edinburgh, where his team was involved in identifying the Presenilin-1 gene that causes Alzheimers Disease. He then moved to Sweden, first to Uppsala University where he received Docenture in Genetics, and then to the Karolinska Institute where he was awarded both Docentureand a Professorship in Genome Research. Professor Brookes now holds a Chair in Bioinformatics and Genomics at the University of Leicester in the UK, and an Adjunct Professorship in the KarolinskaInstitute in Sweden. In these positions he coordinates the European FP7-GEN2PHEN informatics project, and runs the HGVbaseG2P genetic association database. He serves on the HUGO Council and acts as Editor & Chair

of Publications Committee for the HUGO Journal, and he is also a member of both the Public Population Project in Genomics (P3G) and the Open Researcher & Contributor ID (ORCID) Technical Working Group. Additionally, he is a Co-Founder and Board Member of the Human Genome Variation Society, and works as communicating Editor for Human Mutation. Other significant achievements include launching and running the international Meeting series on "Human Genome Variation & Complex Genome Analysis" and inventing the Dynamic Allele-Specific Hybridization genotyping technology.

Elspeth Bruford
Project Co-ordinator
European Bioinformatics Institute (EMBL-EBI)
HUGO Gene Nomenclature Committee (HGNC)

Professor Brookes is an internationally recognized leader in genomics and bioinformatics, having made significant contributions to method development, disease and population genetics, DNA variation analysis, and database systems for organizing gene-disease information. His publications include over 130 peer-reviewed articles and reviews, and he has filed 6 patent applications on cloning, genotyping, and DNA analysis. Professor Brookes began his research career in the UK by achieving a first class degree in Medicine and Medical Biochemistry at Manchester University, and then a molecular genetics PhD at the Imperial Cancer Research Fund and the University of London. He subsequently established a research group in the MRC Human Genetics Unit in Edinburgh, where his team was involved in identifying the Presenilin-1 gene that causes Alzheimers Disease. He then moved to Sweden, first to Uppsala University where he received Docenture in Genetics, and then to the Karolinska Institute where he was awarded both Docentureand a Professorship in Genome Research. Professor Brookes now holds a Chair in Bioinformatics and Genomics at the University of Leicester in the UK, and an Adjunct Professorship in the KarolinskaInstitute in Sweden. In these positions he coordinates the European FP7-GEN2PHEN informatics project, and runs the HGVbaseG2P genetic association database. He serves on the HUGO Council and acts as Editor & Chair of Publications Committee for the HUGO Journal, and he is also a member of both the Public Population Project in Genomics (P3G) and the Open Researcher & Contributor ID (ORCID) Technical Working Group. Additionally, he is a Co-Founder and Board Member of the Human Genome Variation Society, and works as communicating Editor for Human Mutation. Other significant achievements include launching and running the international Meeting series on "Human Genome Variation & Complex Genome Analysis" and inventing the Dynamic Allele-Specific Hybridization genotyping technology.

Alain Bucheton
INSTITUTE of HUMAN GENETICS
CNRS standalone unity (UPR 1142)
Director

Valere Cacheux-Rataboul
Dr Valere CACHEUX, PharmD, PhD
Principal Investigator
Genome Institute of Singapore

Our team has for objectives to identify and characterize the genomic architecture of chromosomal rearrangements in genomes. In cancer genomes these rearrangements can lead to fusion genes and could be used as markers in tumor screening. Then, chromosomal anomalies associated with congenital disorders can aid the identification of new genetic syndromes. Using known technologies such as FISH and microarrays in combination with DNA-PET technology, we can also precisely identify all the breakpoints of chromosomal anomalies and understand the genomic organization of such rearrangements. With the precise description of the breakpoints, a second focus is to understand why these rearrangements occur in cancer genomes and in associated diseases rearrangements.

Benjamin Capps
Assistant Professor in Biomedical Ethics
Director of Graduate Studies, Centre for Biomedical Ethics

Benjamin joined the Centre for Biomedical Ethics in 2008. He was previously a Research Associate at the Centre for Ethics in Medicine, University of Bristol, UK. He read for a Doctorate in Medical Ethics (awarded in 2004) and completed a two-year Postdoctoral Fellowship at the University of Bristol (2004-6), both funded by the Wellcome Trust's Biomedical Ethics Programme. He has coordinated a number of projects, including an ethics review for the UK's Department of Trade and Industry Foresight project: 'Brain Science, Addiction and Drugs' (2004-5); a commissioned report on "'Public Interest' and 'Public Good' as Applied to UK Biobank Access Decision-Making" for the UK Biobank's Ethics and Governance Council (2007-8); and an international multi-centre project on 'New Developments in Neuroscience and Genetics' (2007-8). This latter project was funded by the European Monitoring Centre for Drugs and Drug Addiction, an agency of the European Union. Benjamin is an Honorary Associate at the ESRC Centre for Economic and Social Aspects of Genomics (Cesagen); and has been a Visiting Fellow at the Centre for Biomedical Ethics, at the National University of Singapore (2007 & 2008), University of Brisbane (2008), and the Hastings Centre, New York (2005). His primary research interests are in stem cell science and ethics; 'neuroethics', drugs (mis)use and addiction research. His research also focuses on the development of jurisprudential and political theory in the use of human rights concepts in biotechnology and bio-medicine.

Ruth Chadwick
Director
Cesagen

Ruth Chadwick is Director of the ESRC (Economic and Social Sciences Research Council) Centre for Economic and Social Aspects of Genomics (Cesagen), Cardiff University, UK. She also holds a Link Chair between Cardiff Law School and the School of English, Communication and Philosophy (ENCAP). She has co-ordinated a number of projects funded by the European Commission, including the EUROSCREEN projects (1994-6; 1996-9) and co-edits the journal Bioethics and the online journal Genomics, Society and Policy. She is Chair of the Human Genome Organisation Ethics Committee and has served as a member of several policy-making and advisory bodies, including the Panel of Eminent Ethical Experts of the Food and Agriculture Organisation of the United Nations (FAO), and the UK Advisory Committee on Novel Foods and Processes (ACNFP). She was editor-in-chief of the award winning Encyclopedia of Applied Ethics (1998), of which a second edition is now being prepared. She is an Academician of the Academy of Social Sciences and a Fellow of the Hastings Center, New York; of the Royal Society of Arts; and of the Royal Society of Medicine. In 2005 she was the winner of the World Technology Network Award for Ethics for her work on the relationship between scientific developments and ethical frameworks.

Y.T. Chen
Director and Distinguished Research Fellow
Institute of Biomedical Sciences
Academia Sinica

Professor Yuan-Tsong (Y-T) Chen was originally from Taiwan and have been committed to biomedical research for over 30 years. He recognises the tremendous impact that genetics and genomics have had on the improvement of health and treatment of diseases and has established the Chen Award, with HUGO, to celebrate research accomplishments in Human Genetics and Genomics in Asia Pacific and around the world. Professor Chen received his MD degree from National Taiwan University (Taipei) and a PhD from Columbia University (USA). He is currently a Distinguished Research Fellow and Director of the Institute of Biomedical Sciences, Academia Sinica, Taiwan, and Professor of Pediatrics and Genetics at Duke University Medical Center (USA). Professor Chen is a physician/scientist, recognised for his work on human genetic disorders. His translational research leads to the development of now standard therapies for two devastating inherited metabolic diseases: a simple and effective cornstarch therapy for severe hypoglycemia in glycogen storage diseases and an enzyme replacement therapy, the first ever treatment, for a debilitating, progressive and often fatal myopathy called Pompe disease. Professor Chen has also identified the genetic basis of and developed DNA-based diagnosis for several major heritable diseases, and more recently, his team in Taiwan has uncovered genes/SNPs associated with drug-induced Stevens-Johnson syndrome and warfarin sensitivity. His latest pharmacogenomic studies of adverse drug reactions paved the way for personalised medicine by preventing drug toxicity with a gene test. Professor Chen is an elected member of Academia Sinica and of the Academy Sciences for the Developing World.

Richard Cotton
Director, Genomic Disorders Research Centre
President, Human Genome Variation Society
Co-Editor, Human Mutation; Convenor, Human Variome Project

Richard Cotton AM BAgSc., Ph.D, D.Sc. (Melbourne) initiated the Mutation Research Centre, now renamed the Genomic Disorders Research Centre, in January, 1996 (www.genomic.unimelb.edu.au). He has always been interested in the biochemical genetics of human disease and has recently focussed on mutation. Amongst his more notable scientific achievements are the conception, planning and execution of the fundamental experiment, which proved that when two immunoglobulin producing cells were fused, the immunoglobulin of both parental cells were produced in the hybrid. This laid the experimental and theoretical foundation for the widely used monoclonal antibody technique. He also conceived the widely used tetrohydrobiopterin (BH_4) load test to identify the serious variants of PKU, but BH_4 is currently being trialed in heart disease. He is particularly interested in improving mutation detection technologies to make them cheaper and simpler, so that they can be more widely applied, and holds several patents in the area. A recent development has been a method to detect DNA damaging compounds. He has written two books entitled "Mutation Detection", initiated in 1991 the journal entitled "*Human Mutation*", and in 1991 initiated bi-yearly international workshops on Mutation Detection and in 1998 bi-yearly HUGO Mutation Detection Courses. In 1996 he has also started a worldwide initiative (The HUGO Mutation Detection Database Initiative, recently formed into the Human Genome Variation Society (HGVS) website: www.hgvs.org.) to capture and distribute lists of mutations. In June 2005, he was admitted as a Member of the Order of Australia for service to science through genetic research, particularly through the development of technologies to detect gene mutations that underlie birth defects or cause disease and through efforts to document findings. In June 2006, he convened a Meeting, co-sponsored by WHO, which initiated the Human Variome Project (www.humanvariomeproject.org). This project aims to collect worldwide genetic variation and its associated phenotype affecting human health. He convened a HVP planning meeting in Spain in May 2008 (www.humanvariomeproject.org/meetings/HVP2008/). The third HVP meeting will be held at the UNESCO headquarters in Paris in May 10-14, 2010 (www.humanvariomeproject.org/meetings/paris/). He is the author of over 300 scientific papers and 3 patents.

Albertina De Sario
Group Leader, INSERM

Albertina De Sario, PhD, CNRS Scientist, is group leader since 2003. Her group studied epigenetic modifications in human heterochromatic regions: the group defined a molecular border between heterochromatin and euchromatin in human chromosome 21 and identified an epigenetic tumor marker based upon DNA hypomethylation in heterochromatic genes. This new marker was validated for colon cancer diagnosis. Currently, the general aim of the group is to study the role played by epigenetics in the etiology of rare inherited diseases. A first project focuses on the epigenetic regulation of *CFTR* responsible for cystic fibrosis. A second ongoing project investigates epigenetic changes (DNA methylation and histone modifications) and aberrant transcription of heterochromatic genes in ICF (Immunodeficiency, Centromeric instability, Facial Dysmorphy) syndrome.

John De Vos
Cellular Therapy Processing Facility Director,
Assistant Professor, Faculty of Medicine, Department of Hematology,
University Hospital of Montpellier

Dr. De Vos (M.D., PhD.) is the Cellular Therapy Processing Facility Director of the Universtity Hospital of Montpellier. He is the former manager of the Regional High Density Microarray Core Facility (2003-2010). The focus of his research in the INSERM UNIT U847 is to model the networks controlling pluripotency in human embryonic or reprogrammed stem cells. He has founded the expression atlas AMAZONIA! that provides an access to the expression profile of all human genes in normal or malignant, pluripotent and somatic human cell types.

Marc Delpech
Paris Descartes Medical School (Paris Descartes University)

Marc DELPECH**,** MD, PhD, 58 years old, professor of Biochemistry and Molecular Biology, Paris Descartes Medical School (Paris Descartes University), head of the Cochin hospital Biochemistry and Molecular Genetics laboratory, head of a team of the Department : « Genetics, Development and Molecular Pathology » of Cochin Institute U1016 INSERM**.** The main topics of this research team are molecular Genetics of auto-inflammatory syndromes and amyloidosis.He isPresident of the AFSSaPS (French Drug Agency) “National Committee for Medical in vitro Diagnosis”. Member of several Scientific Committee (University Paris Descartes, “Deafest cystic fibrosis”, “Hemophiliacs French Association”, French Blood Transfusion, French Biomedicine Agency,...). Member of the French Academy of Medicine.

Jean-Louis Guéant
Inserm U954, Medical Faculty of Nancy
Chief Department of Biochemistry-Molecular Biology-Nutrition-Metabolism,
University Hospital Center of Nancy, France

Dr Guéant, MD, DSc, AGAF, is the scientific deputy dean and the director of Inserm U954 « Nutrition-Genetics and Environmental Exposure » at the Medical Faculty of Nancy, Nancy-University, France. He is deputy president of the Inserm scientific committee of Cardiovascular-Nutrition-Metabolism and president of the French Universitary Association of Medical Biochemistry and Molecular Biology. His field of interest is the one-carbon metabolism in nutrigenetics and nutrigenomics. He coordinates the program « Nutrivigene » of the French national agency for research. The Inserm Research Unit 954 aims to evaluate short and long term effects of deficiency in methyl donors/precursors in relation with perinatology, fetal programming, inherited metabolic rare diseases and aging. Its field of interest is focused on the links between the one-carbon metabolism and epigenetic mechanisms involved in cellular stress, inflammation, genotoxicity and immunomodulation, in brain, heart, and digestive organs. The project is designed with a transversal approach, using relevant cellular and animal models and validating the experimental pathomechanisms in population studies. A national reference centre of inherited metabolic diseases is led by the investigators of the Unit, in the University Hospital Center of Nancy.

Christian Hamel
Professor
Director, Institute for Neurosciences of Montpellier,
Genetics and therapy of retinal and optic nerve blindness, Montpellier, France
Genetics of sensory diseases, University hospital, Montpellier

Dr. Hamel (MD, PhD) is Director of the Institute of Neurosciences of Montpellier (Inserm research laboratory) and head of the Department of Genetics of Sensory Diseases at the University Hospital, Montpellier, France. Dr Hamel’s team is involved in gene discovery for retinal dystrophies and optic neuropathies, and run several programs for gene therapy and pharmacological therapy.

Yoshihide Hayashizaki
Director
Omics Science Center
RIKEN Genomic Sciences Center

Yoshihide Hayashizaki received his M.D. and Ph.D. from Osaka University Medical School in 1982 and 1986, respectively.In 1992, he joined RIKEN, and was appointed Project Director for the RIKEN Genome Project in 1995. Since then he has been taken a transversal data-driven approach to analyze transcriptomes by developing unique technologies including a series of full-length cDNA technologies. With this approach, he has established large amount of full-length cDNA clone bank. This activity was followed by organization of FANTOM (Functional Annotation of Mammalian), an international consortium, originally to annotate a large number of cDNA and subsequently expanded to transcriptome and network analysis. FANTOM activities revealed that more than 63% of the genome — not just the ~1.5% fraction that are protein-coding exons — is transcribed as RNA. In 2008, he appointed to the Director of the Omics Science Center. The center was established to link a variety of omics subdisciplines to molecular networks and pathways in order to advance our understanding of biological phenomena as systems at the molecular level.

Martin Hibberd
Senior Group Leader
Assoc Director, Infectious Diseases
Genome Institute of Singapore

Dr Martin Hibberd BSc(Hons) PhD; Senior Group Leader, Infectious Diseases at the Genome Institute of Singapore and has adjunct positions at the National University of Singapore and Imperial College (London, UK). Graduated with Honors from Brunel University in 1985 (West London, UK) and received his Doctorate, on the immune-genetics of the human T-cell antigen receptor, from King’s College, London. Has a broad scientific background spanning both microbial and human determinants of infectious and inflammatory diseases. Previous posts include WHO-funded Senior Microbiologist at

the UK's central Public Health Laboratories, and for seven years prior to his current appointment he was Lecturer and Senior Lecturer in Pediatric Infectious Diseases at the Imperial College School of Medicine, one of the very top-ranking British universities. Current research interests cover both pathogen and host aspects of infectious disease, understanding how microbial agents causes the observed disease (including pathogen identification and sequence characterization) and why specific individuals are susceptible to the disease (using host genetics on a genomic scale). Approaching infectious disease from these two directions also allows specific host pathogen responses to be investigated (utilizing RNA micro and low density arrays). This work aims to identify key host responses to specific pathogens that could be targeted by new therapies.

Gerardo Jimenez-Sanchez
Chief Scientific Officer, BioFields
Chair, Working Party in Biotechnology, Organisation for Economic Co-operation and Development (OECD)

Dr. Gerardo Jimenez-Sanchez was born in Mexico City. He is a M.D. with a Ph.D. in Human Genetics and Molecular Biology from Johns Hopkins and training his business administration from the IPADE Business School. He is a certified Pediatrician and a member of the National Academy of Medicine, the American Society of Human Genetics, the American Society of Gene Therapy, the Society for Inherited Metabolic Disease, the European Society of Inborn Errors of Metabolism. He is founder Director of the National Institute of Genomic Medicine in Mexico and Chairman of the Working Party on Biotechnology at the Organization for Economic Cooperation and Development (OECD). He is the leading investigator in the Mexican Genomic Diversity Project and the Mexican Medical Re-sequencing Initiative. Dr. Jimenez-Sanchez is Professor of Genomic Medicine at the National Autonomous University of Mexico, Member of the National Academy of Medicine. He serves as Council member of the Human Genome Organisation, as well as on the Board the Public Population Projects in Genomics (P3G). He is founder President of the Mexican Society of Genomic Medicine and Regional Editor for Latin America and the Caribbean of the HUGO Journal.

Christian Jorgensen
Head of Medical and Development Genetics
Western General Hospital

Laurent Journot
Research Director, Institute for Functional Genomics, Montpellier
Director, MGX - Montpellier GenomiX, Montpellier

Laurent Journot is currently Research Director at the Institute for Functional Genomics in Montpellier. He obtained his B.Sc. and "Aggrégation" from the Ecole Normale Supérieure de Cachan and his M.Sc. from Montpellier University. During his post-doctoral training at the ZMBH in Heidelberg, he cloned and characterized the G protein-coupled receptor for PACAP, a neuropeptide that displays anti-proliferative and anti-apoptotic activities. He previously was a Senior Investigator at the Centre for Pharmacology and Endocrinology in Montpellier, and focused on the signalling mechanisms elicited by PACAP to prevent neuronal apoptosis. His current research focus on the physiological roles of imprinted genes. Parental genomic imprinting is an epigenetic mechanism of regulation that restrains the expression of a small subset of mammalian genes to one parental allele. Journot's main contribution is the demonstration that, despite seemingly disparate functions, imprinted genes work in a single gene network that controls embryonic development and whose function at the cellular level remains to be identified. Journot is also the Director of MGX - Montpellier GenomiX, a technological core facility that provides service in microarray-based technologies, deep sequencing, bio-statistics, and bio-informatics.

Bernard Klein
Professor of Unversity-Hospital Practitioner
Director of the Institute of Research in Biotherapy

Bernard Klein is the head of the Institute of Research of Biotherapy that groups 100 people working on normal and cancer stem cells with the aim to develop drugs (cell drugs) involved in the field of regenerative medicine. Bernard Klein has founded the Cell and Gene therapy Unit at the University Hospital of Montpellier, coordinates 150 hour courses for teaching the regenerative medicine field and is an active member of national agencies for agreement of cell therapy clinical trials.

Dhavendra Kumar
Consultant, Clinical Genetics at the University Hospital of Wales
Hon. Senior Research Fellow, Institute of Medical Genetics, Cardiff University, Cardiff, Wales, UK
Lead Clinical Geneticist, Clinical Cardiovascular Genetics

Dr. Dhavendra Kumar is Consultant in Clinical Genetics at the University Hospital of Wales and Hon. Senior Research Fellow in the Institute of Medical Genetics, Cardiff University, Cardiff, Wales, United Kingdom. He is the lead Clinical Geneticist for Clinical Cardiovascular Genetics. After qualifying in Medicine from the King George's Medical College, University of Lucknow, India, he completed postgraduate training in Pediatrics with an MD. Since 1980 he has pursued a career in Medical Genetics in the UK. In 1990 he became a Diplomate of the American Board of Medical Genetics. He is a Fellow of the American College of Medical Genetics (FACMG) and as well as Royal Colleges of Physicians (FRCP-London and FRCP-Ireland) and Paediatrics and Child Health (FRCPCH-UK). He has published three books (Genetic disorders of the Indian Subcontinent-Springer/Kluwer, Genomics and Clinical Medicine- Oxford University Press and Principles and Practice of Clinical Cardiovascular Genetics-Oxford University Press) and has several

publications in pediatrics and clinical genetics. He established and developed a new biomedical journal (Genomic Medicine-Springer) that led to setting up the official journal of the Human Genome Organization ('The HUGO Journal'). He serves on the editorial board of the HUGO Journal as Editor in Chief. He is currently a member of the publications committee of HUGO. Dr. Kumar's current clinical and research interests include clinical cardiovascular genetics, applications of genomics in medicine and human health and promoting genetics and genomics for improving the health in emerging economies of the developing world. He is currently actively involved in promoting medical genetics and genomics in developing countries through education programs and establishing a comprehensive database on genomic variation and genetic disorders across several developing countries.

Doron Lancet
Ralph & Lois Silver Chair of Human Genomics
Department of Molecular Genetics
Head, Crown Human Genome Center
Weizmann Institute of Science

Prof. Doron Lancet is a genome and bioinformatics researcher at the Weizmann Institute of Science. He headed there the Department of Membrane Research and Biophysics, and is currently Professor at the department of Molecular Genetics, where he studies the genetics and genome organization of the olfactory receptor gene superfamily. Since 1992 Lancet has been active in the Israeli and international genome scenes, and is currently Director of Israel's National Knowledge Center for Genomics, and of the Crown Human Genome Center at the Weizmann Institute. In this realm, his research has included the discovery of genes for monogenic diseases, as well as high throughput SNP scoring for elucidating polygenic diseases and pharmacogenetics. In the realms of bioinformatics, Lancet and group have developed GeneCards, a widely-used web-based human gene compendium. Lancet is a member of the European Molecular Biology Organization (EMBO) since 1996 and of HUGO since 1994. He has won international prizes in the field of olfaction, and is author of more than 180 papers and reviews.

Marie-Paule Lefranc
Professor University Montpellier 2, Institute of Human Genetics CNRS-UPR1142
Founder and Director of IMGT®, the international ImMunoGeneTics information system®

Marie-Paule Lefranc is Professor Classe Exceptionnelle at the University Montpellier 2, Senior Member of the University Institute of France. She is Founder and Director of IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org, that she created in 1989, at Montpellier, France. She is Head of the Laboratoire d'ImmunoGénétique Moléculaire, created with Prof. Gérard Lefranc and located, since 1998, at the Institute of Human Genetics (CNRS-UPR1142). She has authored over 300 scientific publications in international journals on the molecular immunogenetics of immunoglobulins, T cell receptors, antibody engineering, and on human genetics and immunoinformatics. Three of her major contributions are: firstly, the finding of large chromosomal deletions which encompass several immunoglobulin genes and account for the simultaneous absence of IgG1, IgG2, IgG4 and IgA1 isotypes in healthy individuals, allowing the overall ordering of the human immunoglobulin constant genes; secondly, the full characterization of the human T cell receptor gamma locus, for which she received the ROSEN prize of Cancerology in 1988; thirdly, the definition of the IMGT Scientific chart rules and IMGT-ONTOLOGY concepts, which are the foundations of IMGT® which has become the global reference in immunogenetics and immunoinformatics.

Harris A Lewin
Professor of Immunogenetics
Gutgsell Endowed Chair
Department of Animal Sciences
Director, Institute for Genomic Biology
University Illinois, USA

Professor Lewin's current research interests are in the area of mammalian comparative and functional genomics. His research has resulted in the development of high-density comparative maps for the cattle, human and other mammalian genomes, and novel software for in silico gene mapping using the human genome as a template. In addition, his group produced the first large-scale cDNA microarray and oligoarray for functional genomics of ruminants, which have been applied to a systems level analysis of embryonic development and diet-genotype interactions.

Bing Lim
Senior Group Leader, Genome Institute of Singapore
Associate Professor of Medicine, Harvard Medical School
Associate Physician, Beth Israel Hospital, Harvard Medical School

Research Focus: ES cell as an engine for Pluripotency, Tissue Regeneration and Epigenomic alteration; ES cells hold great promises for cell therapeutics; The focus of our team is to elucidate 3 fundamental paradigm underlying the basis for the practical exploitation of ES cells, combining a gene discovery effort with gene functionalization studies.

Edison T Liu
Executive Director, Genome Institute of Singapore (Biomedical Sciences Institutes)
Professor of Medicine, National University of Singapore
President, Human Genome Organisation

Dr. Edison Liu received his residency training in internal medicine at WashingtonUniversity, St. Louis, and clinical cancer fellowships at StanfordUniversity (Oncology), and at the University of California at San Francisco (Hematology). He then pursued post-doctoral studies as a Damon-Runyan Cancer Research Fellow at the University of California at San Francisco in the laboratory of Dr. J. Michael Bishop. In 1987, he joined the faculty of Medicine at the University of North Carolina at Chapel Hill where he was director of UNC's Specialized Program of Research Excellence (SPORE) in Breast Cancer. In 1996, he joined the NCI as the Director of the Division of Clinical Sciences. In 2001, Dr. Liu assumed the position of Executive Director, Genome Institute of Singapore. His current scientific research investigates the dynamics of gene regulation on a genome scale that can explain biological states in cancer. Dr. Liu has contributed over 245 articles, reviews, and book chapters to the scientific literature. Dr. Liu also is the Executive Director of the Singapore Tissue Network, the national tissue repository in Singapore. Dr. Liu's awards include the Leukemia Society Scholar (1991-1996), the Brinker International Award for basic science research in Breast Cancer (1996), the Rosenthal Award from the American Association for Cancer Research (2000), the President's Public Service Medal for his work in helping Singapore resolve the SARS crisis, and a Doctor of Medicine Sciences honoris causa (2007). In 2007, Dr. Liu was elected the President of the Human Genome Organization (HUGO).

Lance Miller
Associate Professor
Department of Cancer Biology
Wake Forest University School of Medicine

My research is focused on the discovery and characterization of mechanistic and clinically useful aspects of carcinogenesis from a systems biology perspective. Using genomics technologies such as DNA microarrays, my laboratory investigates the transcriptional dynamics and genomic architectures of primary tumors and cell lines at various stages of the oncogenic process and in different clinical contexts. Integrative analysis of genome-wide expression patterns, copy number alterations, and clinicopathologic features allows us to uncover transcriptional programs of mechanistic and prognostic relevance. This strategy has led to the identification and validation of gene expression signatures in liver, breast, ovarian and lung cancers that 1) reflect the activity of specific growth-regulating pathways, 2) define known and novel tumor subtypes, and 3) predict clinical outcomes such as disease recurrence and therapeutic response. Examples include prognostic signatures in breast cancer that reflect the operational configuration of the TP53 pathway (Miller et al, PNAS, 2005) and delineate new prognostic tumor subtypes based on "genetic grade" (Ivshina et al, Cancer Res, 2006). More recently, we have discovered a copy number-related transcriptional signature of disease recurrence in stage I non-small cell lung carcinoma (NSCLC) that outperforms all conventional prognostic factors, and identifies a substantial subgroup of patients that may benefit from adjuvant chemotherapy (Broët, et al, Cancer Res,2009). We have also pioneered novel data mining strategies that integrate multiple forms of clinico-genomic information (expression, copy number, patient survival) and are capable of pinpointing known and novel candidate oncogenes. Using this approach, we have recently identified and validated a novel breast cancer oncogene at chromosome 8p11 that promotes transformation, anchorage-independent growth, invasion through matrigel, and tumor formation in mouse xenograft models with an ability to interact with and activate H-Ras. Other candidate genes identified by this method are currently being prioritized for functional characterization based on their potential for therapeutic targeting.

Michele Ramsay
Head, Molecular Genetics Laboratory, National Health Laboratory Service and
University of the Witwatersrand
Chair, Wits Bioinformatics Steering Group

Professor Michèle Ramsay (PhD) is head of the Molecular Genetics Laboratory in the Division of Human Genetics at the National Health Laboratory Service and University of the Witwatersrand. She is Chair of the Wits Bioinformatics Steering Group, joint champion of a cross-faculty Research Thrust, ***"Molecular Biosciences: Health for Africa***" and joint editor and author of a textbook, "Molecular Medicine for Clinicians" (Wits University Press). In addition she is the Interim Director of the Sydney Brenner Institute for Molecular Bioscience (Wits University) which focuses on a molecular understanding of non-communicable diseases in African populations. Her research interests include the genetic basis and molecular epidemiology of single gene disorders in South African populations and the role of genetic and epigenetic variation in the molecular aetiology of foetal alcohol spectrum disorders (FASD) and other diseases exacerbated by adverse lifestyle choices.

Richard Redon
Senior Researcher
INSERM

Richard Redon has been appointed Senior Researcher by the Inserm in March 2009. As laureate of the Inserm/CNRS young investigator program 2009, he leads a new research group working on the genetics of sudden cardiac death at the 'institut du thorax' in Nantes, France. He obtained his PhD at the University of Strasbourg and was awarded the Young Investigator Award of the European Society of Human Genetics in 2002. He completed his postdoctoral fellowship in 2006 at the Sanger Institute (UK) then focused on Copy Number Variation (CNV) discovery with high-density array CGH. Redon has led several breakthroughs in the field of genomics including establishing the first detailed map of copy number variations in humans.

Agnes Rotig
Research Director at INSERM (Institut National de la Santé et de la Recherche Médicale), France

Dr. Rötig (PhD) is the head of the group working on mitochondrial diseases in Necker Hospital (Paris). This group has initially settled and integrated platform of clinic, biochemistry and molecular analysis to investigate patients with OXPHOS disease. The scientific field of this group is the identification of genes involved in mitochondrial disorders and the investigation of their pathophysiology. They have described the first non-neuromuscular presentation of mitochondrial diseases and characterized the very first mutations in nuclear genes resulting in defects of Krebs's cycle or the respiratory chain.

Ruan Yijun
Senior Group Leader
Assoc Director, Genomic Technologies

My primary interest is to elucidate the structures and dynamics of all functional DNA elements in complex genomes through transcriptome characterizations and genome interrogation. To facilitate such understanding we have developed Paired-End-Tag (PET) sequencing and mapping methodologies. We are applying these sequencing-based technologies to address complex biological questions such as how cancer cells progress and how stem cells maintain their unique properties. Another major interest of mine is to discover previously uncharacterized microbial genes and genomes that are relevant to human health.

Yoshiyuki Sakaki
President, Toyohashi Institute of Technology
Emeritus Professor, University of Tokyo
Special Consultant of RIKEN

Dr. Yoshiyuki Sakaki is the President of Toyohashi Institute of Technology since April 2008. He is also Emeritus Professor of the University of Tokyo and Special Consultant of RIKEN. He has been the Director of RIKEN Genomic Sciences Center from 2004 to 2008(March) and also the President of HUGO (Human Genome Organization) from 2002 to 2005. He has represented Japan in the international Human Genome Project, and the RIKEN team led by him made significant contributions to the completion of the human genome sequence, particularly playing a major role in completing the sequence of chromosome 21 and 11. Currently he is the chief coordinator of Genome Network Project, a Japanese nation-wide project towards the comprehensive study on the transcriptional regulatory networks of the human genome. He was awarded to "Chevalier" from France Government 2001, in recognition of his contribution to the scientific cooperation between France and Japan, the Award of Japanese Society of Human Genetics (2001),and Medal of Purple Ribbon from Japanese Government (2003).

Stephen W Scherer
Director, McLaughlin Centre and the Toronto Centre for Applied Genomics
Professor, Hospital for Sick Children and University of Toronto

Dr. Scherer holds the GlaxoSmithKline-CIHR Endowed Chair in Genetics and Genomics at The Hospital for Sick Children and University of Toronto. He has made numerous contributions to medical genetics including mapping sequencing and disease gene studies of human chromosome 7. He collaborated with the J. Craig Venter Institute to generate the first genome sequence of an individual ushering in an era of genomic medicine. In 2004, his team co-discovered global gene copy number variation (CNV) and has since shown that CNV is the most abundant type of variation of human DNA. His group has also discovered CNV to contribute to the etiology of autism and the Database of Genomic Variants he founded facilitates hundreds of thousands of diagnoses each year. He sits on the Scientific Advisory Board of Combimatrix Diagnostics and Autism Speaks, and he is on the Board of Trustees of Genome Canada and the Human Genome Organization (HUGO). Dr. Scherer has won numerous honors including the 2004 Steacie Prize, an International Howard Hughes Medical Institute Scholarship, and the 2008 Premier Summit Award for Medical Research.

Lawrence Stanton
Senior Group Leader
Deputy Director, Research Affair
Genome Institute of Singapore

Research Focus: Transcriptional control of embryonic stem cells; Human embryonic stem (ES) cells provide a unique avenue to study genes that are involved in the earliest stages of differentiation during embryo development. My lab is primarily focused on identification of genes and regulatory pathways that control embryonic stem cell growth and differentiation. We have generated a comprehensive map of the transcriptome from ES cells using genome-wide expression technologies. We have focused particular attention on a set of transcription factors that are uniquely expressed in ES cells and likely play an important role in determining the differentiation status of the ES cells. Functional genomic approaches are being applied to dissect the transcriptional networks that are governed by these known and novel transcription factors, for example REST, Zic3, and Zfp206. The influence these and other transcription factors have on ES cells is then monitored by transcriptional profiling and functional assessment of differentiation capability. Direct targets for key transcription factors are being identified by chromatin precipitation (ChIP) experiments. We are also looking at the epigenetic changes that are mediated by the binding of these transcription factors. The activity of many of these genes is being exploited to drive lineage-specific differentiation of human ES cells and lead to better understanding of the molecular events that underlie early embryonic development and cellular differentiation. Our long-term goal is to derive cell types in vitro that will provide therapeutic utility for certain degenerative diseases in man.

Sumio Sugano
Professor, Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo

Dr. Sugano (MD, PhD) is a Professor at the Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo. The main emphasis of his research is to identify and collect genes of human en masse in the form of full-length cDNAs. He initiated the FLJ project of collecting and determining the entire sequences of human full-length cDNAs in 1999. He is now using second generation sequencers to characterize transcriptome landscape. He was a Council Member of the Human Genome Organisation (2002-2008) and was Chair of Human Genome Organisation Asia Pacific during the same period.

Todd Taylor
Team Leader
MetaSystems Research Team
RIKEN Advanced Science Institute

Todd Taylor is Team Leader of the MetaSystems Research Team in the Computational Systems Biology Research Group at RIKEN Advanced Science Institute, Yokohama, Japan. His team's main area of research is the analysis of and development of tools for metagenomic sequence data derived from various human health- and environmental-related microbial communities. He has been a research scientist at RIKEN since 1998, and received his Ph.D. in Molecular and Medical Genetics from Oregon Health Sciences University, Portland, Oregon. During graduate school he conducted linkage analysis studies for various disorders and was responsible for the initial identification and mapping of the locus for Hallervorden-Spatz syndrome, a rare autosomal recessive disorder. As an active member of the International Human Genome Sequencing Consortium, Todd served as coordinator for chromosomes 11, 18p, and 21. He also helped to coordinate the finishing of chimpanzee chromosomes 22 and Y. Todd has been a HUGO member since 1999 and has served on the HUGO council since 2006.

Herve Thoreau
Production Leader, Sequencing
Genome Institute of Singapore

Hervé began to work at Genome Institute of Singapore in 2001. He is the Production Leader for Genome Technology and Biology group. The GT&B group is applying cutting edge innovative genomic technologies and next generation sequencing platform to study genome variation and transcription regulation in complex biological systems. Prior to joining Genome Institute of Singapore, Hervé worked for 4 years in the French Biotechnology Company, GENSET S.A., as team leader of the Mapping Production group. Hervé holds a B.Sc. in Industrial Bio-experimentation from E.S.T.B.A, Paris. Concurrently he is also pursuing graduate studies in Biological Engineering at the Université de Technologie de Compiègne (UTC), conferring the *diplôme d'ingénieur*.

Veronica Van Heyningen
Section Head
MRC Human Genetics Unit
Honorary Professor University of Edinburgh, FRS, FRSE, FMedSci

First degree in genetics (Cambridge); DPhil on early gene mapping studies using somatic cell hybrids, with Walter Bodmer (Oxford); Beit Memorial Fellow, MRC Mammalian Genome Unit, Edinburgh; eventually rose through the ranks at MRC Human Genetics Unit, Edinburgh from postdoctoral fellow to group leader and Section Head. Honorary Professor University of Edinburgh, FRS, FRSE, FMedSci. Howard Hughes International Research Scholar 1993-1998. Member of UK Human Genetics Commission 2000-2005. EMBO member 2003. Fellow of the Royal Society 2007. Human geneticist (non-clinical) working on developmental eye anomalies such as aniridia and anophthalmia/microphthalmia. Interested in what detailed analysis of the human phenotypes in individuals with known mutations can tell us about the biological role of the genes involved. Ultimately most human developmental genes need to be explored in model systems and we are currently studying the roles of PAX6 and SOX2 in humans, the mouse and zebrafish. A major area of endeavour since the early 1990s is to understand the mechanisms of long-range regulation of gene expression. Transcription factors like PAX6 and SOX2 show complex spatiotemporal and quantitative control of expression, requiring a large number of strongly sequence-conserved cis-regulatory elements, found upstream, downstream and within the gene. Such regulatory regions can stretch more than a megabase either side of the gene. We are aiming to understand the spatial and functional organisation of these interacting elements (which often function as enhancers in reporter transgenic assays), additively recapitulating the sum of the total gene expression pattern. Expression-associated patterns of open chromatin suggest dynamic genomic organisation. Another emerging interest is in the mechanisms underlying disease with non-Mendelian segregation patterns. We are exploring how environmental factors effect phenotype modification.

Julian Venables
Lead Scientist at the CNRS Institute of Molecular Genetics at Montpellier (IGMM)

Dr Venables gained his PhD from Leicester UniversityU.K. and this was followed by post-doctoral experience at the MRC Human Genetics Unit in Edinburgh (January 1999- June 2001) and then the Institute of Human Genetics in Newcastle. From February 2007 he coordinated a group of

bioinformaticians in a project to characterize the transcriptome of breast and ovarian cancers in Sherbrooke, Québec. Since February 2010 he is lead scientist at the CNRS Institute of Molecular Genetics at Montpellier (IGMM).

Martin Vingron
Director, Max Planck Institute for Molecular Genetics, Germany
Head of Computational Molecular Biology Department

Martin Vingron is a mathematician by education who has done his PhD in computational biology at EMBL in 1991. At the time and for a number of years of postdoctoral training his research has focused on the analysis of protein sequences, sequence analysis, sequence comparison, and molecular evolution. Methods of discrete optimisation were used for the design of comparison algorithms and probability theory was applied to answer questions of significance of computational results. Later, as a department head at the German Cancer Research Center, his focus shifted towards the processing and mathematical analysis of DNA microarrays. Accordingly, the methods largely drew on statistical data analysis techniques. During the last years his research interest lies in utilizing gene expression data as well as evolutionary data for the elucidation of gene regulatory mechanisms. He also acts as a chair of the *RECOMB* <http://recomb.org/> (Research in Computational Molecular Biology) steering committee.

Wei Chia-Lin
Senior Group Leader, Genome Technology & Biology
Group Leader, Clone Production & Management Genome Institute of Singapore

The primary focus of my research is to develop and apply sequencing based genomic technologies for genome and epigenome function elements interrogation pertinent to lineage specification and development. Using various second generation sequencing platforms, my research focuses on integrating the technology development effort to construct global, high resolution and unbiased functional networks constituted by the key genetic modulators in stem cell genomes and their dynamics throughput developmental process. Furthermore, we are also interest to characterize how these modulators can directly impact on gene expressions. Such knowledge will provide fundamental understanding in various nuclear processes including transcription, add new insight on the unique gene expression patterns crucial for cell fate determination and open the possibility of manipulating such structures to achieve cell type specific differentiation.

WORKSHOP ABSTRACTS

The abstracts on the following pages are arranged in order of workshop number. At the end of the abstracts there is an index of all the abstract authors.

1. Computational Biology, Computational Genomics

(19 May 2010, 1.05 PM - 3.05 PM, Salle Pasteur) - Chaired by Doron Lancet and John de Vos

Poster No: P002-W
GENCODE: A reference gene set for the human genome

J. Rajan[1], J. Mudge[1], B. Aken[1], M. Diekhans[7], R. Harte[7], F. Kokocinski[1], M. Lin[3], M. Tress[2] , M. Van Baren[4], J. Gilbert[1], M. Brent[4], M. Gerstein[6], R. Guigo[5], M. Kellis[3], A. Reymond[8], S. Searle[1], T. Hubbard[1], J. Harrow[1]
[1]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK, [2]Spanish National Cancer Research Centre (CNIO), Madrid, Spain, [3]MIT Computer Science and AI Laboratory, Broad Institute, Cambridge, MA, USA, [4]Lab. for Comp. Genomics and Dept. of CS, Washington Univ, St. Louis, Missouri, USA , [5]Centre for Genomic Regulation, Barcelona. Catalonia. Spain, [6]Department of Molecular Biophys. and Biochem. Yale University New Haven, CT USA, [7]Center for Biomolecular Science and Engineering, UCSC, CA, USA. [8]Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland

The HAVANA group at the Wellcome Trust Sanger Institute aim to complete the manual annotation and experimental transcript validation of the human genome by 2011. This work is being carried out as part of the GENCODE consortium, a subgroup of the ENCODE (the Encyclopedia Of DNA Elements) project. Currently over 50% of the genome has been manually annotated by HAVANA, with Ensembl gene objects used in the areas not yet described to create the merged GENCODE genebuild. This annotation contains protein-coding, non-coding and pseudogene loci and is supported by mRNA, EST or protein evidence and is especially important in the accurate description of splice variation and duplicated gene clusters. Our annotation process also identifies sites in the reference genome where improvements are required, including both sequencing errors in CDS as well as genuine polymorphisms. We work with the Genome Reference Consortium (GRC) to ensure such regions are properly represented.

In addition, Havana are an integral part of the consensus CDS (CCDS) project, whereby protein translations for human and mouse genes are agreed upon by HAVANA, Ensembl, UCSC and RefSeq. All CCDSs are contained within the GENCODE reference set which is now shown as the default Ensembl human gene build since Ensembl build 56. Next generation generation technologies are beginning to impact on annotation and we will describe how the 1000 genomes project and RNAseq transcriptome data will be integrated into the reference gene set. Also, we will highlight how proteomics data can be integrated and improve existing annotation.

Poster No: P019-W
Reconstitution of meta-regulation networks by MIR@NT@N using transcriptomics analyses reveals that several miRNAs are antagonistically regulated by retinoic acid and estrogen in breast cancer cells.

Charles-Henri Lecellier[5], Antony Le Béchec[1],*, Anne Saumet[2],*, Elodie Portales-Casamar[3], Guillaume Vetter[1], Pierre-J Zindy[4], Michèle Moes[1], Charles Theillet[2], Wyeth W. Wasserman[3], Evelyne Friederich[1]
[1]Cytoskeleton and Cell Plasticity lab, Life Sciences Research Unit-FSCT, University of Luxembourg, L-1511 Luxembourg. [2]Institut de Recherche en Cancérologie de Montpellier INSERM U896, Université Montpellier1, CRLC Val d'Aurelle Paul Lamarque, Montpellier, F-34298, France. [3]Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, 950 West 28th Avenue, Vancouver, BC V5Z 4H4, Canada. [4]Structure and Function of the Cell Nucleus, Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Montréal (Québec), Canada [5]Institut de Génétique Moléculaire de Montpellier UMR 5535 CNRS, 1919 route de Mende, F-34293 Montpellier cedex 5, France Université Montpellier 2, Place Eugène Bataillon, F-34095 Montpellier cedex 5 Université Montpellier 1, 5 Bd Henry IV, F-34967 Montpellier cedex 2, France.

Identifying meta-regulation networks whose perturbations are causal to the development of cancers is one of the major challenges to understand cell biology and treat cancer cells. We conceived a new computational framework, named MIR@NT@N, which reconstructs such networks of transcription factors and microRNAs. It builds on the JASPAR/PAZAR databases for transcription factor binding profiles with further information derived from miRBase and microRNA.org (for miRNA target predictions). Graph-based algorithms identify specific sub-networks, emphasizing feedback and feed-forward loops. We used this method to analyze independent transcriptomics data of breast cancer cells. We thereby revealed that several miRNAs are antagonistically regulated by retinoic acid and estrogen. Moreover, our analyses characterized additional oncogene and tumour suppressor miRNAs implicated in breast tumorigenesis and unveil some miRNA-based molecular mechanisms of cancer chemotherapies targeting transcription factors. MIR@NT@N is a user-friendly web resource freely available at http://maia.genouest.org/demo/

Poster No: P034-W
Digital Gene Expression Pipeline : from mapping to visualization

[1]**Anthony Boureux**, [2,3]Nicolas Philippe, [2]Laurent Bréhélin, [4]Jorma Tarhio, [3]Thérèse Commes and [2]Éric Rivals
[1]CRBM UMR5237, Université de Montpellier II - CNRS, 1919 route de Mende, 34095 Montpellier, France, [2]Laboratoire d'Informatique, de Robotique et de Microélectronique, Université de Montpellier II, UMR 5506 CNRS, 161 rue Ada, 34392 Montpellier, France, [3]Groupe d'études des transcriptomes, Université de Montpellier II, Institut de Génétique Humaine, UPR1142-CNRS, 141, rue de la cardonille, 34095 Montpellier, France and 4Helsinki University of Technology, PO Box 5400, FI-02015 HUT, Finland

Massive parallel high-throughput sequencing is now widely used to analyze the transcriptome or interactome at unprecedented depth on a genome-wide scale. However, exploration of these massive data remains a challenge. These techniques yield short sequence reads that are then mapped on a genome sequence to predict putatively transcribed or protein-interacting regions. Factors such as background distribution, sequence errors, and read length can impact on the prediction capacity of sequence census experiments. Here we suggest a computational approach to measure these factors and analyze their influence on both Digital gene expression (DGE) and chromatin immunoprecipitation-sequencing (Chip-seq) assays. This investigation provides new clues on both methodological and biological issues. For instance, by analysing Chip-seq read sets, we estimate that 4.6% of reads are affected by SNPs. We show that, although the nucleotide error probability is low, it significantly increases with the position in the sequence. Choosing a read length above 19 bp practically eliminates the risk of finding irrelevant positions, while above 20 bp the number of uniquely mapped reads decreases. With our procedure, we obtain 0.6% false positives among genomic locations. Hence, even rare signatures should identify biologically relevant regions, if they are mapped on the genome. Therefore, we propose a new pipeline to analyze SAGE, DGE and RNA-Seq data. First, sequences are mapped to the genome using the best length (usually between 19 and 20 bp), and sorted in located (none, once or multiple times) or erroneous sequences. Second, located sequences can be visualize on genomic browser as Ensembl (http://www.ensembl.org) or UCSC Genome Browser (http://genome.ucsc.edu) with their occurrences among others genomic tracks.

Poster No: P049-W
Identifying and Prioritizing Human Disease Genes using FunSimMat

Mario Albrecht[1], Andreas Schlicker[1], Hagen Blankenburg[1], Sarah Diehl[1] and Thomas Lengauer[1]
[1]Department of Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Campus E1.4, 66123 Saarbrücken, Germany.

Large-scale experiments such as genome-wide association studies (GWAS) and RNA interference (RNAi) screens associate candidate genes with complex disease processes. Since the number of discovered candidates is frequently quite large, computational methods are applied for identifying and prioritizing the most relevant candidates for further validation experiments. Here, we present our extended FunSimMat web service that offers various functional similarity measures for ranking candidate genes and proteins based on the Gene Ontology. We show that our novel approach achieves excellent results in discovering important disease genes from human GWAS and RNAi screens. FunSimMat is easy-to-use and publicly available free of charge at http://www.funsimmat.de

Poster No: P064-W
Comparing human trophoectoderm with embryonic stem cells: a molecular analysis of the first developmental bifurcation

Bai Qiang[1], Said Assou[1,2] Marilyne Dijon-Grinand[1], Tanguy Le Carrour[2], Samir Hamamah[1,2,3,4], John De Vos[1,2,3,5]
[1]INSERM, U847, Montpellier, F-34000 France ; [2]CHU Montpellier, Institute for Research in Biotherapy, Hôpital Saint-Eloi, Montpellier, F-34000 France ; [3]Université MONTPELLIER1, UFR de médecine, Montpellier, F-34000 France ; [4]CHU Montpellier, Département de Médecine et Biologie de la Reproduction, Hôpital Arnaud de Villeneuve, Montpellier, F-34000; [5]CHU Montpellier, Unité de Thérapie Cellulaire et Génique, Hôpital Saint-Eloi, Montpellier, F-34000 France

Introduction:
The first week of human development is characterized by the differentiation of the early embryo into the inner cell mass (ICM) and the trophectoderm (TE). Whereas the ICM, maintained in vitro as human embryonic stem cells (hESC) has been analyzed by innumerable studies, the human TE at blastocyst stage is by contrast still mostly unelucidated. TE is of prime importance since it latter develops into placenta that will provide the nutrient support for further embryonic development. It plays also a major role in the shaping of the blastocyst stage. The aim of the present study was to compare the TE and the hESC transcriptome and define the first steps of differentiation taking place during the human embryonic development.

Materials and Methods:
Five TE samples were obtained from fresh embryos issued during an in vitro fertilization (IVF) program. At blastocyst stage (day 5), the TE was mechanically seperated. The human embryos were donated to research after informed consent from the couples. This embryo research program directed at deriving new hESC lines has received an IRB agreement. Each TE was individually analyzed by whole genome U133P Affymetrix (Santa Clara, CA) oligonucleotides microarrays. A double in vitro transcription amplification step according to the manufacturer's instructions was used. Gene enrichment is analyzed with Gene Set Enrichment Analysis (GSEA, Broad Institute, Cambridge) and statistical analysis was carried out with Significance Analysis of Microarray (SAM, Stanford, CA).

Results:
Using SAM with a false discovery rate of 0.2774% and a fold change of at least 5, we delineated a TE signature comprising 7521 probeset (PS) overexpressed in TE cells and a hESC signature comprising 1083 overexpressed in hESC. We observed that some pluripotency genes such as OCT4/POU5F1 and NANOG were found in the hESC signature, whereas LIN28, a gene known to repress the let-7 miRNA and to be preferentially expressed in both hESC and oocytes, was also highly expressed in TE cells. The TE signature comprised the transcription factors (TF) GCM1, GATA3, IRX4 and PPARG, suggesting that these TF may have critical roles in TE differentiation. TE cells also displayed a remarkable expression of cell surface molecules such as CD55 and CD53, which could be used to recognize TE cells during in vitro differentiation of hESC into TE cells. A GSEA comparison between the TE samples and 10 hESC samples revealed in TE cells a strong enrichment of gene involved in protein synthesis activity, such as ribosome genes (GO:0005840), structural constituent of ribosome (GO:0003735) and translation (GO:0006412). Finally, specific expression of DNMT genes and cancer/testis antigen genes suggests a unique epigenetic pattern in TE.

Conclusions:

This whole genome expression analysis of TE and its comparison with hESC should improve the understanding of a critical step of the human blastocyst development and the creation of the ICM, i.e. pluripotency. In addition, as TE may play a deleterious role in hESC derivation, the screening of TE specific cell-to-cell signaling could highlight specific pathways to block in order to improve the ICM in vitro development.

Poster No: P081-W
Integrative analysis of transcription factor binding and chromatin modifications data obtained by ChIP-seq in human genome: example of estrogen receptor (ERα) binding sites

[1]**Yuriy L. Orlov**, [1]Mikael Huss, [1]Roy Joseph, [1]Edwin Cheung, [1]Neil D. Clarke, [1]Huck-Hui Ng, [1]Edison T. Liu
[1]Genome Institute of Singapore, Genome #02-01, 60 Biopolis Street, Singapore 138672

Advances in high throughput sequencing technology have enabled the identification of transcription factor (TF) binding sites in genome scale. Such genome wide TF binding maps in human include Oct4, Sox2 and Nanog in stem cells and cancer related transcription factors estrogen receptor (ERα), FOXA1. Key problem of gene expression regulation analysis is detection of functional binding sites responsible for gene activation. TF binds to only a small fraction of sequence motifs or eligible binding sites in the genome. Moreover, this limited number of occupied sites might be significantly different between different cell types. We studied the model of ERα binding in human genome. Previous genome-wide positional analysis showed that a large proportion of in silico predicted ERα binding sites is not occupied in vivo in the MCF-7 human breast cancer cell line. We discuss here interplay between TF binding and chromatin landscape revealed by ChIP-sequencing together with statistical issues. First, the distribution of binding sites across gene is uneven and focused on proximal gene regions. Our analysis has revealed thousands of binding sites in the genome located at gene distal regions for ER binding sites and key developmental regulators. Second, DNA accessibility (nucleosome depletion) facilitates TF binding as maybe measured by different techniques. Epigenetic modifications play important role in regulation of gene expression adding additional complexity to transcription network functioning. We have studied associations between different histone modification using published (GEO NCBI) and in-house data for activation histone marks H3K4me3, H3K4me1, H3K9ac and repressive histone marks H3K27me3 and H3K9me3 together with RNA Pol II sites. We found strong associations between chromatin activation marks and TF binding sites and present it qualitatively. To meet issues of statistical analysis of genome ChIP-sequencing maps we developed computer program to filter out noise signals and find significant association between binding affinity and number of sequence tags. Better computer prediction of TF binding for ERα could be achieved using genome wide chromatin modification data. Our data provide new insights onto dynamical activation of TF binding in human genome by analysis of the changes in chromatin landscape.

2. Genetic and Genomic Databases

(19 May 2010, 1.05 PM - 3.05 PM, Salle Einstein) - Chaired by Yoshiyuki Sakaki and Marie-Paule LeFranc

Poster No: P005-W
THE HUMAN VARIOME PROJECT – PILOT PROJECTS AND PROGRESS

Cotton R.G.H. and collaborators
Convenor, Human Variome Project; Head, Genomic Disorders Research Centre, Howard Florey Institute; Department of Medicine, the University of Melbourne, Melbourne, Australia

The Human Variome Project (2007) (www.humanvariomeproject.org) was initiated in June 2006 (Ring, Kwok et al. 2006) drawing attention to the importance of collection of variation and its phenotype and to develop programs to put this into effect.

The project builds on work and concepts of the HGVS over many years (www.hgvs.org) to focus on all variation associated with disease. The project will include those discovering mutations, their effects and then collect the data making it instantly available for those who need it to inform clinical decisions and research.

The HVP and InSiGHT (www.insight-group.org) has developed a major pilot study to develop procedures and systems to allow effortless flow of de-identified data from the patient/clinic/diagnostic laboratory via curated locus or gene specific databases to central databases/genome browsers such as NCBI, UCSC and EBI. The system will be easily adaptable to other genes and to multiple laboratories, states and countries worldwide.

A country specific collection pilot is underway in Australia and an International Confederation of these countries has been initiated with Korea and China in the application process.

Other pilot studies developed include specific ethical studies related to mutation collection, loading of LSDB content to dbSNP, funding of curation of LSDBs, a system of Microattribution/reward for mutation submission. A high level meeting in Spain (Kaput, Cotton et al. 2009) developed plans to implement the recommendations of the HVP Melbourne meeting (Cotton, Appelbe et al. 2007). The HVP was features recently in relation to Neurogenetic databases in Science (Cotton, Auerbach et al. 2008) and HVP Neurogenetic Database Initiative is being formed after an HVP Neurogenetics forum in Hawaii October 2009.

Poster No: P022-W
Exploring hematological malignancies transcriptome using a comprehensive gene expression profiling resource, Amazonia! Hematology.

[1-3]**Sylvie Tondeur**, [1,2]Tanguy Le Carrour, [1,2]Said Assou, [2]Véronique Pantesco, [1,2]Bernard Klein, [3]Jean-François Schved, [1,2]John De Vos
[1]CHU Montpellier, Institute for Research in Biotherapy, Hôpital Saint-Eloi, Montpellier, F 34000 France ; [2]INSERM, U847, Montpellier, F 34000 France ; [3]CHU Montpellier, Laboratoire d'Hématologie, Hôpital Saint-Eloi, Montpellier, F-34000 France.

Microarray technology can monitor the expression of the whole genome of a cell or a tissue type in one experiment. This tool represents a major technical breakthrough and has known a rapid development in the last years. Its application provided real advances in numerous domains such as in cancerology and has notably contributed to a better understanding of the physiopathology and the evolution of hematological malignancies. Gene expression data accumulate in public repositories such as Gene Expression Omnibus (GEO) or Array Express, in individual web pages or as various supplemental data in the literature. Because of the huge quantity of data generated and the know-how needed for analyses, these published data cannot be routinely accessed.

We have developed the web based tool Amazonia! (http://www.amazonia.transcriptome.eu/), for an easy and free access to a large selection of public transcriptomes. A thematic atlas, Amazonia! Hematology, provides the possibility to query and to visualize the expression of a given gene in representative and selected human transcriptome datasets in the field of haematological malignancies. This atlas provides expression bar plots for single genes, across samples selected from a wide range of normal and malignant cells and tissues. When produced by the same platform type, datasets were renormalized and combined in order to generate series of several hundreds samples. More than 2,000 samples types are annotated, colored and ordered for 9 hematological malignancies and for normal haematopoiesis.

To illustrate the powerfulness of this simple tool, we show how Amazonia! Hematology reveals the specific expression of the T-cell leukemia, homeobox 3 gene TLX3 (HOX11L2) in a subset of T-ALL, or of the Wilm's Tumor gene WT1 in a subset of AML.

Amazonia! Hematology advantageously complements large public repositories by providing a simple way to query a compilation of selected human transcriptome data.

Poster No: P037-W
SNDs (Single Nucleotide Differences) Pollute dbSNP

Juergen Reichardt[1,5], Lucia Musumeci[1,6*], Jonathan W Arthur[2,3*], Florence SG Cheung[1], Ashraful Hoque[4] and Scott Lippman[4]
[1]Plunkett Chair of Molecular Biology (Medicine) Bosch Institute The University of Sydney Medical Foundation Building (K25) 92–94 Parramatta Road Camperdown, NSW 2006, Australia [2] Discipline of Medicine, Sydney Medical School, The University of Sydney, Camperdown, NSW 2006, Australia [3] Sydney Bioinformatics, The University of Sydney, Camperdown, NSW 2006 Australia [4] The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA [5]Presenter

The creation of single-nucleotide polymorphism (SNP) databases (such as NCBI dbSNP) has facilitated scientific research in many fields. SNP discovery and detection has improved to the extent that there are over 17 million human reference (rs) SNPs reported to date (Build 129 of dbSNP). SNP databases are unfortunately not always complete and/or accurate. In fact, half of the reported SNPs are still only candidate SNPs and are not validated in a population. We describe the identification of "bogus" (bs) SNPs in humans, that may contaminate the dbSNP database. These bsSNPs, reported as real SNPs in the database, do not exist as such, but are merely artifacts due to the presence of a paralogue (highly similar duplicated) sequence in the genome. Using sequencing we showed how bsSNPs could originate in two paralogous genes and evaluated samples from a population of 100 individuals for the presence/absence of SNPs. Moreover using bioinformatics, we predicted as many as 8.32% of the biallelic, coding SNPs in the dbSNP database to be bogus.

Our identification of bsSNPs in the database will allow researchers to not only select truly informative SNPs for association studies, but also aid in determining accurate SNP genotypes and haplotypes.

Poster No: P052-W
First genetically characterized mouse model for age-related cataracts is a mutation in the beta-A2-crystallin encoding gene

Jochen Graw[1], Oliver Puk[1], Nafees Ahmad[1], Sibylle Wagner[2], Martin Hrabé de Angelis[2], Helmholtz Center Munich – German Research Center for Environmental Health,
[1]*Institute of Developmental Genetics and* [2]*Institute of Experimental Genetics D-85764 Neuherberg; Germany.*

Age-related cataracts are the major cause of blindness worldwide and have been associated with falls and increased mortality, possibly because of associated systemic conditions. Major risk factors for cataracts are diabetes and UV light. However, the suggested contribution of genetics to the etiology of age-related cataracts is largely unknown.

In a large-scale high-throughput ENU mutagenesis screen we analyzed the offspring of paternally treated C57BL/6J mice for malformation of the eye by non-invasive in-vivo techniques including slit-lamp biomicroscopy, funduscopy and eye size measurement by laser interference biometry. In total, we screened 1700 F1 mice; the most efficient screen was the laser interference biometry at the age of 11 weeks. This test detected 77 variations; 13 mutants could be confirmed so far. One of them was characterized by a clear, but significantly smaller lens without any changes for cornea thickness, anterior chamber depth or aqueous humour size. At the age of 11 weeks, the mean lens axis length is 2.1 mm (+ 0.01 mm) for wild type mice and 1.9 mm (+ 0.03 mm) for heterozygotes. The smaller size of the clear lens was more pronounced in the homozygous mutants (1.7 mm + 0.03 mm), which were fully fertile and viable. The mutation was mapped to chromosome 1 between the markers D1Mit251 and D1Mit253. Using a positional candidate approach, the βA2-crystallin encoding gene Cryba2 was sequenced; a T→C exchange at cDNA position 139 leads to an S47P amino acid exchange.

Histologically, the eye of newborn homozygous mutants showed small vacuoles at the anterior pole of the lens. At the age of three weeks, some clefts appeared at the anterior cortical region; the other main tissues of the eye, cornea and retina, appeared without major changes. Later, at the age of 25 weeks, the lenses of the heterozygous mutants develop a subcapsular cortical cataract, but the lenses of homozygous mutants are completely opaque.

These findings demonstrate the first mutation in the Cryba2 gene in any organism so far. The Cryba2 gene is very closed to the gamma-crystallin gene cluster; mutations in the gamma-crystallin genes have been shown to cause congenital dominant cataracts in mouse and man. Surprisingly, no congenital cataract mutation could be attributed up to now to this gene. If the data on our new Cryba2 allele in the mouse are extrapolated to the situation in human cataract patients, CRYBA2 should be considered as a strong candidate gene for human age-related cataracts, and the slightly smaller size of the lens might be understood as an early biomarker for age-related cataracts. Moreover, data from the Allen Brain Atlas (http://mouse.brain-map.org) demonstrate that Cryba2 is expressed also in the hippocampus. Therefore, mutations in the human CRYBA2 gene are expected also to modify the function of the hippocampus (mainly long-term memory and spatial navigation), and ophthalmologists have to consider additional neurological deficits in these patients.

Poster No: P067-W
A technology platform for isolation and expression profiling of rare cell populations from developing mouse embryos

Sook Peng YAP, Sumantra CHATTERJEE, Wenqing Jean LEE, Hsiao Yun CHAN, Petra KRAUS, Thomas LUFKIN
Stem Cell and Developmental Biology, Genome Institute of Singapore, 60 Biopolis Street, #02-01 Genome, Singapore 138672

Global whole embryo gene expression profiling during mouse embryogenesis has largely been completed. These studies generated data following major fluxes in gene expression but failed to look at specific cell populations. Cells in low abundance are masked by the larger pool of cells found in the developing animal. In order to investigate the molecular mechanisms directing different specialized lineages of cells in vivo, we need to isolate them accurately from other embryonic cells for expression profiling studies.

To achieve that, we set out to establish a technology platform that would allow enrichment for the cells of interest by merging the technologies available from the genomics field with classical molecular genetic approaches using mouse as a model.
A developmental gene with a known expression pattern was engineered endogenously with the enhanced green fluorescent protein (EGFP) by homologous recombination in mouse embryonic stem cells (mESC). The modified mESC were microinjected to generate a transgenic line of mice. The expected EGFP expression pattern was observed in heterozygous embryos at midgestation, and the cell populations expressing the gene of interest were efficiently isolated by Fluorescent Activated Cell Sorting (FACS) to greater than 90% purity. RNA was extracted from the sorted cells and subsequent expression profiling studies done using Illumina Mouse WG-6 Microarrays.

From our work, we have determined the smallest number of cells required for us to obtain the minimum RNA necessary for reproducible gene expression profiling on the Illumina microarray platform. By comparing the expression profiles of EGFP+ and EGFP- cells, we identified a list of genes, including novel and known genes, which are highly expressed in the specific cell lineage isolated.

In summary, a robust strategy to isolate low abundance cell population to high purity for the analysis of spatio-temporal changes in the transcriptome of any specific developing tissue or organ in vivo has been established.

Poster No: P084-W
From HGM10 (1989) to HGM 2010: IG and TR gene concept and IMGT/GENE-DB

[1]**Véronique Giudicelli** and [1]Marie-Paule Lefranc
[1]IMGT®, the international ImMunoGeneTics information system®, Laboratoire d'ImmunoGénétique Moléculaire LIGM, Université Montpellier 2, Institut de Génétique Humaine IGH, UPR CNRS 1142, 141 rue de la Cardonille, 34396 Montpellier cedex 5, France

In 1989, at HGM10 in New Haven (USA), the T cell receptor gamma genes were, for the first time, entered in the newly created Genome DataBase (GDB). This was a first major step as it acknowledged the concept of gene for the antigen receptors, the immunoglobulins (IG) and T cell receptors (TR), despite their unique particularities. Indeed, IG and TR chains are coded by genes belonging to four types, variable (V), diversity (D), joining (J) and constant (C), which show an unusual molecular organization due to the necessity of DNA rearrangements for the IG and TR chain synthesis, in B and T cells, respectively. These rearrangements contribute to the huge diversity and fine specificity of the variable domains of the IG and TR that bind specifically to the antigens, in the adaptive immune response. The potential expressed repertoire is estimated to 2x10 to the twelfth different IG and TR per individual. Owing to these particularities, IMGT®, the international ImMunoGeneTics information system® (http://www.imgt.org), was created in 1989 by Marie-Paule Lefranc at LIGM, Montpellier, France. IMGT® is a high-quality integrated knowledge resource specialized in the IG, TR, major histocompatibility complex (MHC), immunoglobulin superfamily (IgSF), MHC superfamily (MhcSF) and related proteins of the immune system (RPI) of human and other vertebrate species. IMGT/GENE-DB (Nucl. Acids Res., 33, D256-D261, 2005), the IMGT® genome database, was developed to standardize and classify the IG and TR gene data and to manage the related knowledge. The official nomenclature of human IG and TR genes and alleles, based on IMGT-ONTOLOGY, the first ontology for immunogenetics, was approved in 1999 by the HUGO Nomenclature Committee (HGNC) and acknowledged by the WHO-IUIS. Rules for the identification of gene and allele functionality were defined, with each IG and TR gene and allele being represented by an IMGT reference sequence. Another breakthrough is the IMGT unique numbering and its graphical representation, the IMGT Collier de Perles, which allow the standardization per domain type. In March 2010, IMGT/GENE-DB includes the 674 human IG and TR genes (1245 alleles). Five hundred ninety-six genes (1139 alleles) are organized in 7 loci on 4 chromosomes, spanning a total of 6 megabases: IGH (14q32.33), IGK (2p11.2), IGL (22q11.2), TRA (14q11.2), TRD (14q11.2), TRB (7q34) and TRG (7p14) (The Immunoglobulin FactsBook, 2001; The T cell receptor FactsBook, 2001). Seventy-eight orphons (106 alleles) are found outside the main loci. IMGT® gene data are provided on Ensembl Genome Browser (EBI) via a DAS server. IMGT/GENE-DB gene entries are cross-referenced by HGNC database, Genatlas, Entrez Gene (NCBI) and Vega (Wellcome Trust Sanger Institute). IMGT/GENE-DB reference sequences are crucial for the assignment of new alleles of IG and TR from different haplotypes (1,000 genomes project), for gene expression studies in normal and pathologic situations (cDNA high-throughput sequencing) and for biotechnology related to antibody engineering and antibody humanization.

Poster No: P109-W
Antagonism of the Testis-Specific Enhancer of SOX9- an 'anti-testis' effect of DAX1

[1,2]**Vincent R. Harley**, [1,2]Louisa M. Ludbrook, [1]Pascal Bernard, [3]Ryohei Sekido, [4]Dagmar Wilhelm, [1]Stefan Bagheri-Fam, [3]Robin Lovell-Badge
[1]Prince Henry's Institute of Medical Research, Melbourne, Australia, [2]Dept. of Biochemistry & Molecular Biology, Monash University, Melbourne, Australia, [3]Stem Cell Biology and Developmental Genetics, MRC National Institute of Medical Research, London, UK, [4]Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

The role of the orphan nuclear hormone receptor Dosage Sensitive Sex Reversal-Adrenal Hypoplasia Congenita on X chromosome, gene 1 (DAX1) during gonad formation is contentious. Here we investigated an 'anti-testis' effect of DAX1 when it is present at high doses. In humans, duplications of the DAX1 gene at locus Xp21 cause Disorders of Sex Development (DSD), whereby XY individuals develop an ambiguous or female sex phenotype. Exactly how DAX1 duplication causes the failure of typical testicular development is unclear. We hypothesized that, when present in excess, DAX1 must repress the action of early testis-specific genes.

Using a mouse line transgenic for Dax1 with gonad-specific expression, we investigated the expression of the critical testis-specific gene, SOX9. Immunostaining of Dax Tg gonads revealed reduced Sox9 protein levels in gonads with high Dax1 dose. The low Sox9 protein levels may reflect a reduced transcriptional activity of the Testis-Specific Enhancer of Sox9 (TES), (which drives Sox9 transcription in the developing XY gonad). Indeed, TES activity was repressed in vivo and in vitro by higher doses of Dax1. Moreover, in a heterozygous knockout mouse model for Sox9 (50% dose of Sox9 gene); the introduction of extra copies of Dax1 causes the most severe disruption to testicular development as shown by immunohistochemical analysis. In these mice, ovotestes expressing a female gonadal somatic cell marker at both poles are formed in developing XY embryos.

These data support the hypothesis that Dax1 functions in an 'anti-testis' manner by affecting Sox9 expression, when present at higher than typical doses. Using in vitro reporter activation assays and EMSA we provided additional evidence to suggest that the Sox9 transcriptional repression by Dax1 is mediated through a disruption of the ability of the nuclear hormone receptor Steroidogenic Factor-1 to bind to and activate the TES.

With this work we have identified a potential mechanism for disruption of the male-specific sex determination pathway caused by DAX1 duplication and leading to DSD in XY individuals.

Poster No: P144-W
H-InvDB: a comprehensive annotation resource for human transcriptome

[1]**Chisato Yamasaki**, [1]Jun-ichi Takeda, [1]Takuya Habara, [2,3]Makoto Ogawa, [4]Makoto Shimada, [2,5]Kentarou Mamiya, *[1]Akiko Noda, [1]Ryuichi Sakate, [2]Katsuhiko Murakami, [1]Tadashi Imanishi and [1,6]Takashi Gojobori*
[1]Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, [2]Japan Biological Informatics Consortium, [3]DYNACOM Co., Ltd., [4]Fujita Health University [5]Hitachi Software Engineering Co., Ltd., [6]National Institute of Genetics WaterfrontBio-ITResearchBuilding, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan.

H-Invitational Database (H-InvDB: http://www.h-invitational.jp/) is a comprehensive annotation resource for human transcriptome. By extensive analyses of all human transcripts, we provide curated annotations of human genes, transcripts and proteins that include gene structures, alternative splicing isoforms, non-coding functional RNAs, protein functions, functional domains, sub-cellular localizations, metabolic pathways, protein 3D structure, genetic polymorphisms, relation with diseases, gene expression profiling, molecular evolutionary features, protein-protein interactions (PPIs) and gene families/groups. The latest release of H-InvDB (release 7.0) provides annotation for 296,912 human transcripts in 46,499 human gene clusters based on human full-length cDNAs, mRNAs and the reference human genome sequences (NCBI b37.1). H-InvDB consists of three main views, the Transcript view, the Locus view and the Protein view, and six sub-databases; G-integra, H-ANGEL, DiseaseInfo Viewer, Evola, PPI view and Gene Family/Group view. We also provide data mining tools such as "Navigation search", an extended search system that enables complicated searches by

combining 16 different search options (http://www.h-invitational.jp/hinv/c-search/hinvNaviTop.jsp) and "H-InvDB Enrichment Analysis Tool (HEAT)", a data mining tool for automatically identifying features specific to a given human gene set (http://hinv.jp/HEAT/).

We also develop VarySysDB (http://h-invitational.jp/varygene/), a database of annotated genetic polymorphisms on human genes based on gene models of H-InvDB and public polymorphism databases. VarySysDB provides information of human genetic polymorphism and effects derived from a polymorphism on human transcriptome. We analyzed single nucleotide polymorphisms and deletion-insertion polymorphisms from dbSNP, copy number variations from Database of Genomic Variants, short tandem repeats and single amino acid repeats from H-InvDB. We also annotated the polymorphisms altering functional domains (InterPro) and protein 3D-structures (GTOP). VarySysDB consists of two kinds of viewers, GBrowse and Sequence View, to facilitate understanding of the positional relationship among polymorphisms, genome, transcripts, loci and functional domains, providing useful information on polymorphisms affecting gene expression and phenotypes.

3. Genomics, Ethics, Law and Society

(19 May 2010, 1.05 PM - 3.05 PM, Sully 1) - Chaired by Ruth Chadwick / Benjamin Capps

Poster No: P016-W
HUGO Statement on Next Generation Genomic Sequencing

Tamra Lysaght[1]
[1]*Centre for Biomedical Ethics, Yong Loo Lin School of Medicine, #02-04, Clinical Research Centre, 10 Medical Drive, Singapore 117597*

Since HUGO released its Statement on Human Genomic Databases in 2002, developments in next generation sequencing technologies have significantly advanced the possibilities of developing individualized medical treatments and diagnostics along with other biotechnologies that may be used for industrial, military and clinical purposes. With increasing numbers of genome-wide association studies (GWAS), the realisation of these possibilities in the near future is likely to re-invigorate public discussion of the social, political and ethical issues that surround genomics. Many of these issues are familiar; however, some will present new challenges. The Centre for Biomedical Ethics at the National University of Singapore is aiming to assist HUGO's Ethics Committee in generating an in-depth analysis of the issues that are both already under discussion as well as those likely to emerge in response to further developments in GWAS. It is proposed that this analysis will contribute to the formulation of a white paper, tentatively entitled "The Impact of Next Generation Genomic Sequencing on Society, Policy and Law", and a supporting policy statement on whole genome sequencing.

Preparations have commenced on the drafting of these documents, which I aim to discuss at the HUGO's 14th Human Genome Meeting in 2010. To stimulate ongoing discussions in this area, I will address five key issues. The first relates to procedural consent and the right of participants to withdraw from GWAS. The second relates to the new challenges GWAS presents in maintaining confidentiality and protecting the privacy of research participants, their families and the communities they belong to. Third, is the so-called democratisation of DNA and issues arising with the disclosure of GWAS results, not only to research participants, but third parties such as genetically related others and commercial genetics companies. Forth, I shall speak about the governance of GWAS and concepts such as benefit sharing, solidarity, equity and genetic heritage. Finally, I will conclude with some remarks on the deeper cultural impacts of GWAS, particularly with respect to genetic determinism and public understandings of statistical risk and uncertainty.

Attending the HGM 2010 will be tremendously helpful in my research as it will provide me with the opportunity to network with leading scientists in genomics and participate in the presentations and workshops on whole generation sequencing. Attendance will also present opportunities for me to make anecdotal observations about current considerations of the ethical/social issues from the point of view of the scientists working in GWAS, which will help me to undertake the most up-to-date analysis of current concerns in the area. This will benefit my research, and that of others studying the ethical issues surrounding genomics, by not only helping to ensure that my scholarly contributions accurately reflect the state of the art in GWAS, but will open up the work of HUGO to wider audiences interested in genomics research. As a result,

HUGO is set to benefit as the outcomes of my discussions will contribute substantially to the development of the white paper and policy statement on whole genome sequencing. This, in turn, will benefit HUGO in the generation of a public policy position that is defensible on the grounds that it is theoretically-derived and is supported by the extant empirical literature in ethics. I will also contribute a blog on the outcomes of my attendance to HUGO Matters, which will hopefully expand the interest of HUGO's membership in the ethical and social issues surrounding their work.

Poster No: P031-W
Disclosing Genetic Information to Genetic Relatives Without Consent - the Australian Experience

Prof R.J.A. Trent
The National Health and Medical Research Council (NHMRC) of Australia, GPO Box 1421, Canberra ACT 2601, Australia. (Presented by Prof R.J.A. Trent, Chair of the NHMRC Human Genetics Advisory Committee)

The NHMRC is Australia's peak body for supporting health and medical research; for developing health advice for the Australian community, health professionals and governments; and for providing advice on ethical behaviour in health care and in the conduct of health and medical research (www.nhmrc.gov.au).
In 2003 following a joint inquiry by the Australian Law Reform Commission and the NHMRC's Australian Health Ethics Committee a report was issued titled "Essentially Yours - the protection of human genetic information in Australia" - www.austlii.edu.au/au/other/alrc/publications/reports/96/. One of the recommendations of this report was to change the Commonwealth Privacy Act 1988 to allow disclosure of genetic information to genetic relatives without the consent of the patient provided there was reasonable belief that this was necessary to lessen or prevent a serious threat to the life, health or safety of their genetic relatives. Previously this was possible if the threat was both serious and imminent but with DNA predictive testing, the requirement for imminent became a limitation. In this context, it is assumed that for most circumstances a patient would agree to provide this information to family members but, in the unusual case that the patient refused, it was now possible for health practitioners to do this without breaching the Privacy Act. In December 2009, the Office of the Privacy Commissioner and the NHMRC released guidelines that were obligatory before medical practitioners could make use of the amended legislation - www.nhmrc.gov.au/publications/synopses/e96syn.htm.
The guidelines were prepared by an NHMRC working group comprising expertise in medicine, clinical genetics, ethics and a consumer representative. The guidelines took nearly two years to prepare with a number of consultation steps involving the community and health professionals. Two of the NHMRC's principal committees provided input including the Human Genetics Advisory Committee and the Australian Health Ethics Committee.
It is important to note that the changed legislation included specific inclusions and exclusions such as: (1) it is not mandatory for health practitioners to disclose genetic information to genetic relatives, (2) disclosure without consent to non-genetic relatives e.g. spouses or partners is not allowed, (3) a serious threat to an unborn child is not covered, (4) changes impact on privacy legislation but not the health practitioner's duty of confidentiality. (5) In the context of the Australian health system, the changes are to the Commonwealth Privacy Act hence only relate to private patients. To cover public patients, it will be necessary for the different States and Territories to mirror these changes in their own Privacy Acts.

It will be important to monitor how often this provision in the privacy legislation is used, and whether there are problems in relation to this legislation and duty of care or confidentiality.

Poster No: P046-W
A Survey of Scientist and Policy Makers' Attitudes Toward Research on Stored Human Biological Materials in Sri Lanka

[1]**Vajira H. W. Dissanayake**, [1]Madubashini Jayamanne, [1]Dilum Nisansala, [2]Reidar Lie
[1]Human Genetics Unit, Faculty of Medicine, University of Colombo, Sri Lanka [2]Department of Bioethics, National Institutes of Health, Maryland, USA

Ethics issues on research on stored human tissue samples are important in biomedical research. We carried out a survey among biomedical research scientists and policy makers (mainly ethics review committee members) to determine their attitudes concerning controversies surrounding such research. 55 (45.8%) responded. 26 (47.2%) were males. 19 (34.5%) were scientists and 35 (63.6%) were policy makers. 40 (72.7%) favoured giving sample donors the option to provide

advance consent to unspecified future research on their samples, under the condition that future research be approved by an ethics committee. 31 (56.4%) favoured using consent forms that offer a greater variety of consent options for future research than binary yes or no options. Many supported regulations that require that scientists in a sample collections' country of origin be given authorship opportunities and control over research performed on sample collections that are removed from their country of origin with response rates varying from 33 (60.0%) to 50 (90.9%) to the different questions probing these issues. 51 (92.7%) favored involving ethics committees from samples' country of origin in the research approval process when samples have been transported out of their country of origin. 19 (34.5%) felt that when there is a time lag between collection and use of samples the standards at the time of use of samples should be applied for their use. Characterizing attitudes regarding the consent process for collecting stored samples and collaborative research on stored samples such as we have attempted to do in this study may inform policy recommendations.

Poster No: P061-W
Admixture mapping: from paradigms of race and ethnicity, to population history

Sarah E. Ali-Khan[1]; Abdallah S. Daar[1,2,3].
[1]McLaughlin-Rotman Centre for Global Health, University Health Network and University of Toronto, Toronto ON, M5G 1L7, Canada [2]Department of Public Health Sciences and of Surgery, University of Toronto, Toronto ON, M5S 1A8, Canada [3]McLaughlin Centre for Molecular Medicine, University of Toronto, University Health Network and University of Toronto, Toronto ON, M5S 1A1, Canada

Admixture mapping is a whole genome association strategy that takes advantage of population history - or genetic ancestry - to map genes for complex disease. Admixture mapping is beginning to generate important scientific insights. However, because it uses racial/ethnic groupings to examine differential disease risk, there is heightened potential for ethical and social consequences. Here we examine one of the first admixture studies, a scan for Multiple Sclerosis (MS) risk factors in an African American population. We explore perceptions of admixture mapping, its ethical and social implications, and potential risks and benefits. Our data primarily draws on in-depth, one-on-one interviews with leading population-geneticists, bioethicists, and African Americans with MS. Unsurprisingly, our analysis underlined fears about exacerbating racial/ethnic sensitivities as a key concern of admixture mapping. However, a belief in the ongoing scientific value of ancestry and population-based studies, and even the opportunity they may offer for social progress was also highlighted. Thus, such studies may represent a transition point between simplistic notions of race/ethnicity, and more empirically-based understandings of human genomic diversity. Most of the geneticists we spoke to cited social and/or scientific imperatives to move beyond the use of race in genetic research, and clinical practice. To enact this wish, and to make best use of emerging genetic information, our analysis underlined the need for concerted efforts to promote deeper public awareness of population history, and genetics. These undertakings should be co-cultivated by all stakeholders - including social and basic scientists, media, policy-makers, clinicians, and the public themselves.

4. Genome Variation in Disease 1

(19 May 2010, 1.05 PM - 3.05 PM, Sully 2) - Chaired by Christian Jorgensen / Elspeth Bruford

Poster No: P008-W
Complete genome sequencing and analysis of a diploid African-American and Mexican-American genome: implications for personal ancestry reconstruction and multi-ethnic medical genomics

[1]**Francisco M. De La Vega**, [2]Jeremiah D. Degenhardt, [2]Shaila Musharoff, [1]Fiona Hyland, [2]Kasia Bryc, [2]Abra Brisbin, [2]Andy Reynolds, [2]Adam Auton, [2]Simon Gravel, [2]Alon Keinan, [3]Andrew G. Clark and [2,4]Carlos D. Bustamante
[1]Genetic Systems R&D, Life Technologies, Foster City, Ca, USA, [2]Department of Biological Statistics and Computational Biology, Cornell University, [3]Department of Molecular Biology and Genetics, Ithaca, NY, USA, and [4]Department of Genetics, Stanford University.

Understanding the contribution of rare and common genetic genetic variants, environmental exposure, and gene-by-environment interactions to disease susceptibility will likely require multi- and trans-ethnic genome-wide studies that compare completely sequenced genomes of many individuals with and without a particular disease. Of particular importance to this endeavor will be accounting for the role of population stratification at fine scales both in terms of genomic and geographic location. Here, we present results from sequencing, assembly, and genomic analysis of of two diploid genomes from Phase 3 HapMap sequenced to ~20X coverage using LifeTech SoLiD technology. The two donor individuals are of Mexican-American and African-American ancestry and represent the first "admixed" genomes to be completely sequenced to high coverage. We demonstrate that genomic sequencing may provide finer resolution of "admixture breakpoints" by using allele frequency estimates based on Phase 1 of the 1000 Genomes Project. Furthermore, we compare the distribution of functional and putatively neutral genetic variation among 8 sequenced genomes and find that difference in demographic history may account for modest, but statistically significant, differences in distributions of synonymous vs. benign, possibly damaging, and probably damaging non-synonymous coding variants. We also use the comparative personal genomic data sets as well as 1000 Genomes Project data to quantify the extent to which one would expect variants to be private, rare within a given population, or common and segregating among populations.

Poster No: P025-W
Haplotyping and copy number estimation of the highly polymorphic human beta-defensin locus on 8p23 by 454 amplicon sequencing

Marco Groth[1], Stefan Taudien[1], Klaus Huse[1], Andreas Petzold[1], Karol Szafranski[1], Jochen Hampe[3], Philip Rosenstiel[2], Stefan Schreiber[2,3] and Matthias Platzer[1]
[1]Leibniz Institute for Age Research – Fritz Lipmann Institute, Jena, Germany [2]Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Germany [3]Dep. of General Intermal Medicine, Christian-Albrechts-University Kiel, Germany

The beta-defensin gene cluster (DEFB) at chromosome 8p23.1 is one of the most copy number (CN) variable regions of the human genome. Whereas individual DEFB CNs have been suggested as independent genetic risk factors for psoriasis, the role of multisite sequence variations (MSV) is less well understood and so far has been reported for prostate cancer. Studying the correlation of CN and sequence variations with the expression of defensins, a very high density of variations (~1 MSV/41bp) was found within the promoter region of DEFB4.

PCR, cloning and Sanger sequencing allows simultaneous assessment of MSVs and CNs, but is labour and cost intensive as well as prone to methodological bias introduced by bacterial cloning. Here, we report amplicon sequencing of pooled individual PCR products by the 454 technology allowing in-depth determination of MSV haplotypes and estimation of DEFB CNs in parallel.

Six PCR products spread over ~87 kb of DEFB and harbouring 24 known MSVs were amplified from DNA samples of lymphoblastoid cell lines, pooled and sequenced on a Roche 454 GS FLX sequencer. We inferred 22 haplotypes ranging between 2 and 7 per amplicon. Two additional sequence variations were detected. Minimal CNs were estimated from the ratio of haplotype calls (HC) and compared to absolute CNs determined by alternative methods. The 454 haplotyping results were compared to those by cloning/Sanger sequencing. Intrinsic problems related to chimera formation during PCR and differences between haplotyping by 454 and cloning/Sanger sequencing are discussed. Deep amplicon sequencing using the 454 technology yield thousands of HCs per amplicon for an affordable price and may represent an effective method for parallel haplotyping and CN estimation in small to medium-sized cohorts. The presented method represents a valuable approach to facilitate further studies of highly CN variable loci such as the beta-defensin locus, especially by identifying haplotypes in regions interfering with the expression of genes located therein.

Poster No: P040-W
Common and Rare Copy Number Variation in Metabolic Syndrome

Kati Kristiansson[1,2,3], Chris Barnes[2], Markus Perola[1,3], Ida Surakka[1,3], Elisabeth Widen[3], Eero Kajantie[4], Johan Eriksson[4,5], Jorma Viikari[6], Mika Kähönen[7], Terho Lehtimäki[8], Olli Raitakari[9], Anna-Liisa Hartikainen[10], Paul O'Reilly[11], Ulla Sovio[11], Aimo Ruokonen[12], Marjo-Riitta Järvelin[11,13], Anneli Pouta[13], Antti Jula[14], Veikko Salomaa[4], Samuli Ripatti[1,3], Matthew Hurles[2], Leena Peltonen[1,2,3,15,16]
[1]Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland, [2]The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK, [3]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Finland, [4]Dept. of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland, [5]Helsinki University Central Hospital, Helsinki, Finland, [6]Dept. of Medicine, University of Turku and Turku University Hospital, Finland, [7]Department of Clinical Physiology, University of Tampere and Tampere University Hospital, Finland, [8]Department of Clinical Chemistry, University of Tampere and Tampere University Hospital, Tampere, Finland, [9]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and the Department of Clinical Physiology, Turku University Hospital, Finland, [10]Department of Obstetrics and Gynecology, University of Oulu, Finland, [11]Division of Epidemiology, Public Health and Primary Care, Imperial College London, United Kingdom, [12]Department of Clinical Chemistry, University of Oulu, Finland, [13]Department of Children, Young People and Families, National Institute for Health and Welfare, Oulu, Finland, [14]Dept. of Chronic Disease Prevention, National Institute for Health and Welfare, Turku, Finland, [15]Department of Medical Genetics, University of Helsinki, Finland, [16]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

Metabolic syndrome (MetS) is defined as a group of metabolic risk factors, including dyslipidemia, central obesity, insulin resistance, and hypertension. The prevalence of this common age-dependent disorder has greatly increased during the past decades. Biological mechanisms and genetic loci underlying MetS are largely unknown. The overlap between MetS, coronary heart disease (CHD), and Type 2 diabetes suggests that the same genetic determinants may contribute to their complex pathophysiology.

We analyzed copy number variation (CNV) in the genome, with two distinct strategies to assess the contribution of common and rare CNVs as risk factors for MetS. Probe signal intensity data from Illumina human610-quad beadchip, which is enriched for CNV-targeted probes, was analyzed for 1,100 MetS cases and their matched controls. The search for common CNVs utilized a recent high-density genomic map of known CNVs from a study of individuals of European and African ancestry: We analyzed CNV-tagging SNPs from this reference set in our MetS cases and controls. In parallel to our analysis of common CNVs, we searched for rare CNV associations, concentrating on those overlapping with genes or covering large genomic regions, with the hypothesis that these structural variants may have large effects on individual phenotypes.

In our MetS case-control analysis, two of the CNV tag-SNPs indicated loci harboring possible copy number variant association with metabolic syndrome as whole, but this was not replicated in a meta-analysis of three large population cohorts. We also analyzed the CNV tag-SNPs and the distinct component traits of MetS, and identified several interesting susceptibility loci especially for high-density lipoprotein (HDL) levels. Some of these loci, such as ABCA1, have been implicated in previous studies, and several more, such as WWOX, with relevant biological function in lipid metabolism, were observed. Our analysis of rare CNVs identified a novel gene region, implicated in insulin resistance and energy metabolism, where a few of our study subjects, who are located at the far ends of the population distribution of relevant phenotypes, carry large CNVs. These results provide ground for replication studies in large population cohorts and hopefully eventually elucidate the genetic background of MetS, CHD, and their risk factors.

Poster No: P044-T
Aneuploidy and DNA replication stress cooperatively destabilize somatic genome and may lead to neurodegeneration in Alzheimer's disease brain

Yuri B Yurov, Svetlana G Vorsanova and Ivan Y Iourov
Mental Health Research Center, Russia Federation

There are two commonly accepted models for cell death in the Alzheimer's disease (AD) brain based on (i) cell cycle failure and (ii) chromosome missegregation. We propose that aforementioned models may be transformed into a unified one featuring genetic/genomic instability linked to neurodegeneration and is based on intriguing relationship between somatic genome instability and neuronal cell death. According to this unified theory, genetically damaged, abnormal and aneuploid neurons accumulated during neurogenesis in the developing and adult brain have to demonstrate abortive cell cycle re-entry and DNA replication stress. Here we analysed such remarkable biomarkers of genomic instability as mosaic

aneuploidy, which is an indicator of large-scale genomic alterations due to mitotic chromosome non-disjunction, and abortive DNA replication as an indicator of abnormal neurogenesis and DNA replication stress. Different regions of the AD post-mortem brain were monitored by molecular neurocytogenetic approaches (interphase MFISH and MCB FISH, FISH-based DNA replication assays). Aneuploidy was observed in 2-15% of neural cells and replicating DNA signals (singlet/duplet) were present in 0.4-3.7% of neural cells of hippocampus and prefrontal cortex of the AD brain. The frequency of abnormal neural cells (aneuploid and abnormally replicated) affecting different chromosomes was higher in the brain area affected by neurodegeneration (hippocampus, prefrontal cortex) than in less-affected area (cerebellum). These data indicate that molecular cytogenetics allows analyzing genomic landscapes at single-cell level and at molecular resolutions in the postmortem brain and may provide for missing information about genome variation/instability and neuronal dysfunction/death in the AD brain. Authors were supported by Philip Morris USA, Inc.

Poster No: P055-W
Complement receptor 1 variability is associated with increased risk for Alzheimer's disease in an extended Flemish-Belgian population

[1,2]**Nathalie Brouwers**, [1,2]Caroline Van Cauwenberghe, [1,2]Karolien Bettens, [2,3]Sebastiaan Engelborghs, [1,2]Helen Van Miegroet, [1,2]Ana Gil Montoya, [1,2]Karin Peeters, [1,2]Maria Mattheijssens, [4]Rik Vandenberghe, [2,3]Peter P. De Deyn, [1,2]Marc Cruts, [1,2]Kristel Sleegers, [1,2]Christine Van Broeckhoven
[1]Neurodegenerative Brain Diseases Group, Department of Molecular Genetics, VIB, Antwerpen, Belgium [2]Institute Born-Bunge, University of Antwerp, Antwerpen, Belgium [3]Department of Neurology, ZNA Middelheim; Antwerpen, Belgium [4]Department of Neurology, University Hospitals Leuven and University of Leuven (KUL), Leuven, Belgium

Alzheimer's disease (AD), the most frequent cause of neurodegenerative dementia, is a genetically complex disorder for which until shortly only one major genetic risk factor was known, i.e. the apolipoprotein E gene. Recent GWA studies in over 25000 individuals identified several interesting novel candidate genes that reached genome-wide significance levels. One of these genes is the complement receptor 1 gene (CR1), encoding a large transmembrane receptor that is crucial in immunity by regulation of the complement cascade. Here we report an initial follow-up of the GWA finding in a Flemish-Belgian group of 1060 patients and 890 control individuals. We investigated the effect of simple (SNPs) as well as complex (structural) variations in the CR1 locus on the genetic risk of AD. Genetic association analysis of 30 tagSNPs spanning the CR1 locus confirmed the association observed in the GWA studies, with signals located at the 5' end (3 SNPs) and the 3' end (4 SNPs) of the CR1 gene. Two SNPs known to be associated with the number of CR1 molecules present on erythrocytes were not associated with AD. In addition, different isoforms of CR1 protein are known to be expressed that differ from each other in the number of complement factor (C3b and C4b) binding sites. These different isoforms are the result of structural variation at the genomic level. The CR1 locus encompasses a number of copies of a low copy repeat (LCR). Using Multiplex Amplicon Quantification (MAQ) we assessed the copy number polymorphisms (CNP) that arises from this LCR in AD patients and control individuals and preliminary data showed that increased copy number is associated with an increased risk for AD. Taken together, our study confirms and extends the association of CR1 variants with AD. Moreover we showed structural polymphisms to be involved in AD risk, with longer isoforms i.e. isoforms with more complement factor binding sites, conferring a higher risk for AD.

Poster No: P070-W
Genome-wide association study using microsatellites identifies a new psoriasis susceptibility locus on chromosome 6

[1]**Hidetoshi Inoko**, [1]Akira Oka , [1]Hideki Hayashi, [1]Akira Ozawa and [2]Stepahn Beck
[1]Department of Molecular Life Science, Division of Basic Medical Science and Molecular MedicineTokai University School of Medicine, 143 Shimokasuya, Isehara、Kanagawa-Pref、259-1193 Japan [2]UCL Cancer Institute, Paul O'Gorman Builing, University College London, 72 Huntley Street, London WC1E 6BT UK

Psoriasis is a common skin disorder characterized by inflammatory cell infiltration and hyperproliferation of epidermal cells. To identify novel psoriasis susceptibility loci, we conducted a genome-wide association study (GWAS) of 25,946

microsatellites in 1377 cases and 1091 controls, representing four populations of Asian and European ancestry. We find a microsatellite D6SX on chromosome 6 to be significantly associated (P = 3.21E-10, OR = 2.3) with disease risk in the Asian (Japanese and Mongolian) but not European (French and Spanish) cohorts. A single SNP, D6SY around the microsatellite D6SX showed statistical association with the disease, suggesting that this D6SY is directly involved in the susceptibility to psoriasis. D6SY is located 1 kb upstream of the first codon of an 'orphan' G-protein coupled receptor (GPCR-Z). Using quantitative expression analysis of sorted peripheral blood mononuclear cells, we find this GPCR-Z gene expression to be up-regulated in fetal skin and naïve CD4+ T cells in individuals positive for GPCR-Z, implicating a T cell-mediated pathway involving GPCR-Z in susceptibility to psoriasis. Relatively young disease alleles as exemplified by D6SY (5% frequency in healthy Japanese population), which has recently happened to be generated on the frequent SNP haplotype in the history of human population and still remained low in allelic frequency, cannot be detected by conventional tag SNP or SNP haplotype association analysis. Therefore, microsatellite-based genome-wide association study provides a complementary tool for genetic dissection of common and rare diseases.

5. Genome Variation in Evolution

(19 May 2010, 1.05 PM - 3.05 PM, Sully 3) - Chaired by Martin Hibberd / Agnes Rotig

Poster No: P011-W
Genetic Interactions Reveal a Novel B-Cell Signaling Pathway in Systemic Lupus Erythematosus

Angélica M. Delgado-Vega[1]*, Casimiro Castillejo-López[1]*, Jerome Wojcik[2]*, Sergey V. Kozyrev[1], Elena Sánchez[3], David Pöllmann[1], Serena Fineschi[1], Juan R. López-Egido[1,4], Nicolás Domínguez[5], Judith A. James[5], Joan T. Merrill[6], Jennifer A. Kelly[4], Kenneth M. Kaufman[7], Kathy Moser[4], Gary Gilkeson[8], Bernardo A Pons-Estel[9§], Sandra D'Alfonso[10], Torsten Witte[11], José Luis Callejas[12], John B Harley[4], Patrick Gaffney[4], Javier Martin[3]*, Joel M. Guthridge[5], and Marta E. Alarcón-Riquelme*[1,4,13]

[1]Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University [2]Department of Stratified Medicine Bioinformatics, Merck Serono International, SA., Geneva, Switzerland [3]Instituto de Biomedicina y Parasitología López-Neyra, CSIC, Granada, Spain [4]Arthritis and Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA [5]Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA [6]Clinical Pharmacology Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA. [7]US Department of Veterans Affairs Medical Center, Oklahoma City, OK [8]Department of Medicine, Division of Rheumatology, Medical University of South Carolina, Charleston, SC, USA [9]Sanatorio Parque, Rosario, Argentina. [10]University of Eastern Piedmont, Novara, Italy. [11]Clinic of Immunology and Rheumatology, Medical School Hannover. University of Hannover, Germany [12]Hospital Clínico San Cecilio, Granada, Spain [13]GENyO, Andalucian Center for Genomics and Oncological Research Pfizer-University of Granada-Junta de Andalucía, Granada, Spain

BACKGROUND:
Epistasis or genetic interaction might explain larger genetic effects on the susceptibility to diseases than single-gene associations and help define functional pathways with potential therapeutic targets.

AIM:
To identify genes that modify the susceptibility to SLE through their interaction with the B-cell scaffold protein with ankyrin repeats gene (BANK1).

METHODS:
We searched for genetic interactions in the Affymetrix 100k genome-wide scan performed in 256 cases and 515 controls from Sweden. A subsequent replication study included two independent multicenter cohorts of European-Americans (n=676 cases and 850 controls) and Europeans (n=1265 SLE cases and 1506 controls). We developed a genotypic interaction test based on contingency tables for all possible genotype combinations between pairs of SNPs with r2 <.80 and calculated a Pearson S score of interaction association and its chi-squared P value. Each interacting combination was tested against the hypothesis of independence to derive an epistasis score (Se) and a P value (Pe) was obtained through permutation.

RESULTS:
BANK1 showed genetic interactions with 29 genes, including the B-cell tyrosine kinase (BLK) and the inositol 1,4,5-triphosphate receptor 2 (ITPR2). One fifth of SLE patients (21%) vs. 8 % of controls were homozygous for the risk alleles of polymorphisms in these three genes with a significant epistatic effect (Pe < 0.0002). The interactions BANK1xITPR2 and BANK1xBLK were replicated in two independent European-American (P = 2.1 x 10-6) and European sets (P = 4.11 x 10-9). The data was verified using multifactor dimensionality reduction (MDR). Moreover, BLK co-immunoprecipitated and co-localized with BANK1 in co-transfected HEK-293T. Exogenous expression of BANK1 in human Daudi B cells curbed BLK from reaching the plasma membrane with the subsequent accumulation in cytoplasmic compartments. Expression of BANK1 and BLK but not ITPR2 was modulated by IFNα.

CONCLUSIONS.
BANK1, BLK and ITPR2 are genetically and functionally interacting partners and through their protein-protein interactions might result in a novel B-cell signaling pathway regulated by type I interferons. This pathway may affect B-cell responses to self-antigens in human lupus.

Poster No: P028-W

Recombination-based genomics: A new method to exploring genetic variation through inference of recombination events

J. Bertranpetit
Institut de Biologia Evolutiva (UPF-CSIC), Spain

Recombination is a main force shaping genome diversity. It is often overlooked that it contains phylogenetic information: a recombination event creates a junction between two parental sequences that is transmitted to the subsequent generations much as a mutation is. In this paper, we present an algorithm, implemented in the IRiS program, to detect past recombination events in extant sequences, with specificity of parental and recombinant sequences. The algorithm detects recombination events from tree incompatibilities found along the sequence. We have validated and calibrated the algorithm for the human genome given human demographic history and the human recombination model by means of coalescent simulations implementing a standard model of human demography. Among the possible parameters that can be used to run IRiS, we chose the set that optimized simultaneously the false positive rate, sensitivity and accuracy in placing the recombination event in the sequence.

We discovered that this method saturates when the sequence set to which it is applied contains a large number of recombinations. Younger recombination events (such as those happening in the last 500 generations of the coalescent simulation, or created instantly from extant, actual sequences) are detected by IRiS with greater sensitivity, which could provide markers of close relationships between populations or even between individuals. Regions for which sperm typing has been used to recognize recombinants give a very good concordance.

We also applied IRiS to haplotypes for 18 X-chromosome regions in HapMap Phase 3 populations. The regions were selected to be free of genes, segmental duplications, and copy number variants. Recombination events detected for each individual can be recoded as binary allelic states and combined into recotypes, which can be analyzed as haplotypes are. Principal component analysis and multidimensional scaling based on recotypes reproduced the relationships between the 11 HapMap Phase III populations that can be expected from known human population history, further validating IRiS.

Poster No: P043-W
Why is sickle cell uncommon in the Mediterranean?

Bridget S Penman[1], Oliver G Pybus[1], David J Weatherall[2] and Sunetra Gupta[1]
[1]*University of Oxford, Zoology Department, South Parks Road, Oxford, United Kingdom, OX1 3PS* [2]*Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, United Kingdom OX3 9DS*

Several human genetic disorders of haemoglobin have risen in frequency due to the protection they offer against death from malaria, sickle cell anaemia being a canonical example. We have used mathematical models to address the issue of why this highly protective mutant, present at high frequencies in sub-Saharan Africa, is uncommon in Mediterranean populations which instead harbour a diverse range of thalassaemic haemoglobin disorders. We demonstrate that these contrasting profiles of malaria-protective alleles can arise and be stably maintained by two well-documented phenomena: an alleviation of the clinical severity of α and β-thalassemia in compound thalassaemic genotypes, and a cancellation of malaria protection when α-thalassemia and the sickle cell trait are co-inherited. The complex distribution of globin mutants across Africa and the Mediterranean can therefore be explained by their specific intracellular interactions.

(Penman BS, Pybus OG, Weatherall DJ, Gupta S. 2009 Epistatic interactions between genetic disorders of hemoglobin can explain why the sickle cell gene in uncommon in the Mediterranean, PNAS 106: 21242-6)

Poster No: P058-W
Genomic study of the chromosome 9 locus linked with FTLD and ALS

[1,2,6]**Ilse Gijselinck**; [3,6,7]Sebastiaan Engelborghs; [1,2,6]Githa Maes; [1,2,6]Ivy Cuijt; [1,2,6]Jamila Ben-Azza; [1,2,6]Veerle Bäumer; [1,2,6]Karin Peeters; [1,2,6]Maria Mattheijssens; [1,2,6]Geert Joris; [4,8]Patrick Cras; [5,6]Jean-Jacques Martin; [3,6,7]Peter P. De Deyn; [1,2,6]Samir Kumar-Singh; [1,2,6]Christine Van Broeckhoven; [1,2,6]Marc Cruts
[1]Neurodegenerative Brain Diseases Group, Department of Molecular Genetics, VIB, Antwerpen, Belgium; [2]Laboratory of Neurogenetics, Institute Born-Bunge, Antwerpen, Belgium; [3]Laboratory of Neurochemistry and Behavior, Institute Born-Bunge, Antwerpen, Belgium; [4]Laboratory of Neurobiology, Institute Born-Bunge, Antwerpen, Belgium; [5]Laboratory of Neuropathology, Institute Born-Bunge, Antwerpen, Belgium;[6]University of Antwerp, Antwerpen, Belgium; [7]Memory Clinic and Department of Neurology, ZNA Middelheim, Antwerpen, Belgium; [8]Division of Neurology, University Hospital Antwerp, Antwerpen, Belgium

Frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS) are two severe neurodegenerative disorders. FTLD has a prevalence close to that of Alzheimer's disease in the population below age 65 years whereas ALS is the most common motor neuron disorder. FTLD and ALS belong to an overlapping clinicopathological spectrum of disorders, that might have common genetic etiologies. Both diseases share presence of TDP-43 inclusions in affected neurons, which suggests that they share overlapping disease mechanisms. A positive family history is observed in up to 50% of FTLD patients and in 5 to 10% of ALS patients indicating a significant genetic contribution to disease aetiology. Although recent findings have identified several genes, these do not fully explain the genetic aetiology indicating that other genes remain to be identified. One such unidentified gene resides at chromosome 9p13-p21, i.e. the ALSFTD2 locus that is causally linked in 13 families with FTD and ALS worldwide, suggesting that this gene is likely to explain the disease in a substantial fraction of patients.

We identified an extended Belgian family with autosomal dominant FTLD-ALS, DR14. The index patient with clinical FTLD showed TDP-43 positive FTLD pathology. A genome-wide linkage study showed significant linkage to a region at chromosome 9, overlapping with the ALSFTD2 locus (multipoint LOD score = 3.38) and nearly significant linkage with a second locus at chromosome 14q31-32 (multipoint LOD score = 2.79), suggesting that a major modifier gene might reside in this region. We excluded mutations in all genes in the ALSFTD2 locus by standard genomic sequencing and/or transcript analyses. We also sequenced 340kb of the highest conserved elements in mammalians. In addition, we completely sequenced the 130kb linkage disequilibrium block in which association was found in a recently published ALS genome-wide association study, at the same genomic region. The absence of mutations in all these sequenced regions suggested that the underlying genetic defect resides outside these regions or is undetectable using standard mutation identification strategies. Therefore we performed other approaches. First, chromosome-specific copy number variation (CNV) analysis using array-CGH did not reveal segregating pathogenic CNVs. Second, we are currently performing macro-restriction mapping to obtain more information on the presence of more complex rearrangements including inversions. In addition, we have determined the complete genome sequence of selected patients of family DR14 and are currently analyzing these data. Identified genetic variants in the linked region will be analyzed in the complete family and in unrelated patients with FTLD and/or ALS. This strategy will allow for the detection of simple DNA variations and complex structural variations, maximizing the chance of identifying the ALSFTD2 genetic defect and a modifying gene at chromosome 14q. This (these) novel gene(s) will significantly contribute to the understanding of neurodegenerative disease mechanisms in FTLD, ALS and associated neurodegenerative disorders.

Poster No: P073-W
Copy number variations in Korean population and their evolutional and functional implications

Yim SH[1], Hu HJ[1,2], Kim JH[1,2], Jung SH[1,2] and Chung YJ[1,2 *]
[1]Integrated Research Center for Genome Polymorphism, [2]Department of Microbiology, The Catholic University of Korea, School of Medicine, 505 Banpo-dong, Seocho-gu, Seoul, Korea.

Recent discovery of the copy number variation (CNV) in normal individuals has been widening our understanding of genomic variation. However, most of the reported CNVs have been identified in Caucasians, which may not be directly applicable to people of different ethnicities. To profile CNV in East-Asian population, we screened CNVs in 3,578 healthy, unrelated Korean individuals using the Affymetrix Genome-Wide Human SNP array 5.0. We identified 144,207 CNVs using a pooled dataset of 100 randomly chosen, Korean females as a reference. The average number of CNVs per genome was 40.3, which is higher than that of CNVs previously reported using lower-resolution platforms. The median size of CNVs was 18.9 Kb (range 0.2 – 5,406 Kb). Copy number losses were 4.7 times more frequent than copy number gains. CNV regions (CNVRs) were defined by merging overlapping CNVs identified in two or more samples. In total, 4,003 CNVRs were defined encompassing 241.9 Mb accounting for ~ 8 % of the human genome. A total of 2,077 CNVRs (51.9%) were potentially novel. Known CNVRs were larger and more frequent than novel CNVRs. Sixteen percent of the CNVRs were observed in ≥ 1% of study subjects and 24% overlapped with the OMIM genes. A total of 476 (11.9%) CNVRs were associated with segmental duplications. To our knowledge, this is the largest-scale profiling of CNV in East-Asian population using high-resolution SNP arrays. CNVS/CNVRs identified in this study will be valuable resources for studying human genome diversity and its association with disease.

Poster No.: P090-W
Revealing origins of human populations embedded in their genomes.

Mait Metspalu[2], Doron M Behar[1], Bayazit Yunusbayev[2,3], Saharon Rosset[4], Gyaneshwer Chaubey[2], Irene Gallego Romero[6], Toomas Kivisild[6], Karl Skorecki[1,5] and Richard Villems[2]
[1]Molecular Medicine Laboratory, Rambam Health Care Campus, Haifa 31096, Israel [2] Estonian Biocentre and Department of Evolutionary Biology, University of Tartu, Tartu 51010, Estonia [3]Institute of Biochemistry and Genetics, Ufa Research Center, Russian Academy of Sciences, Ufa 450054, Russia [4]Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel [5]Rappaport Faculty of Medicine and Research Institute, Technion, Israel Institute of Technology, Haifa 31096, Israel [6]Leverhulme Centre of Human Evolutionary Studies, The Henry Wellcome Building, University of Cambridge, Fitzwilliam Street, Cambridge, CB2 1QH, UK

The onset of the era of analyses of dense marker sets covering the whole genome has revolutionised the field of (human) population genetics. Driven largely by the needs of biomedical research the new data is helping to unveil our demographic past outlined by the study of mtDNA and Y-chromosome variation during the past ca. 20 years. Here we have analysed (Illumina 650K SNPs) over 500 new Eurasian samples together with the publicly available database (HGDP panel) and illustrate the power of full genome analyses by addressing two specific questions: i) the place of the genetic variation of the Jewish Diaspora in the genetic landscape of West Eurasia, and ii) genetic origins of the Munda speakers of India. We use principal component and structure-like analyses to reveal the structure in the genome wide SNP data.

We have analysed 14 worldwide Jewish communities covering the vast geographic span of Jewish Diaspora together with 68 non-Jewish populations, of which 25 have not been previously reported. The reference populations now represent all the Diaspora host populations as well as current Middle East and North African populations. We show that most Jewish samples form a relatively tight cluster which falls within the genetic variation of the contemporary Levantine populations. The patterns of ancestry components of the Ashkenazi and to a slightly lesser extent Sephardi and Moroccan Jews contain signal for some admixture with the European populations and are very similar to the pattern observed among the Cypriots. In contrast, the most distant Diaspora communities in India and Ethiopia cluster together with their host populations.

Alternative models put the origins of Munda languages speakers either in South Asia (the Munda speakers sport exclusively autochthonous South Asian mtDNA variants) or to Southeast Asia where the other Austro Asiatic languages

are spread. Y-chromosome variation supports the latter model through sharing of hg O2a in both regions. We show that in addition to the dominant ancestry component shared between the Indian Dravidic and Munda speakers the latter retain (up to 30%) an ancestry component otherwise prevalent in East Asia. There is no widespread sign of South Asian ancestry component in Southeast Asia. This provides genomic support to the model by which Indian Austro-Asiatic populations derive from dispersal from Southeast/East Asia followed by an extensive admixture with local Indian populations.

In conclusion we have shown effective use of whole genome genetic variation data in resolving thus far elusive problems in our demographic past.

6. Cancer Genomics I

(20 May 2010, 12.55 PM - 2.55 PM, Salle Pasteur) - Chaired by Philippe Broet / Bernard Klein

Poster Number P003-T
microRNA expression profiling in Mexican breast tumors differentiates normal and tumor tissue and identifies novel microRNAs involved in breast cancer.

Hidalgo-Miranda A[1], Rodriguez-Cuevas S[2], Quintanar-Jurado V[1], Bautista-Pina V[2], Ciceron-Arellano I[1], Carrillo-Sanchez K[1], Jimenez-Sanchez G[1], Uribe-Figueroa L[1], Romero-Cordoba S[1]
[1]Laboratorio de Oncogenomica, Instituto Nacional de Medicina Genomica, Mexico, [2]FundacionMexicana de Fomento Educativo para la Prevencion y Deteccion Oportuna del Cancer de Mama, Mexico.

Breast cancer represents the first cause of cancer-related death in Mexican women, and the incidence of this tumor is increasing in the whole Latin American region, particularly in younger women. However, there is limited information about the genomic alterations involved in the development of breast tumors in Latin American patients. Genomic analysis of breast tumors has led to the development of new tools for a better classification and risk prediction. In particular microRNA expression signatures are capable to differentiate normal and breast cancer tissues and can define specific clinico-pathological phenotypes in breast tumors. However, to our knowledge, analysis of miRNA expression patterns has not been reported in tumors from a Latin American population. In order to determine the microRNA expression profile in breast cancer samples from Mexican patients, we analyzed the expression of 667 miRNAs in 25 tumors and 20 adjacent normal tissues using TaqMan low-density arrays. All samples contained more than 80% tumoral tissue and the RNA integrity number of the total RNA in all samples was higher than 6. The expression of specific microRNAs (miR-21 and miR-10b) was further validated in an independent set of 10 tumors. 122 miRNAs showed significant differential expression in breast tumors compared to the normal adjacent tissue. Importantly, the role of 60% of these miRNAs have not been previously reported in breast cancer. The miRNAs with the most significant (P≤0.01; Fold Change: 2) expression changes were: miR-204, miR-129-3p, miR-488, miR-452, miR-139-3p, miR-145, miR-495, miR-668, miR-136*, miR-380*, miR-206, miR-770-5p, miR-145* (down-regulated), and miR-184, miR-210, miR-492, miR-431, miR-96, miR-142-3p, miR-25*, miR-149*, miR-190b, miR-592, miR-454* (up-regulated). Down-regulation of miR-10b and over-expression of miR-21 was confirmed with an independent TaqMan miRNA assay. Bioinformatic analysis of mRNA targets of the down-regulated miRNAs identified oncogenes like ERBB2, YY1 and several MAP kinases, while the over-expressed miRNAs regulate known tumor-suppressors like FOXA1 and SMAD4. Pathway analysis identified that some biological process which are important in breast carcinogenesis can be affected by the altered profile of miRNA expression, including signaling through MAP kinases and RAS, as well as biological processes like programmed cell death and ERBB2-ERBB3 signaling. Our data represents one of the firsts analyses of microRNA expression patterns in a Latin American population, identifying the altered expression of several microRNAs whose altered expression might have an important impact on cancer-related cellular pathways and whose role in breast cancer has not been previously described.

Poster No: P020-T
Copy number analysis by high-density DNA array revealed common and specific genetic changes in subtypes of colorectal cancer

Chien-Hsing Lin, Shih-Ching Chang, Ya-Hui Chang, Wei-Shone Chen and Shih-Feng Tsai
Division of Molecular and Genomic Medicine, National Health Research Institutes, Taiwan

Colorectal cancer (CRC) is one of the major leading causes of cancer deaths around the world, and is arisen from the accumulations of multiple somatic and germline mutations and environmental factors. CRC tumor is a heterogeneous disease, and ~15% CRC samples are characterized to carry genetic mutations and/or epigenetic modifications, causing mismatch repair deficiency (MMR) and resulting in microsatellite instability (MSI). The majority of CRC samples (85%) have functional MMR systems and microsatellite stabilities (MSS), but show numerical chromosomal abnormality based on previous karyotyping and structural changes as detected by comparative genomic hybridization. Currently, single-nucleotide polymorphism (SNP) array can be used to detect subtle copy number (CN) alteration at the genome-wide scale. CN variations are known to contribute to the levels of gene expressions, and thus fine-scale CN profiling of CRC can improve the knowledge about CRC tumorgenesis.

Due to the lack of high-resolution CN alteration analysis in different CRC subtypes, especially from Han Chinese population, 1173 CRC tumors were collected from the Taiwan population, and can be divided into 75 high-frequency MSI (MSI-H, 6.4%), 96 low-frequency MSI (8.2%) and 1002 MSS (85.4%) based on microsatellite instability assay. Affymetrix SNP 6.0 array, consisting of ~1.8 million probes, was used to detect genome-wide CN alterations of 16 MSI-H and 13 MSS CRC samples and CN variations of 424 randomly-selected general controls from Taiwan. High-density array probes allow the detecting CNVs of 20kb in size. Nearly a quarter of CN alterations were smaller than 100kb, which might have been missed in previous studies due to low-resolution technologies. Some CN alterations were commonly found in all CRC samples, e.g., subtle CN gains in chr2p, chr14q and chr22p, compared to general controls. There were distinctive differences in the CN coverage between CRC MSS and MSI-H subtypes (300Mb v.s. 42Mb per genome, p-value=0.001). Chr7, 8q, 13 and 20 gains, and 8p and 18 losses were frequently found in MSS but not MSI-H. However, there were significant CNV variations among CRC MSS samples. Cancer-associated genes were found to show variable copy number in these CRC samples, e.g., EGFR gain and MALT1 loss, but the biological functions of the majority of these genes remain unknown. These CRC common and subtype-specific CN-altered genes should be seriously considered when investigating the mechanism of heterogeneous CRC tumorgenesis, and might be used as candidate markers in the drug therapy studies.

Poster No: P035-T
Genetic diagnosis of familial breast cancer using clonal sequencing

GR Taylor[1,2], JE Morgan[1], IM Carr[1], E Sheridan[1], CE Chu[2], B. Hayward[1], N Camm[2], HA Lindsay[2], CJ Mattocks[3], AF Markham[1], DT Bonthron[1] &
[1]University of Leeds Institute of Molecular Medicine, Wellcome Trust Brenner Building, St James's University Hospital, Leeds LS9 7TF, [2]Department of Clinical Genetics, Leeds Teaching Hospitals Trust St James's University Hospital, Leeds LS9 7TF, [3]National Genetics Reference Laboratory (Wessex), Salisbury District Hospital, Salisbury, SP2 8BJ.

Using conventional Sanger sequencing as a reference standard, we compared the sensitivity, specificity and capacity of the Illumina GA II platform for the detection of TP53(OMIM *191170), BRCA1(OMIM *113705) and BRCA2 (OMIM *600185) mutations in established tumour cell lines and DNA from patients with germline mutations. A total of 656 coding variants were identified in 4 cell lines and 65 patient DNAs. All of the known pathogenic mutations (including point mutations and insertions/deletions of up to 16 nucleotides) were identified, using a combination of the Illumina data analysis pipeline with custom and commercial sequence alignment software. In our configuration, clonal sequencing outperforms current diagnostic methods, providing a reduction in analysis times and in reagent costs compared with conventional sequencing.

Poster No: P050-T
Genome-wide association study for colorectal cancer in German familial cases, and implication of MAPK signalling pathways in disease susceptibility

[1]**Jesús Lascorz**, [1,2]Asta Försti, [1]Bowang Chen, [3]Pavel Vodicka, [4,5]Barbara Burwinkel, [1,2]Kari Hemminki, German HNPCC Consortium, NGFN Plus CCN Group
[1]Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany [2]Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden [3]Institute of Experimental Medicine, Academy of Sciences of the Czech Republic, Prague, Czech Republic [4]Helmholtz Group Molecular Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany [5]Molecular Oncology, Department of Gynecology and Obstetrics, University of Heidelberg, Germany

Rare high-penetrance germline mutations account for less than 5% of colorectal cancer (CRC) cases. Much of the remaining variation in genetic risk is supposed to be attributable to multiple common susceptibility loci. So far, ten common low risk variants have been identified through genome-wide association studies (GWAS). We performed a GWAS with 371 German familial CRC cases and 1263 healthy controls using the Affymetrix 6.0 Array. After application of quality control criteria, 610.664 SNPs were included in the analysis, and replication studies in four additional case-control sets (4915 cases and 5607 controls) were conducted.

Known risk loci at 8q24.21 and 11q23 were confirmed, and a previously unreported association, rs12701937 ($p = 1.1 \times 10^{-3}$; OR 1.14, 95% CI 1.05 to 1.23, dominant model in the combined cohort) was identified. The association was stronger in familial compared to unselected cases for this SNP ($p = 2.0 \times 10^{-4}$; OR 1.36, 95% CI 1.16 to 1.60, dominant model), as well as for two other unreported polymorphisms, rs6038071 and rs11014993 ($p = 2.5 \times 10^{-3}$, recessive model, and $p = 2.7 \times 10^{-4}$, dominant model). Trying to increase the power to detect association between genes and disease, several bioinformatics tools were used to search for enrichment of pathways or Gene Ontology (GO) categories among the associated polymorphisms. Three different and independent software tools (ConsensusPathDB, ToppGene Suite, and the recently published approach ALIGATOR) successfully pointed to the over-representation of genes related to the mitogen-activated protein kinase (MAPK) signalling pathways among the most strongly associated markers from the GWAS (allelic $p < 10^{-3}$). The risk of CRC increased significantly with an increasing number of risk alleles in these genes (ORper allele 1.34, 95% CI 1.11 to 1.61), and for carriers of more than four risk alleles the risk of disease was increased ~3-fold compared to carriers of ≤4 risk alleles.

Our study replicated two of the known risk loci for CRC and identified three risk polymorphism more specific for familial than for sporadic CRC cases. MAPK signalling pathway genes were enriched among the most significantly associated markers, showing a significant increase in CRC risk with increasing number of risk alleles. This and other similar approaches are clearly needed in order to reveal the large proportion of missing heritability of complex diseases such as CRC.

Poster No: P065-T
transcriptional and translational consequences of genomic re-arrangments in breast cancer

Leena Ukil, Koichiro Inaki, Yuen Fern Ho, Kartiki Desai and Edison Liu
Cancer Biology and Pharmacology 2, Genome Institute of Singapore

In addition to oncogenic transformation, various cancers are often associated with a stochastic selection of chromosomal rearrangements. Some of these events result in the expression of fusion genes that are causal in tumor progression and recurrence. The expression and the function of such fusion genes have been extensively studied in hematopoietic cancers and soft tissue tumors with the prototype rearrangement being the t(9:22)(q34:q11) translocation in chronic myeloid leukaemia (CML) that generates the BCR-ABL1 fusion gene. Recent findings of fusion gene expression in prostate cancer and lung cancer suggest that such fusion transcripts can be found in solid tumors. By contrast, gene fusions in breast cancer have not been well studied. For the comprehensive identification of such fusion genes, we have used massively parallel and high-throughput paired-end sequencing technology (PET) to analyze the transcriptome and the genome in 3 breast cancer cell lines and 5 primary breast tumors and have recorded a high number of potential rearrangements. Herein we show the transcriptional consequences of these structural genomic rearrangements in generating fusion transcripts and further analyze their functionality using a polysomal assay. We find that single

segmental tandem replication spanning several genes is a major source of two gene fusion transcripts. In addition, we find that almost 90% of the fusion transcripts associate with polysomes and are hence translated.

Poster No: P082-T
Cluster Analysis of Risk Factor Genetic Polymorphisms in Gallbladder Cancer

Kshitij Srivastava, Anvesha Srivastava, Balraj Mittal
Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, India-226014

Introduction:
Being relatively rare worldwide, gallbladder cancer (GBC) shows highest incidence and mortality rates in Northern India. Apart from gallstones being the predominant risk factor, the precise etiology of GBC is poorly understood. Cancer being a complex and polygenic disease, multiple genetic variants may interact to cause the disease or act as risk factors. Identification of risk sets of factors contributing to the disease will aide in determining individuals at risk for the disease. In this study, we applied a pathway-based multigenic approach to examine the associations of a comprehensive panel of polymorphisms in DNA-repair, apoptotic and inflammatory pathway genes with GBC risk.

Methods:
In the present population-based case-control study with 230 patients newly diagnosed with GBC and 230 unaffected controls, we applied a multigenic approach to examine the associations with GBC risk of a comprehensive panel of 16 selected polymorphisms in three pathways, DNA repair pathway [ERCC2 Asp312Asn (Ex10-16G>A; rs1799793) and Lys751Gln (Ex23+61A>C; rs13181); MSH2 (IVS1+9G>C; rs2303426) and (-118T>C; rs2303425); OGG1 Ser326Cys (Ex6-315C>G; rs1052133) and (IVS4-15C>G; rs2072668); XRCC1 Arg194Trp (Ex6-22C>T; rs1799782) and Arg399Gln (Ex10-4A>G; rs25487)], apoptotic pathway [CASP8 -652 6N ins/del (rs3834129), Asp302His (Ex13+51G>C; rs1045485) and (IVS12-19 G>A; rs3769818)] and inflammatory pathway [PTGS2 (-1195G>A; rs689466), (-765G>C; rs20417) and (Ex10+837T>C or +8473; rs5275); TLR2 –196 to –174del (TLR2 Δ22); and TLR4 Thr399Ile (Ex4+936C>T; rs4986791)]. Genotyping for all the polymorphisms was done using the PCR-restriction fragment length polymorphism method. Odds ratio (OR) and 95% confidence interval (95% CI) were calculated in multivariate logistic regression analysis for the association of individual SNPs with gallbladder cancer. A fuzzy latent classification approach, grade-of-membership analysis (GoM), was taken to identify risk sets defined by the probabilities of being affected with GBC and for genotypes found at the examined genes.

Results:
Individually, only ERCC2 Asp312Asn, MSH2 IVS1+9G>C, OGG1 Ser326Cys and OGG1 IVS4-15C>G, XRCC1 Arg399Gln, CASP8 -652 6N ins/del, PTGS2 -1195G>A, PTGS2 -765G>C and TLR2 –196 to –174del exhibited statistically significant main effects. However, we found a significant gene-dosage effect for increasing numbers of potential high-risk alleles in inflammatory pathway and all three pathways combined. For the inflammatory pathway, compared with the referent group (fewer than three adverse alleles), individuals with >2 variant alleles [odds ratio (OR)=1.93, 95% CI=1.26–2.96) had increasingly elevated risks of bladder cancer (Ptrend=0.007). Each additional adverse allele was associated with a 1.07-fold increase in risk (95% CI 1.03–1.12). For the combined analysis of DNA-repair, apoptotic and inflammatory pathway SNPs, compared with the referent group (0-10 adverse alleles), the ORs for individuals with >10 adverse alleles was 2.42 (95% CI=1.56–3.77; Ptrend<0.001). GoM analysis revealed three high intrinsic risk sets (I, II and III), having a higher density of various genotypes compared to set IV, V and VI, at low intrinsic risk for GBC.

Conclusion:
This is the first comprehensive study to use a multigenic analysis for GBC, and the data confirm the importance of taking a multigenic pathway based approach to risk assessment.

7. Evolutionary Genomics

(20 May 2010, 12.55 PM - 2.55 PM, Salle Einstein) - Chaired by Harris A Lewis / Christian Hamel

Poster No: P006-T
Combination of metabonomic, transcriptomic and genomic approaches unravels the link between various serum metabolites and an immune response pathway

Johannes Kettunen[2,3,#], Michael Inouye[1,2,#], Pasi Soininen[4,5], Kaisa Silander[3], Samuli Ripatti[3], Linda S. Kumpula[5], Pekka Jousilahti[6], Antti J. Kangas[5], Satu Männistö[6], Markku J. Savolainen[5,7], Aarno Palotie[2,3,8], Veikko Salomaa[6], Markus Perola[3], and Mika Ala-Korpela[4,5,7], Leena Peltonen[2,3,8]
[1]Immunology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia, [2]Department of Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom, [3]Institute of Molecular Medicine FIMM, University of Helsinki and Unit of Public Health Genomics, National Institute for Health and Welfare, Helsinki, Finland, [4]NMR Metabonomics Laboratory, Laboratory of Chemistry, Department of Biosciences, University of Kuopio, Kuopio, Finland, [5]Computational Medicine Research Group, Institute of Clinical Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, Oulu, Finland, [6]Unit of Chronic Disease Epidemiology and Prevention, National Institute for Health and Welfare, Helsinki, Finland, [7]Department of Internal Medicine and Biocenter Oulu, Clinical Research Center, University of Oulu, Oulu, Finland, [8]The Broad Institute of MIT and Harvard, Cambridge, MA, USA, [#]These authors contributed equally

In this study we have integrated the metabonomic, transcriptomic and genomic profiles of 518 individuals sampled from the capital region of Finland. In our previous work we identified a set of highly correlated genes associated with serum lipid levels, the Lipid-Leukocyte module (LL), by using co-expression network analysis. In this work we show that LL module displays significant correlation with over 80 metabolic measures quantified via 1H NMR spectroscopy of serum; these include omega 6 and 7 fatty acids (p = 8*10-6), apolipoprotein B (p = 1*10-9), isoleucine (p = 4*10-16), triglycerides (p = 2*10-24) and different lipoprotein subclasses, for example: very large VLDL particles, p = 1*10-24, small LDL, p = 2*10-4 and medium HDL, p = 2*10-8. We use genetic variation to construct a directed network graph and infer causality. In addition, we show that the connectivity of the module is dependent on metabolite concentrations in serum, evidence that network modules are themselves dependent on the environment. This study is a roadmap for the systematic integration of metabonomic, transcriptomic and genomic data in human population samples.

Poster No: P023-T
Role of rare DNA sequence variants in early-onset myocardial infarction in French-Canadians

[1]**Ken Sin Lo**, [1]Mélissa Beaudoin, [2,3]Manuel A. Rivas, [1,4,5]Michael Phillips, [1,5]Marie-Pierre Dubé, [1]Nathalie Laplante, [2,3]Mark J. Daly, [1,5]John D. Rioux, [1,5]Jean-Claude Tardif, [1,5]Guillaume Lettre
[1]Montreal Heart Institute, Montréal, Canada, [2]Massachusetts General Hospital, Boston, USA, [3]Broad Institute, Cambridge, USA, [4]Génome Québec and Montreal Heart Institute Pharmacogenomics Centre, Montréal, Canada, [5]Université de Montréal, Montréal, Canada

Myocardial infarction (MI), commonly known as a heart attack, is the interruption of blood supply to part of the heart due to occlusion of a coronary. The major risk factors for MI have been identified and include older age, male gender, tobacco use, high triglyceride and LDL cholesterol levels, low HDL cholesterol levels, obesity, diabetes, high blood pressure and chronic stress. A positive family history of MI is also an important risk factor, as shown by heritability studies in twins and large pedigrees. Understanding the genetic etiology of MI would have major implications in terms of prevention, prediction and drug development.

Several genome-wide association studies (GWAS) have identified loci that harbor common SNPs associated with MI risk. In many cases, the causal genes and alleles remain unknown. Because genotyping SNP arrays are specifically designed to capture common genetic variations, GWAS will not address whether rare genetic variants at the associated loci also contribute to disease risk. Here, we set to test this hypothesis by deep re-sequencing of 68 MI candidate genes (369 kb)

in French-Canadian early-onset MI patients and unaffected individuals. These candidate genes were identified from MI GWAS findings and linkage signals.

To find rare mutations that influence MI risk, we used Agilent's SureSelect solution-based capture technology combined with next-generation DNA re-sequencing on an Illumina Genome Analyzer II sequencer (short-insert paired-end 76 bp reads protocol). To minimize costs and maximize power, we re-sequenced 500 French-Canadians early-onset MI patients and 500 matched controls in pools of 50 DNA samples. Sequence reads will be aligned using BWA. DNA polymorphisms will be called using a Bayesian likelihood calculation implemented in the software Syzygy. To simplify downstream analyses and help pinpoint most susceptible loci associated with MI, we have developed a robust pipeline that annotates new DNA sequence variants, assesses their functional effect at the protein level, and performs statistical tests. Associated rare variants will be genotyped in an independent cohort of 2,000 early-onset French-Canadian MI cases and 2,000 matched controls.

Poster No: P038-T
Identification of novel sequence variants using high-throughput re-sequencing of ulcerative colitis genes in French-Canadian patients

[1]**Mélissa Beaudoin**, [1]Ken Sin Lo, [1]Philippe Goyette, [2,3]Manuel Rivas, [1]Caroline Lagacé, [1,5]Andrew M.K. Brown, [1,5]Michael S. Phillips, [4]Alain Bitton, [2,3]Mark J. Daly, [1,5]Guillaume Lettre and [1,5]John D. Rioux
[1]Montreal heart Institute, Montreal, Canada [2]Massachusetts General Hospital, Boston, USA [3]Broad Institute, Cambridge, USA [4]Royal Victoria Hospital, Montreal, Canada [5]Université de Montréal, Montreal, Canada

Ulcerative colitis (UC) is an inflammatory bowel disease (IBD) that causes chronic inflammation of the gastrointestinal tract. The mucosal inflammation is continuous and restricted to the colon and rectum. UC can occur in people of any age, but it usually starts between the ages of 15 and 30, and less frequently between 50 and 70 years of age. It affects men and women equally and segregates within families. A higher incidence of UC is seen in Caucasians and people of Jewish descent. The precise incidence of UC is unknown; however, some studies indicate as many as 10 to 15 out of 100,000 persons have the disease. The most common symptoms of UC are abdominal pain and bloody diarrhea. UC can be difficult to diagnose because its symptoms are similar to other intestinal disorders, such as Crohn's disease, another type of IBD. UC is a complex polygenic disease, influenced by both environmental (e.g. smoking, urban environment) and genetic (e.g. DNA polymorphisms within the major histocompatibility region (MHC)) factors. About 5 percent of people with ulcerative colitis develop colon cancer.

Genome-wide association studies have identified 20 loci associated with UC risk. In most cases, the causal genes and alleles are unknown. To identify causal polymorphisms and other rare DNA sequence variants, we re-sequenced the coding sequence of 55 candidate-genes from these 20 associated regions with the Illumina Genome Analyzer II DNA sequencer. We used a pooled DNA approach (4 groups of 50 UC French-Canadian cases and 3 groups of 50 French-Canadian controls) and PCR for target enrichment.

We re-sequenced 4.6 GB of DNA and obtained an average coverage of 30X per DNA sample. Quality-filtered sequence reads were aligned against the reference human genome using the Burrows-Wheeler Aligner (BWA) software (78% of the reads were on target). DNA polymorphisms were identified using a Bayesian likelihood calculation implemented in the Syzygy software. We identified 2,006 high quality DNA sequence variants, including 2 nonsense (2 novel), 233 missense (126 novel), and 6 splice site (6 novel) variants. Validation in the original re-sequenced samples, as well as replication in independent large UC cohorts will be presented.

Poster No: P053-T
Leveraging the 1000 Genomes Project for Next-Generation Microarrays

Michael A. Eberle, Jennifer L. Stone, Karine Viaud, Luana Galver, Chan Tsan and Ken Kuhn
Illumina, Inc., San Diego, CA, USA

High-throughput sequencing is expanding the catalogue of variation at an unprecedented rate, enabling a more comprehensive understanding of the underlying linkage disequilibrium (LD) patterns within and across populations. This information is required to optimally design the next generation of whole-genome genotyping (WGA) arrays that will allow the exploration of new hypotheses, including the role of intermediate and rare variation in disease. Until recently the primary source of information for developing WGA arrays was the International HapMap Project as it provided detailed information on frequencies and LD between almost four million SNPs in three distinct populations. Selecting an optimal subset of these SNPs allows the creation of arrays that interrogate over 90% of all common HapMap variants either directly or indirectly through LD with only 600-800k tagSNPs. These tools have proven successful in evaluating common variants for association in many diseases and traits, for which, before the genome-wide association era, little was understood. However, the catalog of SNPs available through HapMap represents only a small, and potentially biased, subset of the total variation in the human population. With the advent of high-throughput sequencing, the 1000 Genomes Project has greatly increased the spectrum of known variants and provides an excellent resource for content to develop the next generation of microarrays for “rich” genome-wide association studies (GWAS). These microarrays will interrogate the entire genome, including rarer content down to ~1% MAF. We have evaluated the whole-genome coverage provided by chips whose content was based on the HapMap data. Using the phased genotype data from the 1000 Genomes Project, we calculate that less than 70% of the common (>5%) variants and less than 60% of all variants seen at least twice are tagged by even the most comprehensive arrays currently available. Using a “greedy” approach to selecting tag SNPs, we show that ~2 million intelligently selected markers can effectively cover the SNPs from the 1000 Genomes Project in all three HapMap populations, excluding singletons. Additionally, because the 1000 Genomes Project is sequencing to ~4x depth, many of the available genotype calls have been imputed and/or are based on only a limited number of reads leading to potential false positives and genotype errors. We will also present the results of validation experiments using hundreds of thousands of SNPs to assess false positive rate and categorize potential errors in these imputed genotype calls.

Poster No: P068-T
Ensembl: Recent developments in human genome annotation

[1]**Amonida Zadissa**, [1]Bronwen Aken, [1]Susan Fairley, [2]Julio Fernandez Banet, [1]Felix Kokocinski, [1]Magali Ruffier, [1]Amy Tang, [1]Jan Vogel, [1]Simon White, [1]Jennifer Harrow, [1]Steve Searle.
[1]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK [2]European Bioinformatics Institute, Hinxton, Cambridge, UK.

Ensembl produces automatic gene annotation for the human genome and other vertebrate species. Recent developments in our annotation strategy allow us to provide more complete annotation for the human genome to users: For the human genome we now merge our annotation with manual annotation from HAVANA to produce the GENCODE gene set. We have also developed a pipeline to build transcript models from RNAseq data.

The GENCODE consortium aims to annotate all human protein-coding loci, non-coding loci and pseudogenes that are based on experimental. While HAVANA manual annotation produces high quality curated gene models, the process is time-consuming. Currently manual annotatin covers only 70% of the human genome and Ensembl automatic methods can provide annotation for the remaining 30%. Additionally, Ensembl's automatic annotation can be used to improve the existing manually annotated gene models using new experimental data released in the public domain. Therefore, to create an up-to-date GENCODE gene set, we have developed a pipeline to combine the the HAVANA and Ensembl gene sets to produce a merged gene set.

In addition to the GENCODE project, we still import manually-curated immunoglobulin and T-cell receptor genes from IMGT and we guarantee that all CCDS models are present in our gene sets.

RNAseq transcriptome data provides an opportunity to assess tissue-specific transcription and also to search for isoforms that are in new gene loci or that have not yet been annotated in existing gene loci. We have developed a pipeline to generate transcript models using only RNAseq data with encouraging results. Using 75bp paired-end reads, we are able to build multi-transcript gene models, some of which are in locations that currently do not have transcript annotations in Ensembl or HAVANA. Our transcript models have been entered into the RGASP project for external assessment. We hope to display the results from this pipeline on our website in the near future.

Poster No: P085-T
An atlas of combinatorial transcriptional regulation in mouse and man

Timothy Ravasi[1,2], Carlo Vittorio Cannistraci[1], Shintaro Katayama[3], Vladimir B. Bajic[1,4], Kai Tan[2], David A. Hume[17], Harukazu Suzuki[3], Yoshihide Hayashizaki[3] and Trey Ideker[2]
[1]Red Sea Integrative Systems Biology Laboratory, Division of Chemical & Life Sciences and Engineering, Computational Bioscience Research Center (CBRC), King Abdullah University for Science and Technology (KAUST), Jeddah, Kingdom of Saudi Arabia. [2] Departments of Medicine and Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA. [3] RIKEN Omics Science Center, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho Tsurumi-ku Yokohama, Kanagawa, 230-0045 Japan. [4] South African National Bioinformatics Institute, University of the Western Cape, Private Bag X17, Bellville, 7535 South Africa [5] The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin, EH259PS, UK.

Combinatorial interactions among transcription factors are critical to directing tissue-specific gene expression. To build a global atlas of these combinations, we have screened for physical interactions among the majority of human and mouse DNA-binding transcription factors (TFs) and mapped the active promoters using CAGE coupled with next generation sequencing. The complete networks contain 762 human and 877 mouse interactions. Analysis of the networks reveals that highly connected TFs are broadly expressed across tissues, and that roughly half of the measured interactions are conserved between mouse and human. The data highlight the importance of TF combinations for determining cell fate, and they lead to the identification of a SMAD3/FLI1 complex expressed during development of immunity. The availability of large TF combinatorial networks in both human and mouse will provide many opportunities to study gene regulation, tissue differentiation, and mammalian evolution.

8. Genetics of Health and Disorder

(20 May 2010, 12.55 PM - 2.55 PM, Sully 1) - Chaired by Stephen W Scherer / Veronica Van Heyningen

Poster No: P014-T
Sequence variants at CYP1A2 and AHR loci associates with coffee consumption.

Patrick Sulem [1], Daniel F. Gudbjartsson [1], Frank Geller [2], Inga Prokopenko [3,4], Bjarke Feenstra [2], Katja K.H. Aben [5,6], Barbara Franke [7,8], Martin den Heijer [6,9], Peter Kovacs [10], Michael Stumvoll [11,12], Reedik Mägi [3,4], Lisa R. Yanek [13], Lewis C. Becker [13], Heather A. Boyd [2], Simon Stacey [1], Gudmundur B. Walters [1], Adalbjorg Jonsdottir [1], Gudmar Thorleifsson [1], Hilma Holm [1], Sigurjon A. Gudjonsson [1], Thorunn Rafnar [1], Gyda Bj örnsdottir [1], Diane M. Becker [12], Mads Melbye [2], Augustine Kong [1], Anke Tönjes [11], Thorgeir Thorgeirsson [1], Unnur Thorsteinsdottir [1,14], Lambertus A. Kiemeney [5,6,15], Kari Stefansson [1,14]
[1] deCODE genetics, Sturlugata 8, 101 Reykjavik, Iceland. [2] Department of Epidemiology Research, Statens Serum Institut, 2300 Copenhagen S, Denmark. [3] Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, OX3 7LJ, UK. [4] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK. [5] Comprehensive Cancer Center IKO, 6501 BG Nijmegen, the Netherlands. Department of Genetics, [6] Department of Epidemiology, Biostatistics & HTA, [7] Department of Genetics, [8] Department of Psychiatry and [9] Department of Endocrinology, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, the Netherlands. [10] Interdisciplinary Centre for Clinical Research, [11] Department of Medicine, and [12] LIFE Study Center, University of Leipzig, 04103 Leipzig, Germany. [13] The Johns Hopkins GeneSTAR Research Program, Division of General Internal Medicine, The Johns Hopkins Medical Institutions, Baltimore, MD21287, Maryland. [14] University of Iceland, Faculty of Medicine, 101 Reykjavik, Iceland. [15] Department of Urology, Radboud University Nijmegen Medical Centre.

Coffee is the most commonly used stimulant and the amount of coffee consumed has been shown to be heritable. There has been substantial debate on the effect of coffee consumption on health, e.g. whether it increases risk of hypertension and protects against Alzheimer's disease. We performed a genome-wide association scan of 6,611 coffee drinkers of

European descent (From Iceland, The Netherlands , Germany and the US) followed by a replication in an additional set of 4,050 individuals (from Iceland and Denmark) and identified two genome wide significant associations (P< 1.10-7). The first association is with a SNP located in a cluster of the dioxin response element between CYP1A1 and CYP1A2 , the main caffeine metabolizing enzyme . The second association with coffee consumption was detected with a SNP directly upstream of the Aryl Hydrocarbon receptor (AHR) locus, a gene that induces CYP1A1 CYP1A2 through their dioxin response element . For both SNPs, the observed effect size was around 0.2 additional cups of coffee per day per allele. We postulate that the effect on coffee consumption operates through an effect on caffeine metabolism, such that individuals with a higher caffeine clearance are able to drink more coffee than those with lower caffeine clearance. Given that caffeine is used as a probe to assess the pharmacokinetics of several therapeutic drugs these variants are candidates for having an effect on drug metabolism.

Poster No: P029-T
The RetChip v1.0 resequencing array - application of high-throughput analysis in diagnostic testing of hereditary retinal degenerations

[1*]**Heidi L. Schulz** , [1*]Lars Fritsche, [1]Kerstin Meier, [1]Britta S. Fiebig, [1]Bernhard H.F. Weber
[1] *Institute of Human Genetics, University of Regensburg, Germany*
**Authors contributed equally to the work*

Retinal degeneration (RD), the term generally used to describe a common phenotype found in more than two dozen different disease phenotypes, affects approximately one in 2.000 individuals worldwide. These inherited dystrophies, characterized by malfunction or loss of cone and/or rod photoreceptors, have unparalleled high genetic and phenotypic heterogeneity. Due to the significant clinical overlap between the disorders, the fact that more than 150 genetic loci harbor causative mutations, the large gene sizes and lack of mutational hotspots, it has been difficult to comprehensively survey the disease-causing mutations in clinical practice. Current diagnostic approaches, such as conventional Sanger sequencing or use of approaches that test for known mutations, regularly fail to detect the complete mutational spectrum. To address this limitation, while simultaneously reducing cost, time and personnel required for routine DNA testing, we have designed and implemented a fast and reliable genotyping platform, known as RetChip v.1.0. This array contains more than two million probes matching the coding and flanking intronic sequences of 72 RD genes as well as additional 702 common variants and 587 known indel mutations. To streamline the set-up process, a multiplex PCR approach was implemented to amplify the 1.147 amplicons.

Chip performance, accuracy, and reproducibility were assessed by analysing 14 controls and three RD subjects. The combined use of two analysis softwares (GSeq and SeqC) led to consistent call rates higher than 99.97%. All substitutions identified by the chip analysis were confirmed by dideoxy sequencing. To further validate the method, seven patients with known mutations in the ABCA4 gene were genotyped and all 44 known variations they carried were successfully identified. Currently, the RetChip is being used in routine molecular diagnosis of Stargardt disease and cone-rod dystrophy patients. The advantage of this platform over current technologies is illustrated by the fact that in 50 Stargardt disease patients at least one disease-associated allele was detected in 68% of cases, an increase of 14% over results obtained with a mutation-specific microarray. Even though 547 mutations have been published in the three genes investigated (ABCA4, CNGB3, ELOVL4), we found 18 novel variants of probable pathogenicity in the 50 patients. The significance of the RetChip platform will be even higher for less well-characterized RD phenotypes.

The RetChip is a powerful tool that will significantly contribute to clinical practice. It will not only facilitate differential diagnosis but also enable predictive testing, identification of modifier variants and determination of genotype-phenotype correlations. Additionally, it will provide clinicians with information essential for prognosis and selection of therapeutic strategies.

Poster No: P059-T
tri e tocotrienol supplementation affects gene expression of peripheral blood mononuclear cells from different aged individuals

Wan Zurinah Wan Ngah[1], Eng Chee Heng[1], Noor Aini Abdul Hamid[1], Noriah Ramli[1]and Zalina Hamid[2]
UKM Medical Molecular Biology Institute/Department of Biochemistry, Faculty of Medicine, Universiti Kebangsaan Malaysia[1], Sime Darby Bioganic, Kuala Langat, Malaysia[2]

The different isomers of vitamin E have been reported to be protective against ageing, cardiovascular disease and cancer. However, human studies involving supplement of vitamin E are limited. In this study, the effect of Tri E tocotrienol supplementation was determined. Fourteen healthy female volunteers were recruited and divided into two age groups: 6 young subjects aged 32+2 (3 received placebo & 3 Tri E) and 8 older subjects aged 52+2 (4 received placebo & 4 Tri E) at a dose of 150mg tocotrienol rich fraction. Fasting blood (15mL) was taken before and after 3 and 6 months supplementation. Total RNA was extracted from PBMC using Qiagen RNeasy Mini Kit. Analysis using GeneSpring 10 GX suggested 134 genes and 257 genes differentially expressed in younger group after 3 and 6 month Tri E supplementation. About 94 genes were down-regulated and 40 genes up-regulated at 3 month supplementation. The number of genes affected increased to 182 genes up-regulated and 73 down-regulated at 6 month supplementation. The most highly down-expressed genes in the younger group were involved in the cell cycle arrest (GADD45A, GADD45B, DDIT4, PCNA), apoptosis (PLEKHF1, BAK1), inflammation (ICAM1, ICAM2), angiogenesis (S100A4, NR4A2, NR4A3) and cholesterol synthesis (DHCR7, SC5DL). In the older group, Tri E supplementation caused changes in 21 genes and 591 genes between 3 and 6 months. At 3 month supplementation, 8 genes were up-regulated and 13 genes down-regulated in contrast to 507 genes down-regulated and 84 genes up-regulated at 6 month supplementation. Cytokines (IL-1A, IL-6, SOCS1, SOCS3), cell cycle arrest (DDIT4), apoptosis (PLEKHF1, CTSD, IER3, BAK1), inflammation (ICAM1, ICAM3), angiogenesis (S100A4), and cholesterol biosynthesis (DHC7) were down-regulated while histone binding protein (H2AB), antioxidant enzyme (GSTM1L), and G protein regulator (RGS18) were up-regulated in the older group. Genes such as ICAM1, DDIT4, PLEKHF1, BAK1, S100A4, and DHCR7 were down-regulated in both groups after Tri E supplementation. Up-regulated and down-regulated were analyzed using Adriadne Pathway Studio 7 shown enrichment of pathway related to NF-kB pathway. In conclusion, differences in gene expression were more profound in the older group compared to younger group after 6 month supplementation.

Poster No: P074-T
Personalized Nutrition: Will Livestock Lead the Way?

Harris A. Lewin,
Department of Animal Sciences and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, USA

The sequencing, annotation and functional characterization of the genomes of cattle and pigs have been rapidly translated into new products and services for the livestock industry. The animal breeding industry adopted marker-assisted breeding in the late 1990s and is now implementing genomic selection for complex traits, such as milk production and carcass composition. In addition, the mapping and positional cloning of genes responsible for a large number of monogenic and polygenic traits has resulted in the creation of new models for human diseases and has informed our understanding of the genetics of complex traits. One of the most important areas of interface between human and livestock genomics is in the area of nutrition. In human and animal health, there is a dynamic interplay between host genotype and diet. However, for many diseases, the specific genes, pathways and networks that are influenced by diet are not well understood. In cattle and other livestock species it has been common practice for the past 50 years to develop special diet formulations based on breed and/or environmental conditions. Recently, we have discovered what may be the first example of a diet-regulated QTL (DGAT1) in dairy cattle that influences the development of metabolic disorders associated with the onset of lactation. Our data suggest that controlling energy intake during the prepartal period in DGAT1 genotyped animals can be used to optimize milk production performance while minimizing the incidence of postpartal metabolic disorders, such as fatty liver disease and ketosis. For human diseases, understanding of diet-genotype interactions may be critical to promoting wellness in individuals genetically predisposed to diabetes, obesity, heart disease, gastrointestinal and immunological disorders and cancer. By implementing dietary restrictions in genetically-defined individuals, it may be possible to reduce the incidence and severity of such diseases. Thus, “personalized nutrition” could take on a significant role as a first line of defense against the “diseases of affluence” currently plaguing Western and developing nations. Research using non-inbred cattle and pigs as models may contribute critical knowledge that will pave the way toward the implementation of genetics-driven nutritional management of human diseases.

Poster No: P088-T
The mechanisms of action of Sirtuin 1

[1]**Sascha Sauer**, [1]Chung-Ting Han, [1]Susanne Holzhauser, [2]Thomas Manke, [2]Stefan Haas and [2]Martin Vingron
Max-Planck-Institute for Molecular Genetics, Ihnestrasse 63-73, D-14195 Berlin, Germany, [1]Department Otto-Warburg-Laboratory, [1]Department of Computational Molecular Biology

Sirtuin 1 (SIRT1), a NAD+-dependant deacetylase, has been implicated in the prevention of aging processes and metabolic diseases such as type 2 diabetes. SIRT1 modulates cellular gene expression by deacetylating a number of regulatory proteins such as nuclear receptors, transcriptional cofactors and histones. Resveratrol, a natural product that was found in red wine, was shown to activate SIRT1 directly. Synthetic SIRT1 activators have been described as promising new antidiabetic therapeutics. However, it was shown that the previously reported enzymatic activation of SIRT1 by small molecules were simply assay artefacts generated by fluorescently labelled substrate peptides. We discovered novel mechanisms of SIRT1 “activation” by resveratrol. Our results indicate specific effects of resveratrol on several histone modifications (namely H1K26, H3K9, H4K16) and on global gene expression. We used chromatin immunoprecipitation followed by second generation sequencing (ChIP-Seq) to compare patterns of histone modification between untreated and resveratrol-treated murine skeletal muscle cells. In general, resveratrol treatment results in global deacetylation of histones. But surprisingly, resveratrol induces dense acetylation of H1K26 on the exons of important metabolic sensors, in particular SIRT1, PPARgamma, and LXRalpha. Moreover, we found specific acetylation of H3K9 at transcriptional start sites (TSS) of a number of genes, which has a striking correlation with SIRT1 binding. After integrating the chromatin modification data with gene expression data, we observed significant changes in gene expression of various genes involved in glycolysis, gluconeogenesis and fat oxidation. Amongst others, the H1K26 acetylation pattern was associated with upregulated SIRT1, PPARgamma, and LXRalpha gene expression. Moreover, we found up regulations of uncoupling proteins UCP1 and UCP3 that are involved in thermogenesis. Finally we observed up regulation of the regulatory kinase PDK4 that inhibits the formation of the pyruvat dehydrogenase complex and prevents the entry of pyruvate into the tricarboxylic acid (TCA) cycle. These data to date lead to the hypothesis that resveratrol treatment results in energy dissipation in muscle cells.

For future SIRT1 small molecule screening projects we developed a novel, unambiguous mass spectrometry-based assay including unlabelled peptide substrates.

Using a structurally diverse library of ~10,000 natural products and their analogues we identified a number of new molecules that inhibit SIRT1 activity with low IC50 values. We have furthermore characterised the physiological function of these molecules in different cell culture models.

In general, our data demonstrate how dietary compounds and potent synthetic molecules are likely to influence the expression of metabolically active genes through histone modifications. Novel SIRT1 inhibitors may be used as reference compounds to study SIRT1 biology. Our results shed light on the mechanism of SIRT 1 activation and provide new concepts for the development of small molecules in preventing and treating type 2 diabetes or in slowing aging processes.

Poster No: P101-T
OPA8 is mutated in dominant families with mild optic atrophy and severe deafness

[1]**Laëtitia Buret**, [1]Guy Rebillard, [1]Benjamin Delprat, [1]Marie Péquignot, [1]Karim Chekroud, [2]Karen Cornille, [3]Ghislaine Hamard, [4]Odile Boespflug-Tanguy ,[1]Christian Hamel, [1]Guy Lenaers, [1]Cécile Delettre.
[1]Institute of Neurosciences of Montpellier, St Eloi hospital, Montpellier, France. [2] UMR866 Cellular Differentiation and Growth, SupAgro-INRA, Montpellier, France. [3] Technical service "homologous recombination", Cochin Institute, Paris, France. [4] CHRU Clermont-Ferrand, France.

Dominant optic atrophy (DOA) is the most common form of inherited optic neuropathy. DOA affects retinal ganglion cells (RGCs) and the axons forming the optic nerve, leading to progressive visual loss. Today, there is no treatment able to stop the degenerative process. Unless several loci are implied only two genes have been identified: OPA1 the major gene of DOA and OPA3 associating a DOA with a cataract.

Using a candidate gene strategy on a cohort of OPA1-OPA3 negative patients, we found that a new gene OPA8 is mutated in 3 dominant families with mild optic neuropathy and severe deafness. OPA8 protein is preponderant in sensory neurons: it is expressed in the RGCs and the bipolar cells of retina and in the spiral ganglion neuron in the cochlea. OPA8 plays a role in the control of apoptosis. Consistent with this function we have shown that OPA8 mutated cells are more sensitive to apoptotic stimuli. Moreover, overexpression of the protein is protecting against apoptosis, while that of the mutated protein is not.

In order to understand the physiopathology of optic and auditory neuropathy linked to OPA8 mutations, we have generated a mutant mouse reproducing a pathogenic mutation. We have tested the visual acuity by electroretinogram (ERG) and visual evoked potential (VEP) and the auditory acuity by auditory evoked potential (AEP) and endocochlear potential (EP). The mouse presents at 2 months age an auditory loss between 10 and 15 dB according to the frequencies, in comparison with wild type mice. This auditory loss does not increase during the life of the animal. We observed on ERG a decreasing of b wave in mutant mice from 8 months. We will perform electronic microscopy to identify damaged cells in retina and cochlea and we will check for cells degeneration by immunofluorescence. Visual and auditory function will be studied on mice until 24 months.

Identification of the new gene OPA8 allowed us to build a unique mouse model presenting similar phenotypes to those observed on patients mutated in OPA8. Insights provide by the study of this mouse will help us to understand the physiopathological mechanism of OPA8 mutations.

9. Genomic Medicine and Pharmacogenomics

(20 May 2010, 12.55 PM - 2.55 PM, Sully 2) - Chaired by Lance Miller / Marc Delpech

Poster No: P032-T
From «Bedside to Bench and Back » : Innovative Diagnostic and Therapeutic Technologies for this New Decade.

[1]**B. Reversade**, [2]B. Merriman, [3]A. Masri, [4]H. Kayserili, [2]H. Lee, [1]M. Shboul, [1]J. Tian, [1]C. Bonnard, [1]N. Escande, [1]S. Chng, [2]S. Nelson, [5]H. Hamamy.
*[1]Institute of Medical Biology, A*STAR, Singapore, Singapore. [2] Dpt. of Human Genetics, University of California, Los Angeles, Los Angeles, USA. [3] Dpt. of Paediatrics, Jordan University Hospital, Amman, Jordan. [4] Medical Genetics Department, Istanbul Medical School and University, Istanbul, Turkey. [5]Dpt. of Genetic Medicine & Development, Geneva University Hospital, Geneva, Switzerland.*

A novel high-throughput strategy comprising SNP-chip based mapping, array-based genomic pull-down and massively parallel re-sequencing was used to efficiently screen for causative mutations in ten rare syndromes. In each case, the responsible gene was identified. Disease pathogenesis was then examined by functional studies involving patient's induced-Pluripotent Stem (iPS) cells and animal modeling. Genetic therapy via homologous recombination in iPS cells by means of Zinc Finger Nuclease (ZFN)-mediated gene repair, is underway to validate the much anticipated "bedside to bench and back" paradigm.

Poster No: P047-T
Translating functional genomics into the clinics: Characterization and isolation of differentiating mesenchymal stem cells in live cell chips

Christian Maercker[1,2], Karen Bieback[3], Dirk Breitkreutz[1,2], Michael Angstmann[1,2]
[1]Mannheim University of Applied Sciences, Mannheim, Germany; [2]German Cancer Research Center, Heidelberg, Germany; [3]University of Heidelberg, University Hospital Mannheim.

Mesenchymal stem cells (MSCs) are a promising tool for regenerative medicine and tissue engineering due to the ease of their isolation and differentiation potential. They also can be easily transfected with viral vectors making them an interesting vehicle in gene-therapy approaches. However, MSC isolates are mixed populations containing diverse cellular precursors in different differentiation stages with diverse differentiation end points. Moreover, no specific biomarkers have been identified so far. To improve this situation, we have developed a method to monitor MSC differentiation by measuring cell adhesion to a substratum. In live cell chips in a multiwell format (ibidi and xCELLigence/Roche), cell adhesion is monitored by measuring the alternating-current resistance (impedance) on gold electrodes. Our experiments showed that the impedance profiles are not only specific for MSCs vs. other cell types, but also for developmental stages of stem cells in vitro differentiating into adipogenic or osteogenic precursors. We took advantage of these results by testing the influence of different components of the extracellular matrix (ECM) on cellular behavior. As expected, the adhesion profiles changed on different coatings. However, more interestingly, certain cell differentiation pathways obviously are promoted by certain matrix molecules. A five day measurement of cell adhesion in cell differentiation media showed that on type I and type IV collagens, the impedance values of MSCs in osteogenic induction medium are very high compared to cells in adipogenic medium. On fibronectin, we detected a high shoulder in the profile with adipogenic cells between 8 and 45 h after start of the experiment, with values even higher than in osteogenic medium. This result also could be confirmed by oil-red staining for microscopy after 5 days, showing a high rate of adipogenic cells. Therefore, this high-content screening approach, beside other possible applications in pharma research, is a promising tool to differentiate and enrich MSCs for molecular medicine.

Poster No: P062-T
Genome-wide expression profiling at maternal-fetal interface in case of recurrent miscarriage (RM)

[1]**Kärt Tomberg**, [1,2]Kristiina Rull, [3]Sulev Kõks, [1]Jaana Männik, [1]Liis Uusküla, [4]Meeli Sirotkina, [5]Märt Möls, [6]Signe Värv, [7]Robert K. Campbell, [1]Maris Laan.
[1]Human Molecular Genetics group, Institute of Molecular and Cell Biology, University of Tartu, Estonia, [2]Department of Obstetrics and Gynecology, University of Tartu, Estonia, [3]Department of Physiology, University of Tartu, Estonia, [4]Department of Pathology, Tartu University Hospital, Estonia, [5]Institute of Mathematical Statistics, University of Tartu, Estonia, [6]Department of Cell Biology, Institute of olecular and Cell Biology, University of Tartu, Estonia, [7]Department of Molecular Biology, EMD Serono Research Institute, MA, USA

Maternal-fetal interface plays a crucial role to ensure a successful pregnancy. RM (≥3 consecutive pregnancy losses) occurring in 1-3% of fertile couples has a heterogeneous background with contribution from both genetic and environmental factors. As several physiological processes are affected in the pathogenesis of RM, it serves as a good model to study the processes at the maternal-fetal interface.

We aimed to map differentially expressed genes and pathways affected in case of RM. Affymetrix® GeneChip® HG-U133 Plus 2.0 Array was applied to placental tissue from 4 RM cases (mean gestational age 63 days) and 6 elective abortions as controls (mean gestational age 62.8 days). Between the two groups 30 transcripts representing 27 genes showed differential expression. 10 genes with the highest fold-change were chosen for validation with Taqman® RT-qPCR (13 RM cases, 23 controls) and significant differential expression was replicated in 4 of the genes. RM patients exhibited an upregulation of tumor necrosis factor family member TRAIL ($p=0.005$) and downregulation of transcriptional repressor SNAI2 ($p=0.043$), both involved in apoptotic pathways. S100A8, encoding for inflammatory marker calprotectin, showed a significantly ($p=0.005$) higher expression in case of RM while ASMTL ($p=0.055$), gene of unknown function, was downregulated. Combinatory effect of TRAIL, S100A8 and ASMTL provided a highly sensitive test distinguishing RM cases from controls (ROC analysis, area under curve=0.967). Immunohistochemical staining detected TRAIL, SNAI2 and ASMTL mainly in trophoblastic cells and S100A8 in myeloid cells of maternal blood at maternal-fetal interface of first trimester placenta.

In conclusion, genome wide expression profiling distinguished four differentially expressed genes in RM placentas: TRAIL, S100A8, SNAI2 and ASMTL. Although the detected gene expression alterations related to various pathways could be primary (causing) or secondary (consequence) events associated with the process of RM, the joint contribution of identified markers may provide a highly predictive test for detection of early pregnancy complications.

Poster No: P079-T

Genetic polymorphisms of SLC22A3-LPAL2-LPA gene cluster has direct and indirect effects on coronary atherosclerosis and coronary heart disease

[1]Motoji Sawabe, [2]Noriko Tanaka, [1]Tomio Arai, [3]Makiko Naka Mieno, [4]Shinobu Ikeda, [4]Masaaki Muramatsu
[1]Department of Pathology, Tokyo Metropolitan Geriatric Hospital, 35-2 Itabashi, Tokyo 173-0015, Japan, [2]Department of Biostatistics, Harvard School of Public Health, Boston, USA, [3]Department of Medical Informatics, Center for Information, Jichi Medical University, Shimotsuke, Japan, [4]Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

[Objective] High serum lipoprotein(a) [Lp(a)] level is a genetic risk factor for cardiovascular events and LPA gene encodes apolipoprotein(a), a characteristic component of Lp(a). Not surprisingly, a recent genome-wide haplotype analysis identified a SLC22A3-LPAL2-LPA gene cluster as a susceptibility region of coronary heart disease (CHD) in European population [Nat Genet, 2009]. Since the physiological role and pathological role in atherogenesis of Lp(a) is unknown, we tried to analyze a complex relationship among serum Lp(a) level, coronary atherosclerosis, and myocardial infarction using a large number of autopsy cases [Heart, 2009]. We here analyze the effect of genetic polymorphisms of SLC22A3-LPAL2-LPA gene cluster on CHD to clarify the contribution of these two intermediate phenotypes [serum Lp(a) level and coronary atherosclerosis] of CHD using the same cohort. [Methods] The subjects comprised 1,150 consecutive autopsy cases with an average age of 80 years. The coronary stenotic index (CSI) was studied using autopsy hearts. Unfrozen serum Lp(a) level was measured in 865 cases. Two genetic polymorphisms were analyzed; rs2048327 in SLC22A3 (C/T) and rs10755578 (C/G) in LPA. We performed a path analysis with adjustment of gender, age, hypertension, dyslipidemia, and smoking and drinking habits. [Results] Frequencies of rs2048327-T and rs10755578-C were 0.469 and 0.534, respectively and their linkage disequilibrium was low (D' = 0.331). The frequencies of four haplotypes were as follows: 0.320 in T-C haplotype, 0.153 in T-G, 0.201 in C-C, and 0.326 in C-G. The path analysis confirmed significant effects of Lp(a) to CSI, Lp(a) to CHD, and CSI to CHD in each haplotype. Serum Lp(a) level was significantly higher in T-C haplotype (path coefficient = 0.073, t = 2.13, p = 0.033) and lower in C-G haplotype (-0.091, t = 2.69, p = 0.0073). T-G haplotype had a direct effect on CSI (0.087, t = 2.59, p = 0.0098) and CHD (0.079, t = 2.52, p = 0.012). [Conclusions] T-C haplotype works as a risk factor for coronary sclerosis and CHD through high serum Lp(a) level, and T-G haplotype had a direct positive effect on coronary sclerosis and CHD. Thus the genetic interaction between these two polymorphisms on CHD was observed.

Poster No: P091-T

Pharmacogenetics of CYP3A4 in Bangladeshi population including TB patients

[1]**Abdullah Al Maruf**, [3]Masum Ahmed, [2]M.A.K. Azad, [1]Md. Ashik Ullah, [1]Abul Hasnat
[1]Department of Clinical Pharmacy & Pharmacy, Faculty of Pharmacy, University of Dhaka, Dhaka-1000, Bangladesh.
[2]Department of Pharmaceutical Technology, Faculty of Pharmacy, University of Dhaka, Dhaka-1000, Bangladesh.
[3]National Institute of Diseases of Chest and Hospital, TB unit, Mohakhali, Dhaka-1212, Bangladesh

The purpose of this study is to investigate the genotypes and phenotypes of CYP3A4 genes in Bangladeshi population and to compare these in healthy individuals with TB patients. Cytochrome P450 3A4 (CYP3A4) is the major cytochrome involved in metabolizing of >60% of all drugs used in humans. CYP3A4 exhibits a remarkable interindividual activity variation as high as 20-fold. Detection of CYP3A4 variant alleles and knowledge about their allelic frequencies in Bangladeshi healthy subjects and TB patients are important to lead to individualized drug dosing and improved therapeutics. Genotyping was done using the extracted genomic DNA from 200 unrelated healthy Bangladeshi subjects as well as from 90 TB patients followed by amplification of target alleles by PCR. Amplified alleles was digested by restriction enzymes (MboII, XcmI, BsmAI, ClaI, HinfI, HpyCH4III & HpaII) followed by gel electrophoresis & sequencing to, identify the targeted alleles namely CYP3A4*1B, CYP3A4*2, CYP3A4*4, CY3A4*5, CYP3A4*6, CYP3A4*10 & CYP3A4*18. In both groups (healthy individuals and TB patients), no mutations were detected for CYP3A4*2, CYP3A4*4, CYP3A4*5, CYP3A4*6, CYP3A4*10 & CYP3A4*18 alleles. Two samples with CYP3A4*1B allele were found to be heterozygous (1%) (n=200) in healthy subjects, where as only one sample with CYP3A4*1B allele was found to be heterozygous (1.11%) (n=90) in TB patients. This is the first investigation establishing CYP3A4 genotypes and demonstrating the absence of common CYP3A4 genotypes in Bangladeshi population. For phenotyping, morning spot

urine samples were collected from 93 healthy subjects. Cortisol and 6β-hydroxy-cortisol was extracted and quantified by HPLC. For cortisol, the maximum and minimum concentration found was 819.26 ng/ml and 19.64 ng/ml respectively with an average of 123.75 ± 118.37 (SD). On the other hand, the maximum and minimum concentration of 6β-hydroxy-cortisol was 466.65 ng/ml and 8.12 ng/ml respectively with an average of 101.57 ± 91.23 (SD). The ratios of 6β-hydroxy-cortisol and cortisol ranged from 0.24 to 7.49 with an average of 1.84 ± 1.66 (SD). The within-day coefficient of variation (C.V.) was 4.5 % (n=93), while the between-day C.V. was found to be 8.7% (n=93). Correlation between phenotyping and genotyping results and statistical analysis will be done after phenotyping rest of the healthy subjects and TB patients. Clinical relevance of these genetic variants is under investigation. This is the first investigation establishing CYP3A4 phenotypes and demonstrating the absence of common CYP3A4 genotypes in Bangladeshi population.

10. Genetics and Genomics in Developing Countries

(20 May 2010, 12.55 PM - 2.55 PM, Sully 3) - Chaired by Dhavendra Kumar / Michele Ramsay

Poster: P009-T
MtDNA and Y-chromosomal variation in populations of Sakha (Yakutia)

[1]Sardana A Fedorova, [2]Elza K.Khusnutdinova
[1]*Laboratory of molecular genetics, Yakut Research Center of Complex Medical Problems, Yakutsk, Russia, 677010*
[2]*Department of Genomics, Institute of Biochemistry and Genetics, Ufa, Russia, 450054*

We have characterized, at high phylogenetic resolution, mtDNA variation (n=694) and Y chromosome haplotypes diversity (n=318) in populations of Sakha (Yakutia) - the Autonomous Republic situated in northeastern part of the Russian Federation that comprises 1/5 part of Russia's total land area. The results were analyzed in a broader context of the Eurasian mtDNA and Y-chromosomal variability. Extended analysis confirms that Yakutia was colonized from the regions west and eastward of Lake Baikal with minor gene flow from Lower Amur/Southern Okhotsk region and/or Kamchatka. The genetic portraits of studied ethnic groups (Sakha or Yakuts, Evenks, Evens, Yukaghirs, Dolgans) were obtained and scenarios of ethnogenesis suggested by historians and archaeologists were compared with genetic reconstructions. We considered our results in connection with some epidemiologic and molecular genetic researches of hereditary diseases characterized by a high prevalence in the region such as spinocerebellar ataxia type 1, myotonic dystrophy, 3-M syndrome.

Poster No: P026-T
Genetic diseases and founder mutations in the Tunisian Population: Implications for diagnosis in North Africa and Middle East

[1]**Sonia Abdelhak**, [1]Lilia Romdhane , [1]Rym Kefi-Ben Atig and [1]Safa Romdhane
[1] *Research Unit on Molecular Investigation of Genetic Orphan Diseases. Institut Pasteur de Tunis.13, Place Pasteur, 1002, Tunis. Tunisia*

Tunisia is a North African country of 10 million inhabitants at the cross road between Europe and Africa. Throughout its history, it has been the seat of the invasion and immigration of different ethnic groups. As its neighbouring and Middle Eastern countries, its population shows a high prevalence of consanguinity and endogamy, thus leading to emergence of genetic disorders at higher rates. We report, here on the spectrum of genetic diseases in Tunisia and on founder mutations. The review of the literature, including other available information (gray literature) showed that 346 genetic disorders for which cases have been identified in the Tunisian population. Among these, 62,9 % are autosomal recessive, 23 % autosomal dominant, 5,4 % X-linked and the remaining are of Y-linked, mitochondrial and unknown mode of transmission. Fifty percent of the reported conditions in this study are caused by at least one mutation. For autosomal recessive diseases, most of the mutations were identified at homozygous state among the affected individuals. Part of the mutations was the result of a founder effect. Two classes of founder mutations have been identified in the Tunisian

population. The first includes founder mutations that are reported so far only among Tunisian patients, they are mainly the result of the high rate of consanguinity and endogamy. The second founder mutations are shared with other populations originating mainly from other North African or Middle Eastern countries and in certain cases from both shores of Mediterranean. These mutations have captured historical events in the region and are particularly useful for the development of easy and cost effective tools for molecular diagnosis.

Poster No: P041-T
High-altitude adaptation of highland natives, human or yak: NOS3 gene as an example

Qadar Pasha MA[1] and Ghulam Mohammad[2]
[1]Functional Genomics Unit, Institute of Genomics and Integrative Biology, Delhi, India [2]Department of Medicine, Sonam Norboo Memorial Hospital, Leh, Ladakh, J&K, India.

High-altitude (HA), time immemorial, has attracted humans for its beauty, adventure and, of course, natural laboratory setup owing to its hypobaric hypoxia environment. The attenuated hypoxic pulmonary vasoconstriction (HPV) and high nitric oxide (NO) level of highland-natives suggest the evolutionary (genetic) adaptation as they are continuously subjected to a stressful environment. NO is the most implicated molecule as it is the endogenous vasodilator required for maintaining vascular homeostasis and counteracting the effect of HPV. Hence, higher NO level in highland-natives certainly provides endurance for living and also indicates the selection pressure of those genes involved in vascular homeostasis. The native Himalayan Yak (Bos grunniens) and Tibeto-Burman Ladakhi population both being successfully adapted to the chronic hypoxia of HA (>3000m) served as cogent models to study the molecular basis of adaptation and endurance.

Out of the >4000 SNPs analyzed in a genome-wide screen of native humans, we preferred to restrict to Endothelial nitric oxide synthase gene (NOS3), which codes for the enzyme eNOS that produces NO. Screening of NOS3 in yak, surprisingly, showed the complete homozygous selection of –922A and –786T promoter alleles when compared with highland-natives. The same alleles were over-represented significantly in highland-natives compared to sealand-natives ($p=0.000$). Genotype and allelic distribution of G894T and 4b/4a polymorphisms was significantly different in the two groups ($p=0.00$). Strong selection pressure was also demonstrated by highest r2 value ($r2=1.00$) when compared with rest of the world's major populations. The circulating NO levels were significantly higher at HA compared to sealand ($p<0.0001$).

The NOS3 variants, its haplotypes and higher NO levels in highland natives like human and yak strongly favor the involvement of this gene in evolutionary adaptation.

Poster No: P056-T
Genomic scale labeling of persistent genes

Jian-Dong Huang[1], Yuanying Miao[1], Rory M. Watt[2], Antoine Danchin[3], David K. Smith[1], Hongzhe Sun[4], Zhongjun Zhou[1],Junwen Wang[1]
[1]Department of Biochemistry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Pok Fu Lam, Hong Kong SAR, People's Republic of China [2]Oral Biosciences, Faculty of Dentistry, University of Hong Kong, Pok Fu Lam, Hong Kong SAR, People's Republic of China [3]Amabiotics, Genoscope/CEA, 2, rue Gaston Cremieux, Evry Cedex, 91057, France

[4]Department of Chemistry, University of Hong Kong, Pok Fu Lam, Hong Kong SAR, People's Republic of China

Systematic characterization of the locations of proteins is important for understanding genome and proteome functions. This may be achieved by visualizing fluorescent protein (e.g. EGFP)-tagged proteins within live cells using fluorescence microscopy. Recombinnering technology allows the efficient integration of transformed linear dsDNA, containing short homologous regions (35-55 bp), into bacterial chromosome. To produce EGFP fusion proteins under the control of the target gene's endogenous promoter, the PCR product containing the EGFP coding region, a selectable marker and 50 bp of sequence homologous to the chromosome regions immediately upstream and downstream of the target gene was

electroporated into an E. coli that contained recombineering functionality. Correctly EGFP-tagged strains were identified using selective medium and were confirmed by PCR. The localization patterns of the target gene products within immobilized viable cells were determined by fluorescence microscopy. We have tried to label all 611 persistent genes of E. coli. 253 of them have been tagged successfully among which 179 have been located successfully. Successfully located proteins have different cellular localization: the majority localized diffusely throughout the cell; some localized at either or both poles or at the center of the cell, some were membrane localized. Several proteins exhibited localization patterns that were different from previously results. The fluorescent intensities of the EGFP-tagged proteins varied. This might be caused by differing expression levels of the target proteins. Some gene products showed dynamic localization patterns at the different stages of the growth phase. The process of tagging genes with EGFP directly on the E. coli chromosome has revealed known localization patterns and new information. As the fluorescence intensity of EGFP-tagged genes varied from gene to gene, recombineering technology might allow the study gene expression levels. Our findings also show that the growth phase of strain should be considered when determining the localization patterns of proteins.

11. Cancer Genomics II

(21 May 2010, 10.30 AM - 12.00 NN, Salle Pasteur) - Chaired by Wei Chia-Lin / Albertina de Sario

Poster No: P004-F
Genomic and functional analysis of the tumor cell of origin in HCC.

Maia Chanrion[1], Eric Sawey[1], Scott Powers[1]
[1]*Cold Spring Harbor Laboratory*

There is substantial heterogeneity of the gene expression profiles and clinical outcome of human hepatocellular carcinomas (HCCs). While some of this is likely caused by the diversity of acquired genetic alterations, it is also possible that there could be diverse cells of origin for HCCs. Previously, we determined that hepatoblasts from day 14 embryos could be malignantly transformed by a combination of overexpression of MYC and loss of p53, but that more differentiated hepatoblasts from day 18 embryos were resistant to such malignant transformation (1). We used genomic analysis to investigate underlying mechanisms. By comparing expression profiles of D14 and D18 mouse liver hepatoblasts, we found that the most significantly altered signaling pathways involve the nuclear receptor RXRα, upregulated in D18. By silencing the expression of the nuclear hormone receptor RXRα in D18 hepatoblasts using shRNA, we found that we can induce D18 hepatoblasts to become malignantly transformed by MYC and loss of p53 after subcutaneous and in situ injections. Thus, day 14 hepatoblasts require only two oncogenetic “hits” while day 18 hepatoblasts require three hits. We validated the relevance of RXRα in human HCC as it relatively underexpressed in poor prognosis associated HCC patients (2). We also showed that the 9-cis-retinoic acid, an RXRα agonist, slowed exclusively the growth of differentiated HCC cell lines expressing RXRα. Our results pinpoint RXRα as an oncogenic determinant of the cell of origin and highlight its potential as a predictive marker of the response to the 9-cis-retinoic acid.

(1) Zender et al., CSH symposia on Quantitative Biology, 2005
(2) Lee JS et al., Nature genetics, 2006

Poster No: P021-F
Insulin like growth factor binding protein-7 reduces growth of human breast cancer cells and xenografted tumors.

[1,4] **Arun Seth**, [1]Tania Benatar, [1]Yutaka Amemiya, [1]Wenyi Yang, [2]Claire Holloway, [3]Sharon Nofech-Mozes
[1]*Division of Molecular and Cellular Biology, Sunnybrook Research Institute, Sunnybrook Hospital, Canada* [2]*Odette Cancer Center, Sunnybrook Hospital, Canada* [3]*Department of Anatomic Pathology, Sunnybrook Hospital, Canada* [4]*Department of Laboratory Medicine and Pathobiology, University of Toronto, Canada*

To identify genes correlated with breast tumorigenesis and eventual metastasis, gene expression profiles were generated comparing primary patient tumors with subsequent metastatic breast tumors developed in the NOD/SCID mouse model.

Using 28K whole genome microarrays we compared the gene expression profile of tumors that grew and metastasized in the NOD/SCID model with other primary breast tumors with different growth and metastatic potentials. Five hundred and eighty two genes were significantly differentially expressed by our statistical comparison criteria using the GeneSpring GX 7.3.1 software suite. In the metastatic set eight genes were selectively overexpressed (YB1, MMP7, MMP9, RAB5A, RABGDIB, EPHRB3, WNT2B, CSF1R) and 12 were selectively underexpressed (IGFBP7, GATA3, CST5, CDK6, SERBP1, MGP, TGF1L4, ESE3, ELF3, EDNRB, HECTD1, TINP1). Validation by qRT-PCR analysis of IGFBP7, GATA3 and CSF1R expression has been confirmed in the primary tumors metastatic to bone and host tissues. IGFBP-7 expression has been shown to be inversely correlated with disease progression in breast cancer. In order to investigate the role of IGFBP-7 in breast tumor suppression, it was overexpressed in the triple negative MDA-MB-468 human breast cancer line. Ectopic overexpression of IGFBP-7 clearly reduced the growth of the MDA-MB-468/IGFBP-7 cells compared to the parental MDA-MB-468 cells. Investigation of downstream signalling pathways affected by IGFBP-7 revealed that ectopic overexpression of IGFBP-7 strongly suppressed the phosphorylation of the ERK-1/2, while simultaneously increasing the phosphorylation of the ATK/PKB, a downstream effector of the PI3K pathway, suggesting that IGFBP-7 mediates its anti-proliferative effects through modification of multiple downstream signalling. When injected subcutaneously into NOD/SCID mice, the increased expression of IGFBP-7 in the MDA-MB-468/ IGFBP-7 cells reduced the rate of tumor growth in comparison to the parental MDA-MB-468 cells. These results suggest that the growth of breast cancer could be prevented by the forced expression of IGFBP-7 protein. Work in progress will test how systemic or local IGFBP-7 treatment affects human primary breast tumors and breast cancer cell-line growth and metastasis in our NOD/SCID models. Our studies are expected to characterize critical mediators of the IGFBP-7 senescence response, its potential co-regulatory function in breast tumor suppression, and its value as a novel drug.

Poster No: P036-F
Digital Gene Expression Profiling of Acute Lymphoblastic Leukemia Cells

[1]**Jessica Nordlund**, [1]Anna Kiialainen, [1]Olof Karlberg, [1]Anders Lundmark, [2]Mads Sønderkær, [2]Kåre Lehmann Nielsen, [3]Gudmar Lönnerholm & [1]Ann-Christine Syvänen
[1]Department of Medical Sciences, Uppsala University, Uppsala, Sweden, [2]Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Aalborg, Denmark, [3]For the Nordic Society of Pediatric Hematology and Oncology (NOPHO), and the Department of Children's and Women's Health, Uppsala University, Uppsala, Sweden

Acute lymphoblastic leukemia (ALL) is one of the most common malignancies and a leading cause of illness-related death of children in developed countries. ALL is a heterogeneous disease with distinct molecular subtypes, which differ in their response to chemotherapy and subsequent risk for relapse. Genome-wide transcription profiling studies of ALL with microarray technology are successful in identifying expression signatures for subtype classification and risk of relapse. However, microarray studies are hampered by a priori knowledge of the gene sequences, high background, and a limited dynamic range. Recent advances in next generation sequencing (NGS) technologies have enabled highly quantitative sequencing-based approaches for transcriptome analysis. This advancement has allowed us to overcome the limited accuracy of hybridization based methods, thus facilitating a new scale of transcriptome measurement. To detect differences in gene expression profiles of ALL cells from different cytogenetic subtypes and chemotherapeutic treatment response of patients, we generated genome-wide transcriptome profiles from 24 well characterized ALL patients. These patients consist of four with T-cell lineage ALL and 20 from different subclasses of B-cell lineage ALL: High Hyperploidy (HeH), BCR-ABL1, TEL-AML, and dic(9;20). For transcriptome analysis, high quality RNA was isolated from bone marrow or peripheral blood samples taken from the patients at diagnosis. We used Digital Gene Expression (DGE) tag profiling on the Illumina Genome Analyzer (GA) to profile poly-a selected mRNA. DGE is analogous to serial analysis of gene expression (SAGE), which measures the frequency of a tag in the sequence library population, providing a digital count read out of overall gene expression.

Each DGE library was sequenced on one lane of the Illumina GA. DGE analysis generated on average 9.5 million reads per sample, of which >80% mapped to the human transcriptome. Using bioinformatics approaches, we have annotated and quantified the expression of ~19,000 genes in ALL cells, and observed anti-sense transcription for approximately 50% of the genes. By pairwise comparison, we have detected differential gene expression between the distinct ALL subtypes and between patients in continuous remission compared to patients who have subsequently relapsed.

Poster No: P051-F

Detection of significant DNA copy number alterations in breast tumors from Mexican patients with high-resolution SNP arrays and GISTIC analysis

Hidalgo-Miranda A[1], Rodriguez-Cuevas S[2], Rebollar-Vega R[1], Romero-Cordoba S[1], Quintanar-Jurado V[1], Bautista-Pina V[2], Ciceron-Arellano I[1], Carrillo-Sanchez K[1], March-Mifsut S[1], Jimenez-Sanchez G[1], Uribe-Figueroa L[1].
[1]Laboratorio de Oncogenomica, Instituto Nacional de Medicina Genomica, Mexico, [2]FundacionMexicana de Fomento Educativo para la Prevencion y Deteccion Oportuna del Cancer de Mama, Mexico.

Breast cancer represents the second cause of cancer related death in Mexican women, and also represents an important health problem worldwide. Specific patterns of DNA copy number changes in breast tumors have been associated with particular tumor types and clinical behaviors. Analysis of these DNA copy number alterations have been carried out using different methods, like comparative genomic hybridization (either on metaphase chromosomes or different microarray platforms), and more recently, with high density oligonucleotide and SNP arrays. Increased array resolution, as well as the introduction of statistical methods to evaluate the significance of the copy number aberrations, has allowed the detection of particular gene targets in several human tumors. In order to define the pattern of statistically significant DNA copy number changes in breast tumors from Mexican patients, we used a high resolution SNP array to analyze the DNA copy number profiles of 89 breast normal/tumor pairs from Mexican patients. Tumor tissue and peripheral blood lymphocyte DNA from each patient was analyzed with the Affymetrix SNP 6.0 array. All tumor samples contained more than 80% tumor cells. Array normalization and paired copy number analysis using the SNP probes in the array was done using DChip, followed by segmentation using the GLAD algorithm and significant regions were identified using the Genomic Identification of Significant Targets in Cancer (GISTIC) method, both implemented in the GenePattern platform. GISTIC analysis identified 49 significant regions with copy number changes in the tumor compared to the paired normal tissue, including 22 amplifications and 27 deletions. In the amplifications events, seven were considered as “broad” (average: 72 Mb), 11 were “focal” (average: 91 kb) and four were considered as “both”. In the case of the deletions, 15 were broad, nine were focal and three were both. Significant amplified regions were located at, 1q23.2, 1q42.2, 3q26.33, 5q23.1, 6q21, 7q34, 8p12, 8q24.11, 9p21.3, 10p14, 11q13.3, 12p12.1, 13q32.3, 14q21.1, 16p13.13, 17p11.2, 17q11.2, 17q12, 17q21.32, 17q23.3, 17q25.1 and 20q13.2. These regions contained 421 genes. Significant regions with deletion were found at 1p36.12, 2q24.3, 3p21.1, 4p13, 5q15, 6q22.1, 6q26, 8p23.2, 9p24.3, 10q26.11, 11p15.5, 11q23.3, 13q14.2, 13q31.1, 13q33.2, 14q32.12, 15q12, 16q21, 16q23.3, 17p12, 17p11.2, 17q21.2, 18p11.31, 19p13.3, 21q21.1, 22q13.2, Xq21.31. Some of these regions have not been previously reported as common areas of copy number change in breast cancer making the analysis of the genes inside these significant regions an interesting research task. To our knowledge, this is one of the first reports regarding high-resolution DNA copy number aberrations in breast cancer from a Latin American population, as well as one of the first analysis using the SNP 6.0 array in combination with a statistical method for the detection of significant regions with DNA copy number changes in breast cancer.

Poster No: P066-F
Transcriptional profiling by massively parallel ligation DNA sequencing reveals much about alterations that occur during the development of oral cancer

[1]**David I Smith**, [1]Rebecca Laborde, [2]Brian Tuch, [2]Melissa Barker, and [2]Francisco De La Vegas
[1]Division of Experimental Pathology, Mayo Clinic, Rochester, MN, USA [2]Life Technologies, Foster City, CA, USA

We utilized massively parallel ligation DNA sequencing on the SOLid Platform to comprehensively analyze transcriptomic alterations that occur during the development of cancer of the oral tongue. We utilized a protocol which preserves the standedness of each transcrip present in three oral tongue tumors as well as in matched normal oral mucosa from the same patients. Over 200 million 50 base pair reads were obained fromeach sample and then the transcriptional profilie of oral tongue tumor was directly compared to the matched normal oral mucosa sample. This enabled us to ascertain transcripts that had altered expression during tumor formation. In addition, we could characterize allele-specific changes in expression for the more abundantly expressed genes. The genes that had alterations in allele-specific expression were members of the same molecular pathways as genes that had change in expression. We could also examine the more abundantly expressed genes to directly identify mutaitons presaent in numerous gene transcripts. These genes were also members of the same molecular pathways as genes that either had changes in overall expression or in allele-specific expression. Finally, using a novel program to identify and prioritize potential transcripts which were fusions between different genes we have identified a series of potential nvel gene fusions, several of which wre preent in at least two of the

three oral tumors characterized. This work demonstrates the power of sequence-directed transcriptional profiling to more completely characterize the genomic alterations that underlie the development of cancers of the oral tongue.

12. Genome Variation in Disease II

(21 May 2010, 10.30 AM - 12.00 NN, Salle Einstein) - Chaired by Y T Chen / Valere Cacheux-Rataboul

Poster No: P030-F
Identification of recent admixture in an Indian population of African ancestry: prospects for complex disease mapping

Mitali Mukerji[1], Ankita[2], Pankaj Jha[1], Vimal Rawat[2], Debasis Dash[2], Arijit Mukhopadhyay[1], Indian Genome Variation Consortium[1,2] and Analabha Basu[3]
[1]Genomics and Molecular Medicine, [2]G.N. Ramachandran Knowledge Centre for Genome Informatics, Institute of Genomics and Integrative Biology (CSIR), Delhi, India; National Institute of BioMedical Genomics, Kalyani, Kolkata, India

Study of genetic variation in recently admixed populations not only reveals historical population events but is also a powerful approach for mapping disease genes. This has been successfully demonstrated in a number of admixed populations like Africo-American, Hispanic and Uyghur. India occupies a center stage in human evolution. The genetic structure of Indian populations that represents a substantial fraction of global diversity, has been shaped by multiples waves of migration and local admixture events. We studied an out-group population (OG-W-IP) of African-Indian origin that are reported to be the descendants of slaves from the Bantu speaking parts of East Africa and have been residing in the western part of India for nearly 500 years. We have carried out this study using a common data of 18534 autosomal markers that have been genotyped on 26 Indian populations using 50 K Affymetrix array and are also available in the HGDP panel and HapMap populations. Principal component analysis revealed clear separation of African population from Indian populations. STRUCTURE and ADMIXTURE analysis revealed that overall the OG-W-IPs derive 53% of their genomic ancestry from their African past with very little inter-individual ancestry variation (sd = 0.09). The block sizes also reveal that the admixture event has been recent. These results corroborate with the history of the population. Therefore OG-W-IP is a potential candidate population for admixture mapping. Since OG-W-IP reside in a high salt area of Gujarat they provide opportunity to design and test the “African salt-retention gene theory” and studies related to hypertension in recently admixed population.

Poster No: P045-F
Dried reagents for multiplex genotyping by tag-array minisequencing in microfluidic devices

[1]**Annika Ahlford**, [2]Bastian Kjeldsen, [2]Jakob Reimers, [1]Anders Lundmark, [3]Massimo Romani, [2]Anders Wolff, [1]Ann-Christine Syvänen and [2]Monica Brivio
[1]Department of Medical Sciences, Molecular Medicine, Uppsala University,Uppsala, Sweden. [2] Department of Micro- and Nanotechnology, Technical University of Denmark, DTU Nanotech, Kongens Lyngby, Denmark .[3] Department of Advanced Diagnostic Technologies, Division of Tumor Genetics, National Cancer Research Institute (IST-Genova), Genova, Italy.

By integrating miniaturized components and automating assay functionalities, lab-on-a-chip technologies offer promising tools for fast, cost-effective and efficient point-of-care DNA analysis. External reagent supply to the microfluidic system is one difficulty for integration of complex DNA-based assays and storing reagents on the microfluidic device could reduce the complexity. In this study we have addressed reagent storage by freeze-drying all reagents required for multiplex genotyping using PCR amplification and tag-array minisequencing with fluorescence detection, to be implemented in microfluidic chambers for cancer diagnostics using the tumor protein 53 (TP53) gene as a model.

We performed tests to evaluate enzyme activity after freeze-drying in the presence of lyoprotectants and enzyme stability

after storage and established optimal reaction protocols. The activity of the thermostable DNA polymerase remained high when stored at -21ºC for 6 month after freeze-drying. By accelerated stability tests, the half-life of activities for exonuclease I and shrimp alkaline phosphatase, used for PCR ”clean-up”, were estimated to 50 and 80 days, respectively. We carried out a systematic genotyping experiment comparing freeze-dried and standard liquid reagents in each step of the protocol, genotyping the five most common cancer mutation sites (codons 175, 248 and 273) and 13 SNPs in TP53. An accuracy of 99.1% was achived for successful genotypes generated by freeze-dried reagents when compared to liquid reagent. As a proof of principle the complete genotyping protocol was successfully performed in cyclic olefin copolymer test chambers using freeze-dried reagents.

Our results indicate the feasibility to apply freeze-drying for reagent storage in the establishment of simple devices for multiplex DNA-analysis. Currently we are implementing the reagent deposition strategy in the development of a fully integrated and automated microfluidic chip for a complete genetic analysis comprising DNA extraction and amplification, genotyping and result readout. The system consists of a disposable biochip for all biochemical reactions and re-usable control components and functionalities.

Poster No: P060-F
Dissection of genes in the type I interferon pathway reveals two novel risk loci for SLE

[1]**Johanna K. Sandling**, [1,§]Sophie Garnier, [1,#]Snaevar Sigurdsson, [1]Chuan Wang, [2]Gunnel Nordmark, [3]Iva Gunnarsson, [3]Elisabet Svenungsson, [3]Leonid Padyukov, [4]Gunnar Sturfelt, [4]Andreas Jönsen, [4]Anders A. Bengtsson, [5]Lennart Truedsson, [6]Catharina Eriksson, [7]Solbritt Rantapää-Dahlqvist, [8]Anders Mälarstig, [8]Rona Strawbridge, [8]Anders Hamsten, [9]Robert R. Graham, [9]Timothy W. Behrens, [2]Maija-Leena Eloranta, [10]Gunnar Alm, [2]Lars Rönnblom and [1]Ann-Christine Syvänen
[1]Molecular Medicine, Department of Medical Sciences, Uppsala University, Uppsala, Sweden, [2]Section of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden, [3]Rheumatology Unit, Department of Medicine, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden, [4]Section of Rheumatology, Department Clinical Sciences, Lund University, Lund, Sweden, [5]Department of Laboratory Medicine, Section of MIG, Lund University, Lund, Sweden, [6]Department of Clinical Immunology, Umeå University Hospital, Umeå, Sweden, [7]Department of Rheumatology, Umeå University Hospital, Umeå, Sweden, [8]Atherosclerosis Research Unit, Department of Medicine, Solna Karolinska Institutet Stockholm, Sweden, [9]Genentech Inc., South San Francisco, CA, USA, [10]Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Uppsala, Sweden, [§]Current address INSERM UMR S937, Génétique Epidémiologique et Moléculaire des, Pathologies Cardiovasculaires, Université Paris VI, Paris, France, [#]Current address Broad Institute of Harvard and MIT, Boston, MA, USA

Systemic Lupus Erythematosus (SLE) is a systemic autoimmune disease in which the type I interferon (IFN) pathway plays a crucial role. We have previously shown that five genes in this pathway, IRF5, TYK2, STAT4, IFIH1 and IRF8 are associated with risk for SLE, as well as successfully replicated the previous reports on association for IRAK1, TNFAIP3, TNFSF4 and the IRF7 region (KIAA1542). Here we investigate 78 genes involved in the type I IFN pathway to identify additional SLE susceptibility loci. First, we genotyped SNPs in these 78 genes and 14 other candidate genes in Swedish SLE patients and controls. Genes with $P < 0.01$ in the initial screen were then followed up in an additional Swedish cohort. SNPs in five genes were nominally associated with SLE in this extended cohort. To replicate these findings we extracted data from a genome-wide association study on SLE performed in a US cohort. Combined analysis of the Swedish and US data confirmed two of these genes as SLE susceptibility loci. Our study highlights additional genes from the type I IFN system for further functional analysis, and more specifically points to the importance of genes in the IFIH1/DDX58 pathway which is activated in cells other than the plasmacytoid dendritic cells, for example monocyte derived dendritic cells, in response to viral infections.

Poster No: P089-F
SNPome: Human Polygenic Disease SNP Network

Yasha Hasija and Samir K. Brahmachari
Council of Scientific and Industrial Research, Anusandhan Bhawan, 2 Rafi Marg, New Delhi-110001, INDIA

Common complex disorders follow an intricate pattern of inheritance, characterized by incomplete penetrance, genetic heterogeneity, as well as polygenic predisposition. In addition, environmental factors also play an important role in contributing to disease phenotypes. A gain in understanding of the association between genetic variations and their phenotypic effects is a step towards comprehending the etiology of these disorders. While genome-wide association studies are, to some extent, seeming to explain genetic predisposition to common diseases, the significance of allelic overlap in complex disorders cannot be undermined. With this preview, in the present study, the complex disorders were analyzed for the shared nature of their associated SNPs, and a network of complex disorders having common associated SNPs was constructed. These complex disorders were then classified into discrete groups based on the pathology and organ system involved. The biological systems sharing maximum overlap of associated SNPs were identified. We tried to address two basic questions through our analysis: (a) Are there any common SNPs that are essential to disorders of a particular biological system? and (b) Are there any generalized SNPs that may affect multiple human biological systems?

Through this study, we identified SNPs associated with more than one disorder of the same human system (designated as 'Intra-Systemic Overlapping SNPs', indicative of genetic differences that make an individual prone to the disorders of a particular human biological system) as well as SNPs associated with disorders of multiple biological systems (designated as 'Inter-Systemic Overlapping SNPs'). The genes harboring these inter- systemic overlapping SNPs may be central to the core intrinsic pathways, or may have an effect by interacting with other genes, that would involve multiple biological systems. A comprehensive network of these SNPs associated with more than one polygenic disorder, including both inter- and intra- systemic overlapping SNPs, was thereby constructed, hence building human polygenic disease SNP network, the SNPome.

The study emphasizes the shared nature of common alleles in related complex disorders, such as Type 2 diabetes mellitus and obesity, schizophrenia and bipolar disorder, etc., as well as in non-related polygenic disorders, such as the ones associated with cardiovascular and musculoskeletal system disorders.

13. Epigenetics and microRNA-mediated gene regulation

(21 May 2010, 10.30 AM - 12.00 NN, Sully 1) - Chaired by Lim Bing / Jean-Marie Blanchard

Poster No: P007-F

A mutation in the 3'UTR of the HDAC6 gene abolishing the post-transcriptional regulation mediated by hsa-miR-433 is linked to a new form of dominant X-linked chondrodysplasia.

[1,2]**Benoit Arveiler**, [1]Delphine Simon, [1]Malika Barillot, [3]Benoit Laloo, [1]Thomas Barnetche, [1,2]Caroline Rooryck, [1]Camille Blanchard, [1]Ingrid Burgelin, [1]Michèle Marche, [1]Isabelle Coupry, [4,5]Nicolas Chassaing, [6]Brigitte Gilbert-Dussardier, [3]Christophe Grosset, [1,2]Didier Lacombe;
[1]Université Victor Segalen Bordeaux 2, Bordeaux, France [2]CHU de Bordeaux, Bordeaux, France, [3]INSERM U889, Bordeaux, France, [4]INSERM U563, Toulouse, France, [5]CHU de Toulouse, Toulouse, France, [6]CHU de Poitiers, Poitiers, France.

A family with dominant X-linked chondrodysplasia was previously described. The disease locus was ascribed to a 24 Mb interval in Xp11.3-q13.1. We have identified a variant (c.*281A>T) in the 3'UTR of the HDAC6 gene that totally segregates with the disease. The variant is located in the seed sequence of hsa-miR-433. Our data showed that, in MG63 osteosarcoma cells, hsa-miR-433 (miR433) down-regulated both the expression of endogenous HDAC6 and that of an eGFP-reporter mRNA bearing the wild-type 3'UTR of HDAC6. This effect was totally abrogated when the reporter mRNA bore the mutated HDAC6 3'UTR. The HDAC6 protein was found to be over-expressed in thymus from an affected male foetus. Concomitantly, the level of total alpha-tubulin, a target of HDAC6, was found to be increased in the affected foetal thymus, whereas the level of acetylated alpha-tubulin was found to be profoundly decreased. Skin biopsies were obtained from a female patient who presented a striking body asymmetry with hypotrophy of the left limbs. The mutated HDAC6 allele was expressed in 31% of left arm-derived fibroblasts, whereas it was not expressed in the right arm. Overexpression of HDAC6 was observed in left arm-derived fibroblasts. Altogether these results strongly suggest that this HDAC6 3'UTR variant suppressed hsa-miR-433-mediated post-transcriptional regulation causing the overexpression of HDAC6. This variant is likely to constitute the molecular cause of this new form of X-linked chondrodysplasia. This

represents to our knowledge the first example of a skeletal disease caused by the loss of a miRNA-mediated post-transcriptional regulation on its target mRNA.

Poster No: P024-F
Dynamic Changes in the Human Methylome During Differentiation

Eleanor Wong[3, 7 *], Louise Laurent[1, 2 *] and Chia-Lin Wei[3, 7 #]
[1]UCSD Medical Center, Department of Reproductive Medicine. 200 West Arbor Drive, Mail Code 8433. San Diego, CA 92103 [2] Center for Regenerative Medicine, Dept of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 [3] Genome Technology & Biology Group, Genome Institute of Singapore, Singapore 138672. [7] Department of Biological Sciences, National University of Singapore

DNA methylation is a critical epigenetic regulator in mammalian development. In our study, we present a whole genome comparative view of DNA methylation using bisulfite sequencing of three cultured cell types representing progressive stages of differentiation: human embryonic stem cells (hESCs), a fibroblastic differentiated derivative of the hESCs, and neonatal fibroblasts. We observed many notable common and cell type-specific features among all cell types. Promoter hypomethylation (both CG and CA) and higher levels of gene body methylation were positively correlated with transcription in all cell types. Exons were more highly methylated than introns and sharp transitions of methylation occurred at exon-intron boundaries, suggesting a role for differential methylation in transcript splicing. Developmental stage was reflected in both the level of global methylation and extent of non-CpG methylation, with hESC highest, fibroblasts intermediate, and monocytes lowest. Differentiation-associated differential methylation profiles were observed for developmentally regulated genes, including the HOX clusters, other homeobox transcription factors, and pluripotence-associated genes such as POU5F1, TCF3, and KLF4. Our study highlight the value of high-resolution methylation maps, in conjunction with other systems-level analyses, for investigation of previously undetectable developmental regulatory mechanisms.

Poster No: P039-F
Integrated analysis of microRNA and mRNA expression profiles in physiological myelopoieis: role of hsa-mir-299-5p in CD34+ progenitor cells commitment

Elena Tenedini[1,4], Enrica Roncaglia[1,4], Francesco Ferrari[2], Claudia Orlandi[1], Elisa Bianchi[1], Silvio Bicciato[1,3], Enrico Tagliafico[1,3,4] and Sergio Ferrari[1,3,4]
[1]Department of Biomedical Sciences, University of Modena and Reggio Emilia, Modena, Italy; [2]Department of Biology, University of Padova, Padova, Italy; [3]Center for Genome Research, University of Modena and Reggio Emilia, Modena, Italy; [4]BioPharmaNet, Emilia-Romagna High-Tech Network, Ferarra, BioPharmaNet, Emilia-Romagna High-Tech Network, Ferarra, Italy

Hematopoiesis entails a series of hierarchically organized events that proceed throughout cell specification and terminates with cell differentiation. Commitment needs the transcription factors effort that, in concert with microRNAs, drives cell fate and responds to promiscuous patterns of gene expression by turning-on lineage-specific genes and repressing alternate lineage transcripts. We obtained microRNAs profiles from human CD34+ hematopoietic progenitor cells and in-vitro differentiated erythroblasts, megakaryoblasts, monoblasts and myeloblasts precursors, that we analyzed together with their gene expression profiles. The integrated analysis of microRNA-mRNA expression levels highlighted an inverse correlation between microRNAs specifically up-regulated in one single cell progeny and their putative target genes, which resulted down-regulated. Among the up-regulated lineage-enriched microRNAs, hsa-miR-299-5p emerged as having a role in controlling CD34+ progenitors fate, grown in multilineage culture conditions. Gain- and loss-of-function experiments revealed that hsa-miR-299-5p participates the regulation of hematopoietic progenitors fate, modulating megakaryocytic-granulocytic versus erythroid-monocytic differentiation.

Poster No: P054-F
Prediction of novel microRNAs on human Y-Chromosome

Arijit Mukhopadhyay[1,2], Saurabh V Laddha[1] and Subhashree Nayak[2]
[1]G.N. Ramachandran Knowledge Centre for Genome Informatics, [2]Genomics & Molecular Medicine, Institute of Genomics & Integrative Biology (CSIR), Delhi, India

MicroRNAs are endogenous non-coding RNAs of ~22 nucleotides and shown to regulate gene expression at the posttranscriptional level by directly cleaving targeted mRNAs or repressing translation. There are almost 1000 mature miRNAs (miRBase, release-14) discovered in the human genome which are distributed non-randomly. So far not even a single miRNA has been reported from the human Y-chromosome. In this study we wanted to explore the possibility of Y-chromosome specific miRNA in human.

We have taken a sliding window of 125 nucleotides to scan the entire sequence of Y-chromosome available in the public domain (GRCh37 build, NCBI). We then predicted the possible secondary structure for each window using recently reported tool CID-miRNA (http://mirna.jnu.ac.in/cidmirna/). This was further filtered using a more stringent prediction tool, Vir-Mir (http://140.109.42.4/cgi-bin/miRNA/miRNA.cgi) to reduce false positives. To achieve higher specificity we have used the sequences specific to Y-chromosome followed by presence of Dicer substrate site using PHDcleav (http://imtech.res.in/raghava/phdcleav/submission.html).
The initial prediction using CID-miRNA gave 1685 sequences with 'grammar' cutoff -0.609 for human specific prediction. Vir-Mir predicted a subset of 1135 sequences to be positive. When searched for specificity to Y-chromosome this was further reduced to 798 sequences. Dicer substrate site prediction gave 176 out of 798 sequences after using a cutoff score of 2.0. These were then matched against EST records and for 59 sequences there were at least an EST showing evidence for expression.

Our informatics analysis shows that there are many sequences on the human Y-chromosome that are potential precursor microRNA. The experimental validation is in process and will be presented in the meeting. The identification of Y-chromsome specific miRNA might lead to yet unknown ways of regulating male specific expression of genes and/or gender specific epigenetic regulation.

Poster No: P069-F
When 11p15.5 genomic imprinting meets pluripotency factors

[1]**Julie Demars**, [1]Mansur Ennuri Shmela, [2]Sylvie Rossignol, [2]Irène Netchine, [1]Jun Okabe, [2]Salah Azzi, [2]Sylvie Cabrol, [3,4]Cédric Le Caignec, [3]Albert David, [2]Yves Le Bouc, [1]Assam El-Osta, [1]Christine Gicquel
[1]Epigenetics in Human Health and Disease, Baker IDI Heart and Diabetes Institute, Melbourne, 3004, Victoria, Australia. [2]APHP, Hopital Armand Trousseau, Laboratoire d'Explorations Fonctionnelles Endocriniennes, Unité mixte de recherche INSERM UMPC U938, Paris, 75012, France. [3]Service de Génétique Médicale, CHU, Nantes, 44093, France. [4]INSERM, UMR915, Institut du thorax, Nantes, 44035, France.

The paternally expressed IGF2 and maternally expressed H19 genes are located on the human 11p15.5 chromosomal region. A 4.2kb, paternally methylated, germline imprinting control region 1 (ICR1) between IGF2 and H19 is responsible for monoallelic expression of both IGF2 and H19. The ICR1 domain works as a methylation-sensitive chromatin insulator by binding the CCCTC-binding factor (CTCF) in a parent-specific manner. Many studies point to a crucial role of CTCF in maintaining rather than establishing imprinting of the ICR1 region. However, the germline-specific processes that establish the differential ICR1 methylation in the female and male germlines remain unknown.
DNA methylation defects involving the IGF2/H19 domain result in two fetal growth disorders with opposite phenotypes: the Beckwith-Wiedemann syndrome (BWS), an overgrowth disorder (maternal ICR1 gain of methylation in 10% of BWS cases) and the Silver-Russell syndrome (SRS), a growth retardation disorder (paternal ICR1 loss of methylation in 60% of SRS cases). Maternally-inherited deletions involving CTCF binding sites located within the ICR1 domain have been reported in a few BWS familial cases; however, attempts to identify such deletions or mutations in ICR1 CTCF binding sites in SRS patients, including a few familial cases, were negative.

We therefore hypothesized that other sequences than CTCF binding sites located within the IGF2/H19 domain could be implicated in patients with ICR1 DNA methylation defects. We investigated extensively the ICR1 domain in a group of patients (21 BWS and 16 SRS) with an ICR1 gain/loss of methylation and no genetic abnormalities involving CTCF binding sites.

We identified three novel ICR1 imprinting anomalies in BWS patients. Those defects consisting of small deletions and a single mutation do not affect one of the seven CTCF binding sites. Interestingly, all of them involved pluripotency factor binding sequences, more particularly OCT4 and SOX2 binding sites. We also demonstrated by gel shift assay analysis that i) OCT4 and SOX2 proteins bind the identified sequences at the IGF2/H19 domain; ii) genetic anomalies identified in the three BWS patients with an ICR1 gain of methylation alter the binding of OCT4 and SOX2 pluripotency factors. Altogether, those results highlight the importance of other transcription factors (OCT4 and SOX2) than CTCF in the regulation of genomic imprinting at the 11p15.5 region. Further studies will investigate whether pluripotency factors play a role in the ICR1 differential acquisition of DNA methylation in germlines.

Poster No: P083-W
MicroRNA function in regulating pluripotency and differentiation of Human Embryonic Stem cells

Lay Teng Ang[1,2], Ludovic Vallier[1] and Bing Lim[2,3]
[1]*Laboratory for Regenerative Medicine, University of Cambridge, Department of Surgery. Cambridge UK* [2]*Stem Cell and Developmental Biology Group, Genome Institute of Singapore, Agency for Science, Technology and Research, 138672*
[3]*Beth Israel Deaconess Medical Center, Harvard Medical School*

MicroRNAs (miRNAs) are key regulators of gene expression and are essential for early development in mouse embryos. The requirement of miRNAs is supported by genetic studies in the mouse showing that loss of Dicer or Dgcr8, key components of miRNA processing machinery, resulted in developmental arrest and embryonic lethality. In addition, mouse ESCs mutant for either enzyme, are defective in downregulating pluripotency markers upon differentiation, indicating that miRNAs are necessary to inhibit pluripotency during differentiation. Furthermore, knockouts of individual miRNA families such as miR-1 and miR-17-92b resulted in severe developmental defects in the mouse. Despite the importance of miRNAs in early development and the rapid identification of miRNAs, the possible function of many of these miRNAs remains to be precisely defined. This project aims to investigate the function of miRNA in regulating mechanisms governing pluripotency and early cell fate specification of mammals, using human ESCs as an in vitro model of development.

MiRNAs that could potentially regulate pluripotency and differentiation were identified by performing miRNA expression profiling of the specification of hESCs into definitive endoderm. Importantly, high level of homogeneity was obtained using a fully chemically defined system and an efficient differentiation protocol. Using this approach, more than 20 miRNAs were found to be up or down regulated upon differentiation of hESCs. The expression data was validated with Taqman microRNA assay and miRNAs showing significant change in expression were functionally characterized by over-expression and inhibition studies.

Together these results will provide a comprehensive model explaining the function of miRNA in the transition between pluripotency and the endodermal pathway and expand the key gene regulatory network regulating these processes. The understanding of these mechanisms is essential for the generation of stable and fully functional endodermal cell states for future regenerative medicine applications.

14. Structural Variations and Chromosomal Aberrations

(21 May 2010, 10.30 AM - 12.00 NN, Sully 2) - Chaired by Lawrence Stanton / Richard Redon

Poster No: P017-T
Increased Sensitivity of Copy Number Variation Detection Using a High-Density Array CGH Platform

Christine J. Shaw, Kathleen O'Moore, Joel Geoghegan, Nan Jiang, Todd A. Richmond, Vanessa L. Ott, Rohaizah James, Rebecca R. Selzer
Roche NimbleGen, 504 S. Rosa Road, Madison, WI 53719, USA

Array CGH methods have been widely used to investigate genome-wide DNA copy number variation associated with complex disorders. Disease association studies have become increasingly focused on CNVs, and several recent reports show links between CNVs and schizophrenia (Stefansson 2008; Stone 2008), autism (Sebat 2007; Marshall 2008; Glessner 2009), and cancer, among others. More recently, the focus of CNV research has migrated to detection of rare variants, with an allelic frequency < 5% (Conrad 2009, Manolio 2009). In an effort to increase detection of rare variants as well as more common CNVs, we sought to develop the highest density oligo array available for CNV detection, as well as a more sensitive algorithm for CNV detection on NimbleGen arrays.

The 2.1M and 3x720K CNV-focused arrays contain empirically optimized probes, with the most comprehensive collection of targeted regions available, including Asian population-specific CNV regions. Utilization of the CNV-focused arrays enabled detection of hundreds of CNVs per individual. When compared to whole genome tiling arrays, 2- to 3- fold more CNVs were detected with the CNV-focused arrays. When the Roche-NimbleGen CNV-focused arrays were compared to competitive array platforms, we found an average of 300% and 245% more CNVs were detected with the 2.1M and the 3x720 platforms, respectively. Additionally, the higher probe density of Roche-NimbleGen arrays allowed 2-fold higher detection of CNVs < 1,000 bp in size when compared to competitive array platforms. The increased sensitivity of CNV detection was aided by an improved segmentation algorithm (segMNT), included in NimbleScan v2.6 software. A comparison between the segMNT v1.1 and segMNT v1.2 algorithms shows increased sensitivity in detection of CNVs resulting in about 2-fold more CNVs detected per sample with segMNT v1.2.

We demonstrate here the variability in sensitivity of CNV detection based on the platform and algorithm utilized. We show the increased detection of total CNVs, as well as CNVs of smaller size with the Roche NimbleGen CNV-focused arrays and NimbleScan v2.6 software.

Poster No: P018-F
Stoichiometry of base excision repair proteins contributes to the tissue selectivity of somatic CAG expansion in Huntington's disease mice.

Merienne K[1] Goula A-V [1], Berquist, BR [2], Wheeler VC [3], Wilson DM 3rd[2] and Trottier Y [1]
[1]*Department of Neurobiology and Genetics, Institut de Génétique et Biologie Moléculaire et Cellulaire, UMR 7104-CNRS/INSERM/ULP, BP10142, 67404 Illkirch Cédex, CU de Strasbourg, France* [2]*Laboratory of Gerontology, National Institute on Aging, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA* [3]*Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston MA 02114, USA*

Huntington's disease (HD) is an inherited progressive neurodegenerative disorder caused by an unstable CAG repeat in the Huntingtin gene. Instability of the mutation affects both germline and somatic cells. Somatic instability is tissue-specific and increases with age. In particular, the CAG repeat sequence in the striatum, the brain region that preferentially degenerates in HD, is highly unstable, whereas it is rather stable in the disease-spared cerebellum. The mechanisms underlying the age-dependence and tissue-specificity of somatic CAG instability remain elusive. Recent studies have suggested that DNA oxidation and OGG1, a glycosylase involved in the repair of 8-oxoG lesions, contribute to this process. We show that in HD mice oxidative DNA damage alone is not sufficient to trigger somatic instability. Protein levels and activities of major base excision repair (BER) enzymes were compared between striatum and cerebellum of HD mice. Strikingly, 5'-flap endonuclease activity was much lower in the striatum than in the cerebellum of HD mice. Accordingly, Flap Endonuclease-1 (FEN1), the main enzyme responsible for 5'-flap endonuclease activity was also significantly lower in the striatum compared to the cerebellum. Finally, chromatin immunoprecipitation experiments revealed that POLb was specifically enriched at CAG expansions in the striatum, but not in the cerebellum of HD mice. These in vivo data fit a model in which POLb strand displacement activity during LP-BER promotes the formation of stable 5'-flap structures at CAG repeats representing pre-expanded intermediates structures, which are not efficiently removed when FEN1 activity is constitutively low. We propose that the stoichiometry of BER enzymes is one critical factor underlying the tissue selectivity of somatic CAG expansions.

Support: CNRS/INSERM/Université de Strasbourg, Hereditary Disease Foundation

Poster No: P033-F
TaqMan® real time PCR based assays for detecting mouse copy number variations

[1]**Xiaoqing You**, [1]Annie Titus,[1]Fawn Wang, [1]Kelly Li, [1]Yu Wang
[1]Life Technologies, 850 Lincoln Centre Drive, Foster City, CA 94404, USA

Copy number variations (CNV) are found in many species including human and mouse. CNVs have been directly linked to biological functions and have been associated with genetic disorders and diseases. Mouse is the most commonly used mammalian model organism for genetics and disease research. The study of mouse copy number variations plays an important role in understanding functions and biological impact of genomic variations.

We developed quantitative TaqMan® real time PCR assays that can be used to detect copy number variations in mice and to screen transgenic and knockout mice. Each assay reaction consists of a duplex assay: a FAM™ dye-based target assay designed to detect a target-of-interest on genome and VIC® dye-based reference assay that detects a sequence known to be present in two copies in a diploid genome. The copy number of the target of interest is determined based on the relative quantification of the target assay versus the reference assay.

We designed mouse TaqMan® copy number reference assays and 28 TaqMan® copy number assays that target 18 mouse genes using proprietary computational algorithms. We also designed assays for reporter genes (including neomycin resistance (Neo) gene, beta-D-galactosidase LacZ and green fluorescent protein (GFP)) that are frequently used as markers in targeted mouse mutagenesis and gene manipulation. The assays were tested on 48 DNA samples (including aneuploidy samples with known chromosome abnormality) from 17 mouse strains and 40 transgenic and knockout samples. The copy number results were subsequently compared with the expected genotypes of mouse samples. Our studies demonstrate that the assays detect copy number variations of different mouse strains/sub-strains with high accuracy and specificity.

Poster No: P048-F
Copy number variation in severe early onset childhood obesity

[1]**Bochukova E.G.***, [2]Huang N., [1]Keogh J.,[1]Henning E.,[1] O'Rahilly R.,[2] Hurles M.E.,[1] Farooqi I.S.
[1] University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK [2] Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK

Severe obesity is highly heritable and genetically heterogeneous disorder. We set out to investigate the contribution of copy number variation to obesity in the Genetics of Obesity Study cohort (GOOS), using Affymetrix SNP 6.0 microarrays. Our discovery set consisted of 300 Caucasian patients with severe early-onset obesity, approximately half of whom also had developmental delay. Large (>500 kb), rare (<1%) deletions were significantly enriched in patients compared to 7,366 controls (P<0.001). Multiple rare copy number variants (CNVs) were found recurrent in patients but absent or at much lower prevalence in controls.

The most common recurrent CNV was identified in five patients with overlapping deletions on chromosome 16p11.2, only found in 2 out of 7,366 controls (P<5x10-5). The 16p11.2 deletion co-segregated with severe obesity in three of the patient's families. Two of the patients harboured a larger de novo 16p11.2 deletion, extending through a neighbouring CNV region where deletions have previously been associated with autism and mental retardation phenotypes; both of these patients had mild developmental delay in addition to severe obesity. The 220 kb region of minimal overlap of the 16p11.2 deletions contains several genes including SH2B1, known to be involved in leptin and insulin signalling. Deletion carriers exhibited hyperphagia and severe insulin resistance disproportionate for the degree of obesity. The phenotype of the patients was consistent with SH2B1 haploinsufficiency as an underlying molecular mechanism. We have then broadened our investigation into the extent of CNV contribution to early-onset obesity phenotype in the SCOOP cohort: 1,000 UK Caucasian patients with severe early-onset obesity from GOOS. This dataset was analysed for the presence of recurrent CNVs, previously reported by us. We found many of the recurrent CNVs present in this replication cohort, including the SH2B1-containing deletion.

In conclusion, we show that copy number variation contributes significantly to the genetics of human obesity, warranting a further investigation into the influence of CNV-contained genes on the obesity phenotype.

Poster No: P063-F
Use of DNA –PET technology in human congenital disorders: a new whole genome approach

Valere Cacheux[1], Kagistia.Hana Utami[1], Elaine Chew[1], Kim Ng[1], Zhang Zhenshui[1] Seong Soo Lim[1], Pascal de Santa Barbara[2], Sylvain Briault[3], Robyn Jamieson[4], Edison Liu[1], Yijun Ruan[1], Axel Hillmer[1],
[1]Genome Institute of Singapore, [2]InsermERI25 Montpellier, France,[3] CHROrleans France,[4]The Children's Hospital at Westmead, Sydney, Australia.

Identification and characterization of chromosomal anomalies and DNA alterations in human congenital disorders are still one of the biggest challenges in human genetics. Considering the limitations of current technologies (resolution for karyotype, coverage and inability to detect balanced rearrangements for array-CGH, time-consumption and low resolution for breakpoint mapping techniques), most of the cases are not detected or not widely explored. Recently, the development of whole genome paired-end sequencing approaches (DNA-PET) has demonstrated that such technologies can identify copy number variations, balanced and unbalanced DNA rearrangements, and then map the breakpoints at the gene level in normal and cancer genomes.

In order to evaluate the ability of this technology to detect rearrangements and pinpoint the putative causative gene in congenital disorders, we applied the DNA-PET technology to 5 independent patients presenting a birth defect with or without known chromosomal anomalies (selected by karyotype).

The DNA of these 5 patients was subjected to DNA-PET and high throughput sequencing processes. After random shearing, the DNA fragments were processed for extraction of PET construct and paired end sequencing. The PET sequences were then mapped to the human reference genome (Hg18) and structural variations detected as discordant PETs. Then, a Copy Number Variations (CNVs) filtration and a cross comparison using normal DNAs libraries were done and revealed specific and unique structural variations for each sample (between 50-100). Results obtained from the 5 DNAs allowed the detection of specific breakpoints of chromosomal rearrangements and the causative disease gene in 2 of the patients, the identification of candidate genes for 2 patients and a technical issue of interpretation for one patient. We present here the results, and discuss the advantages and current improvements of DNA-PET technology as a complete tool for identification, mapping and sequencing of chromosome rearrangements to identify new genetic syndromes.

Poster No.: P145-F
Deletion of SULF1 and SLCO5A1 genes at 8q13 causes mesomelia-synostoses syndrome

Cédric Le Caignec[1,2,3,4], Bertrand Isidor[1,13], Olivier Pichon[1,13], Richard Redon[1,2,3,4], Debra Day-Salvatore[5], Antoine Hamel[6], Lena Kjellén[7], Cornelia Kraus[8], Jules G. Leroy[9], Geert R. Mortier[10], Anita Rauch[11], Alain Verloes[12], Albert David[1]
[1],CHU Nantes, Service de Génétique Médicale, Nantes, F44093, France [2]INSERM, UMR915, l'institut du thorax, Nantes, F44007, France [3]CNRS, ERL3147, Nantes, F44007, France [4]Université de Nantes, Nantes, F44035, France [5]Institute for Genetic Medicine, Saint Peter's University Hospital, New Brunswick, New Jersey, NJ08901, US [6]CHU Nantes, Service d'Orthopédie Pédiatrique, Nantes, F44093, France [7]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, SE75123, Sweden [8]Institute of Human Genetics, University of Erlangen-Nuremberg, Erlangen, D91054, Germany [9]Department of Medical Genetics, Ghent University Hospital, Ghent, B9000, Belgium [10]Department of Medical Genetics, Antwerp University Hospital, Antwerp, B2610, Belgium [11]Institute of Medical Genetics, University of Zurich, Schwerzenbach-Zurich, CH8603, Switzerland [12]Department of Clinical Genetics and INSERM U676, Robert Debré University Hospital, Paris, F75019, France

Mesomelia-synostoses syndrome (MSS) or mesomelic dysplasia with acral dysostoses Verloes-David-Pfeiffer type is a rare autosomal dominant disorder characterized by mesomelic limb shortening, acral synostoses and multiple congenital malformations. So far, five patients in four unrelated families have been reported worldwide with MMS. Using whole genome oligonucleotide array CGH, we have identified an interstitial deletion at 8q13 in all patients. The deletions vary from 582 kb to 738 kb in size, but invariably encompass only two genes: SULF1, encoding the heparan sulfate 6-O-endosulfatase 1 and SLCO5A1, encoding the solute carrier organic anion transporter family member 5A1. SULF1 acts as a regulator of numerous growth factors in skeletal embryonic development while the function of SLCO5A1 is yet unknown. Breakpoint sequence analyses performed in two families showed non-recurrent deletions. Our results strongly suggest that haploinsufficiency of SULF1 contributes to this mesomelic chondrodysplasia, highlighting the critical role of endosulfatase in human skeletal development. As co-deletion of SULF1 and SLCO5A1 - which does not result from a low-copy repeats (LCRs)-mediated recombination event - was found in all patients, we suggest that haploinsufficiency of SULF1 combined with haploinsufficiency of SLCO5A1 (or the altered expression of a neighbouring gene through a position effect) could be necessary in the pathogenesis of MSS.

15. Functional Genomics

(21 May 2010, 10.30 AM - 12.00 NN, Sully 3) - Chaired by Sumio Sugano / Alain Bucheton

Poster No: P010-F
A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing

Xavier Estivill[1,2,6], Eulàlia Martí[1,2], Mónica Bañez-Coronel[1,2], Lorena Pantano[1,2], Franc Llorens[3,4] and Isidre Ferrer[4,5]
[1]Genetic Causes of Disease Group, Genes and Disease Program, Centre for Genomic Regulation (CRG), Barcelona, Catalonia, Spain; [2]Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública; [3]Molecular and Cellular Neurobiotechnology, Institute of Bioengineering of Catalonia and Department of Cell Biology, University of Barcelona; [4]Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas; [5]Institut Neuropatologia, Servei Anatomia Patològica, IDIBELL-Hospital Universitari de Bellvitge, Universitat de Barcelona; [6]Experimental and Health Sciences Department, Pompeu Fabra University, Barcelona, Catalonia, Spain

Huntington disease (HD) is a neurodegenerative disorder that predominantly affects neurons of the forebrain. We have applied the Illumina massive parallel sequencing to deeply analyze the small RNA populations of two different forebrain areas, the frontal cortex (FC) and the striatum (ST) of healthy individuals and individuals with HD. More than 80% of the small-RNAs were annotated as miRNAs in all samples. Deep sequencing revealed length and sequence heterogeneity (IsomiRs) for the vast majority of miRNAs. Around 80-90% of the miRNAs presented modifications in the 3'-terminus mainly in the form of trimming and/or as nucleotide addition variants, while the 5'-terminus of the miRNAs was specially protected from changes. Expression profiling showed strong miRNA and isomiR expression deregulation in HD, most being common to both FC and ST. The putative targets of the seed-region of deregulated miRNAs/isomiRs strongly suggest that their altered expression contribute to the aberrant gene expression of HD. Many REST (RE1-Silencing Transcription Factor) modulated miRNAs were downregulated in HD, suggesting that repressed REST target miRNAs play a role in aberrant gene expression in HD. Our results show that miRNA variability is a ubiquitous phenomenon in the adult human brain, which may influence the mechanism of gene expression modulation.

Poster No: P027-F
Common genetic determinants of vitamin D insufficiency: a meta-analysis of genome-wide association studies

Elina Hypponen, The SUNLIGHT consortium*
**King's College London, UCL Institute of Child Health, Imperial College, and Whipps Cross Rheumatology Department, London. University of Cambridge, Addenbrooke's Hospital, Wellcome Trust Sanger Institute Cambridge. University of Oxford, University of Aberdeen, University of Southampton and University of Liverpool (UK). Harvard Medical School,*

Boston University, Institute for Ageing Research, Tufts University, and Massachusetts General Hospital, Boston, Broad Institute, Cambridge, MA. University of Washington, Seattle, WA; University of Maryland Baltimore MD; National Institute on Aging, National Institutes of Health, Bethesda MD; University of Pittsburgh, Pittsburgh PA; Cedars-Sinai Medical Center, Los Angeles CA; Wake Forest University, Winston-Salem, NC; University of Miami, FL, Indiana University (US). McGill University, Montreal, Health Canada, Ottawa, Public Health Agency of Canada, Toronto (Canada). National Institute of Health and Welfare, University of Oulu, Finnish Institute of Occupational Health, Oulu. University of Helsinki (Finland). University of Gothenburg, (Sweden). Erasmus Medical Center, Rotterdam (Netherlands), Klinikum der Johann Wolfgang Goethe University, Frankfurt (Germany). Genome Institute of Singapore, Biopolis (Singapore).

Background: Vitamin D is crucial for maintaining musculoskeletal health. Recently, vitamin D insufficiency has been linked to a number of extraskeletal disorders, including diabetes, cancer, and cardiovascular disease. Determinants of circulating 25-hydroxyvitamin D (25-OH D) include sun exposure and dietary intake, but its high heritability suggests that genetic determinants may also play a role.

Methods: We performed a genome-wide association study of 25-OH D among ~30,000 individuals of European descent from 15 cohorts. Five cohorts were designated as discovery cohorts (n=16,125), five as in silico replication cohorts (n=9,366), and five as de novo replication cohorts (n=8,378). Association results were combined using z-score-weighted meta-analysis. Vitamin D insufficiency was defined as 25-OH D <75 nmol/L.

Findings: Variants at three loci reached genome-wide significance in the discovery cohorts, and were confirmed in the replication cohorts: 4p12 (overall P=1.9 x 10-109); 11q12 (P=2.1 x 10-27); 11p15 (P=3.3 x 10-20). Variants at an additional locus (20q13) were genome-wide significant in the pooled sample (P=6.0 x 10-10). All confirmed variants were biologically relevant, showing direct associations with 25(OH)D synthesis, clearance, or binding. Participants in the top quartile of genotype scores had 2.5-fold elevated odds of vitamin D insufficiency (P=2.3 x 10-48).

Interpretation: Variants near genes involved in synthesis, hydroxylation, and vitamin D transport influence vitamin D status. Genetic variation at these loci identifies individuals of European descent who have substantially elevated risk of vitamin D insufficiency.

Poster No: P042-F
Chromosomal conformation patterns govern the formation of conjoined genes in eukaryotic genomes.

Tulika Prakash Srivastava, Ritsuko Ozawa, Todd D. Taylor.
MetaSystems Research Team, Computational Systems Biology Research Group, RIKEN Advanced Science Institute (ASI), Yokohama 230-0045, Japan.

Conjoined genes (CGs) are the transcripts formed by joining at least one complete or partial exon from two distinct (child) genes which lie on the same chromosome, are in the same orientation, and sometimes independently encode for different proteins. Previously, we identified several hundred CGs in the human and other eukaryotic genomes, including mouse and fruit fly, and proposed their functional roles including formation of novel proteins and regulation of the expression of the child genes. We also found that the number of CGs conserved from human to other vertebrate genomes including chimpanzee, mouse, and dog is correlated with the order of complexity of the genomes. Three obvious questions that arise from these observations are the following: 1) are all the human child gene pairs resulting in the formation of CGs also conserved in other genomes, 2) is their arrangement in human syntenic in other genomes, and 3) do all those human child gene pairs which are syntenic in other genomes give rise to a CG in that genome? In the present analysis we found that 64% of the child genes are conserved from human to mouse. We also found that although the fraction of those human child gene pairs which are syntenic in mouse (60%) is similar to that of all the gene pairs in the human genome (68%), only a very small fraction (12.6%) of all known CGs in mouse were found in these human syntenic regions. Similar observations were made for the chimpanzee, dog and fruit fly genomes. Therefore, a pair of child genes giving rise to a CG in human may not necessarily result into a CG in other genomes in spite of their syntenic arrangement. Together, these observations indicate a selective conservation of human child genes along with increasing genomic complexity. Also, it is evident that the formation of CGs in each genome is a context dependent process governed by its own genomic requirements like genetic regulation and its overall complexity. Chromosomal folding patterns have been postulated to affect the formation of CGs. Our observations provide strong support for this hypothesis because, although the human

child gene pairs remain syntenic in lower organisms, they may not give rise to a CG in that genome due to a different conformation of its chromosomes.

Poster No: P057-F
Genome-wide association studies of hematological, biochemical and physiological traits in a Japanese population identified 49 novel loci.

[1,2]**Yukinori Okada**, [3]Yoichiro Kamatani, [2]Atushi Takahashi, [3]Koichi Matsuda, [4]Naoya Hosono, [2]Hiroko Ohmiya, [3]Yataro Daigo, [1]Kazuhiko Yamamoto, [4]Michiaki Kubo, [3]Yusuke Nakamura, and [2]Naoyuki Kamatani
[1]Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Japan, [2]Lab for Statistical Analysis, CGM, RIKEN, Japan, [3]Lab of Molecular Medicine, HGC, IMS-UT, Japan, [4]Lab for Genotyping Development, CGM, RIKEN, Japan

Recent development of genome-wide association studies (GWASs) led to the identification of numerous quantitative trait loci (QTL). Here, we report GWASs for hematological, biochemical and physiological traits in the Japanese populations. We enrolled around ~20,000 Japanese individuals under the support of BioBank Japan Project (http://biobankjp.org). Genotyping was performed using Illumina Human610-Quad BeadChip or HumanHap550v3 BeadChip. The stringent quality control filters were applied, including principal component analysis (PCA) and the estimation of identity-by-descent (IBD). Quantitative traits were adjusted with covariates (age, sex, smoking, drinking, etc.) and normalized as Z scores. Associations of the SNPs with the adjusted quantitative traits were evaluated using linear regression analysis. We identified 62 associations for 9 hematological traits (white blood cell count: WBC, neutrophil count, red blood cell count: RBC, hemoglobin concentration: Hb, hematocrit: Ht, mean corpuscular volume: MCV, mean corpuscular hemoglobin: MCH, mean corpuscular hemoglobin concentra¬tion: MCHC, and platelet count: PLT), 29 associations for 12 biochemical traits (urate: UA, high-density lipoprotein: HDL, triglyceride: TG, gamma glutamyl transferase: GGT, alkaline phosphatase: ALP, aspartate aminotransferase: AST, alanine aminotransferase: ALT, creatine kinase: CK, total protein: TP, albumin: ALB, blood urea nitrogen: BUN, and serum creatinine: sCr), and 8 associations for adult height at genome-wide significance levels (P < 5×10^–8). Of these, 49 associations were novel and 50 were the replications for the previous reports.

In addition to the identifications of the newly associated loci, our study provides the following topics. (i): Different genetic backgrounds of QTL among different populations. Although IGF1 is a well-known gene that plays an important role in skeletal development, the association of the variation in IGF1 with adult height had not been established in Europeans. Our study identified that the multiple SNPs were independently associated with adult height in IGF1 locus, and that their minor allele frequencies were high in Asian populations but quite low or monomorphic in Europeans. (ii): Most of the identified QTL included the genes that were functionally related with the traits, such as CSF3 gene as QTL for WBC and neutrophil count, ABO gene as QTL for RBC, Hb and Ht, and GGT1 gene as QTL for GGT. (iii): Even though a number of QTL were identified, most of the QTL would be still unidentified. We evaluated the combined effects of the height-associated loci incorporating a total of 51 loci that were previously or newly identified. However, the explained proportion of the variance of adult height was as low as 4.8%.

Our study would enhance our knowledge for QTL and motivate the accumulations of the further studies.

Poster No: P072-F
Digital gene expression data, cross-species conservation and noncoding RNA

[2]**Thérèse Commes**,[1,2]Nicolas Philippe, [2]Florence Ruffle, [2]Elias Bou-Samra, [3]Anthony Boureux, [1]Eric Rivals
[1]Laboratoire d'Informatique, de Robotique et de Microélectronique, UMR 5506 CNRS, équipe MAB, Université de Montpellier II, 161 rue Ada, 34392 Montpellier, France, [2]Groupe d'études des transcriptomes, Université de Montpellier II, Institut de Génétique Humaine, UPR1142-CNRS, 141, rue de la cardonille, 34095 Montpellier, France, [3]CRBM UMR5237, Université de Montpellier II - CNRS, 1919 route de Mende, 34095 Montpellier, France

Recently developed sequencing technologies offer massively parallel production of short reads and become the technology of choice for a variety of sequencing-based assays, including gene expression. Among them, digital gene expression analysis (DGE), which combines generation of short tag signatures for cellular transcripts with massively parallel sequencing, offers a large dynamic range to detect transcripts and is limited only by sequencing depth. As recently described (Philippe et al, 2009), tag signatures can easily be mapped to a reference genome and used to perform gene discovery. This procedure distinguishes between transcripts originating from both DNA strands and categorizes tags corresponding to protein coding gene (CDS and 3'UTR), antisense, intronic or intergenic transcribed regions. Here, we have applied an integrated bioinformatics approach to investigate tags'properties, including cross-species conservation, and the ability to reveal novel transcripts located outside the boundaries of known protein or RNA coding genes. We mapped the tags from a human DGE library obtained with Solexa sequencing, against the human, chimpanzee, and mouse genomic sequences. We considered the subset of uniquely mapped tags in the human genome, and given their genomic location, determined according to Ensembl if they fall within a region annotated by a gene (CDS, UTR and intron) or an intergenic region. We found that 76.4 % of the tags located in human also matched to the chimpanzee genome. The level of conservation between human and chimpanzee varied among annotation categories: 85 % of conserved tags in the CDS, 81 % in the UTR, 76% and 73% respectively in intron and intergenic regions. With the same procedure applied to human and mouse, we obtained 11% of conserved tags in the CDS, 7% in UTR, 1% and 3% respectively in intronic and intergenic regions. We analysed in depth the common CDS and UTR tags in human and mouse for their functional relevance: 90% of them correspond to orthologous genes with a common HUGO. We used DAVID database to extract biological features, the gene clustering revealed specific molecular functions belonging to transcription cofactor and regulator activity, nucleotide binding, ligase and proteine kinase, hormone receptor, histone methyltransferase or GTPase activity, and also important signaling pathways like WNT pathway. Indeed, intergenic transcription includes mainly new, non protein-coding RNAs (npcRNAs), which could represent an important class of regulatory molecules. By integrating also SAGE genie and RNA-seq expression data, we selected intergenic tags conserved across species and assayed experimentally the npcRNA transcriptome with Q-PCR. We validated 80% of the 32 tested biological cases. These results demonstrate that considering tag conservation helps to identify conserved genes and functions, which is of great relevance when investigating expressed tags located in intergenic regions.

POSTER ABSTRACTS

Poster Sessions

On each day of the meeting the following time periods have been specifically reserved for to viewing and discussing of posters.

Poster Session 1	19 May 2010 (Wednesday)	11.35 AM – 1.05 PM
Poster Session 2	20 May 2010 (Thursday)	11.25 AM – 12.55 PM
Poster Session 3	21 May 2010 (Friday)	12.00 NN – 1.30 PM

Poster Number

PO15 – W (Day)

Poster Board Number

W – 19 May 2010
T – 20 May 2010
W – 21 May 2010

All posters must remain on the poster boards for the duration. Posters should put up your poster no later than 4.00 pm on Tuesday 18 May 2010. HUGO will provide Velcro tape for hanging the posters on the boards. Each presenter is required to present at the time specified by their poster number. Presenters not at their posters on the specified date/time will not be eligible for the Poster Awards.

Poster No: P001-W
On the utility of biobanks linked to electronic medical records in genome-wide association studies

Logan Dumitrescu[1,2], Marylyn D. Ritchie[1,2], Joshua C. Denny[3,4], Jill M. Pulley[5], Melissa A. Basford[5], Andrea H. Ramirez[4], Daniel R. Masys[3,4], Jonathan L. Haines[1,2], Dan M. Roden[4,5,6], and Dana C. Crawford[1,2]
Departments of [1]Molecular Physiology and Biophysics, [2]Center for Human Genetics Research, [3]Biomedical Informatics, [4]Medicine, [5]Office of Personalized Medicine, [6]Pharmacology, Vanderbilt University, Nashville, TN 37232

Low levels of high-density lipoprotein cholesterol (HDL-C) are predictive of cardiovascular disease and myocardial infarction. Both genetics and the environment contribute to the variability of HDL-C trait distribution in the general population. To identify the genetic variants associated with HDL-C, candidate gene and genome-wide association studies (GWAS) have been performed in epidemiologic studies drawn from general population settings. Collectively, these studies identified more than ten genes or genomic regions associated with HDL-C levels in populations of European-descent, many of which are highly replicable in subsequent studies. An alternative strategy to the well-characterized epidemiologic study is the use of large DNA repositories or biobanks linked to electronic medical records (EMRs) as a source of data suitable for genome-wide association studies for gene discovery and replication. The advantages of biobanks are several-fold, including rapid accrual of samples, multiple phenotypes and traits linked to each DNA sample, and dense pharmacologic data for drug exposure assessment. To explore this alternative strategy, the National Human Genome Research Institute's electronic Medical Records and Genomics (eMERGE) Network aims to assess the utility of EMRs coupled to DNA repositories as a tool for GWAS and other genome studies. BioVU, the Vanderbilt DNA Databank, is one such repository of DNA samples extracted from discarded blood samples collected for routine clinical testing. These DNA samples are linked to a de-identified image of the EMR called the Synthetic Derivative. To date, BioVU contains >78,000 DNA samples in the repository. As a member of eMERGE, a subset of BioVU DNA samples was genotyped on the Illumina Human660W-Quadv1_A by the Center for Genotyping and Analysis at the Broad Institute. Natural language processing algorithms were developed to select DNA samples for genotyping from patients with normal electrocardiograms without evidence of cardiac disease, as the QRS duration (a trait of the electrocardiogram) was the primary trait for analysis. Data were cleaned using the quality control pipeline developed by the eMERGE Genomics

Working Group, and a total of 514,841 SNPs in 2,337 samples were available for the primary trait analysis. A subset of the genotyped samples also contained trait information on HDL-C. As a secondary analysis, single SNP tests of association for median HDL-C were performed on these 1,079 genotyped samples. In unadjusted linear regressions assuming an additive genetic model, transformed HDL-C was associated with CETP SNPs rs1532624 and rs1800775 at p-values (betas) 1.79x10-8 (0.083) and 2.62x10-8 (0.082), respectively. Two other CETP SNPs were also associated with HDL-C levels at a significance threshold of 10-7. Thus, our "top hits" in this EMR-based GWAS replicate the findings of several GWAS and candidate gene studies performed in epidemiologic cohorts for HDL-C, suggesting EMR-based DNA repositories will be useful for discovery and replication for a variety of algorithm defined phenotypes and traits linked to DNA samples.

Poster No: P012-W
Association of CD226 gene with systemic lupus erythematosus through impaired mRNA processing in T cells

Angelica M. Delgado-Vega*[1], **Marta E. Alarcón-Riquelme**[1], and **Serge V. Kozyrev**[1], Sara E. Löfgren*[1], Caroline J. Gallant[1], Elena Sánchez[2], Johan Frostegård[3], Lennart Truedsson[4], Bernardo A. Pons-Estel[5], Sandra D'Alfonso[6], Torsten Witte[7], Bernard Lauwerys[8], Emöke Endreffy[9], Laszlo Kovacs[9], Carlos Vasconcelos[10], Berta Martins[11], Javier Martin[2], *[1]Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden. [2]Instituto de Biomedicina y Parasitología López-Neyra, CSIC, Granada, Spain [3]Department of Medicine, Karolinska University Hospital, Huddinge, Stockholm, Sweden [4]Clinical Microbiology and Immunology, Lund University, Lund, Sweden [5]Department of Rheumatology, Sanatorio Parque, Rosario, Argentina [6]University of Eastern Piedmont, Novara, Italy [7]Hannover Medical School, Hannover, Germany [8]Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium [9]University of Szeged, Szeged, Hungary [10]Hospital Santo Antonio, Largo Abel Salazar, Porto, Portugal [11]ICBAS, Immunogenetic, Largo Abel Salazar, Porto, Portugal*

Objective.
The variant Gly307Ser/rs763361 of CD226 gene has recently been associated with a number of autoimmune diseases. We here report a fine mapping of the gene locus and the genetic association of several SNPs with systemic lupus erythematosus in Europeans, together with functional analyses to give a better understanding on the mechanisms behind the gene association and potential contribution of these variants.

Methods.
Genetic association of 12 SNPs in the CD226 gene was conducted using 1163 SLE patients and 1481 healthy controls from a European multicenter collection. Genotyping was performed using a GoldenGate Custom Genotyping Assay and a BeadXpress Reader from Illumina or Taqman assays and the statistical association analyzed with PLINK v1.07. Gene expression was analyzed by quantitative real-time PCR and SYBR Green for signal detection, using total RNA purified from PBMCs from healthy donors. Surface detection of the protein was performed by a three-color flow cytometry using leukocytes isolated from healthy donors, where total lymphocytes, CD3+CD4+, CD3+CD8+, CD3-CD56+, CD3-CD19+ and CD3+CD56+ cells were analyzed. Expression analysis of reporter plasmids including different alleles of the risk haplotypes of the CD226 3'-UTR region was assessed by transfection of HEK293 cells and dual luciferase assay.

Results.
A risk haplotype (P = 1.7 x10-7) was detected in the 3'-UTR region of the gene, and revealed rs727088 as the strongest associated variant (PCMH-adjusted = 0.0115). Expression analysis showed that the risk haplotype correlated with decreased levels of CD226 transcripts and protein levels in the surface of T helper cells (CD4+), Cytotoxic T cells (CD8+) and NK T cells (CD3+CD56+), but not NK cells (CD3-CD56+) or B cells (CD19+). Luciferase assays suggest that rs727088 is the main polymorphism responsible for altered gene expression.

Conclusion.
Our data does not support Gly307Ser as main functional variant within CD226 gene and indicates rs727088 located in the 3'UTR region as the potential causative SNP, by a mechanism that alters protein expression in T cells, potentially involving mRNA processing and/or stability.

Poster No: P013-W
Molecular genetic analysis of hereditary diseases in Bashkortostan Republic of Russia

E.K. Khusnutdinova, I.M. Khidiyatova, R.I. Khusainova, A.S. Karunas, V.L. Akhmetova, L.U. Dzhemileva, R.R. Valiev.
Institute of Biochemistry and Genetics, Ufa Science Center, RAS

Hereditary diseases are characterized by clinic polymorphism and genetic heterogeneity, that's why the decision of the most important problem of medical genetics is determination of population singularity and structural features of genes, determining development of these diseases. Here we report the main results of molecular-genetic analysis of Wilson disease (WD), Duchenne muscular dystrophy (DMD), congenital adrenal hyperplasia (CAH), Marfan syndrome (MFS) and non-syndromic hereditary deafness (NSHD) in Bashkortostan Republic.

WD is an autosomal recessive disorder characterized by toxic accumulation of copper in liver, brain and other organs. Using SSCP analysis followed by direct sequencing, among 38 unrelated patients we have characterized the molecular defect in 81.8% of WD chromosomes and identified 9 different mutations and 3 polymorphic variants. The common His1069Gln mutation was represented in 43.9% of WD chromosomes. Three novel mutations (p.Ala718Pro, p.Leu1057Pro and p.Lys1315_Arg1316delinsGlu) were revealed. Mutation p.Lys1315_Arg1316delinsGlu was detected in 18.2% of Tatar chromosomes and 22.2% of Bashkir chromosomes.

DMD is caused by mutations in the dystrophin gene consisting of 79 exons. We studied mutations spectrum of twenty exons (3, 4, 6, 8, 13, 17, 19, 42-53, 60) and promoter region of DMD gene in patients from 60 unrelated families. Exon deletions of the gene were revealed in 31.75% of families. We revealed also 2 new point mutations not described before and 1 polymorphism: p.Thr134ThrfsX7 (3.17%); p.Lys2210ArgfsX11 (1.6%); c.7728T>C (Asn2575Asn) (1.6%). CAH is a group of autosomal recessive disorders of adrenal steroid genesis in which 21-hydroxylase deficiency accounts for over 95% of cases. We studied 83 patients with CAH. Mutations of the CYP21A2 gene were revealed in 69.19% of the studied CAH-chromosomes. The mutations were distributed as follows: delA2/LGC (26.74%), p.Arg356Trp (16.28%), I2splice (11.05%), p.Ile172Asn (7.56%), p.Gln318X (5.23%), p.Pro30Leu (1.16%) and p.Val.281Leu (1.16%). We found 7 patients who carried 3 mutations, two from which formed a cluster: p.Gln318X+p.Arg356Thr (N=5), p.Ile172Asn+p.Gln318X (N=1), p.Ile172Asn+Arg356Thr (N=1). This complex of alleles probably resulted from large conversions or multiple mutations events.

MFS is an inherited autosomal dominant connective tissue disorder. We analyzed 30 exons of FBN1 gene and 4 exons of TGFBR2 gene in 80 unrelated patients with MFS. We identified two missense mutations (p.Gly1176Tyr and p.Cys2489Tyr) which affect cbEGF-like motifs of fibrillin-1 protein in two patients with classical MFS. We also found 9 polymorphisms both in coding and non coding regions of FBN1 gene, five of them are not previously described. One novel mutation (p.T223M) has been found in TGFBR2 gene in two unrelated patients.

NSHD is one of the most common disorders of sensorineural function. We analyzed the GJB2 gene in 100 unrelated patients. Four different frequent mutations 35delG (44% of GJB2 alleles), 312del14 (4%), del167T (3%) and 235delC (1.5%) were observed. The 35delG mutation of the GJB2 gene is responsible for deafness in 53% of the independent Russian families tested.
These findings became a base for optimal scheme of DNA-diagnostics development of these hereditary diseases in the Volga-Ural region and ascertainment of the mutations origin, leading to these diseases.

Poster No: P015-T
Effect of Tumour Necrosis Factor Gene Polymorphism on Risk of Metabolic Syndrome in Young Obese Males

SH Yap, Alicia PL Chong, M Uttamchandani, Rita YY Yong, Eric PH Yap
Defence Medical and Environmental Research Institute, DSO National Laboratories, Singapore

Obesity is on the rise worldwide and is associated with the development of significant medical illnesses and consequences. One example is the metabolic syndrome or insulin resistance, a condition that consists of multiple, interrelated risk factors of metabolic origin strongly associated with risk of heart disease and type 2 diabetes. Adipose cell enlargement in obesity induces a proinflammatory state in the tissue with the infiltration of macrophages, which increases the level of cytokines such as tumour necrosis factor (TNF). The increased level of cytokines promotes ectopic lipid storage and induces insulin resistance locally and in other peripheral tissues.

In our study of the health profile of 713 local Chinese males aged 18 to 21 years old, metabolic syndrome was found to be prevalent (25.2%) in the obese individuals (n=377, body mass index, BMI>28kg/m^2), while absent in the normal weight individuals (n=336, BMI<23kg/m^2). Hence, we sought to find out whether associations exist between genetic variation in cytokine genes and metabolic syndrome or any of its five components in young obese males. DNA extracted from venous blood of the 377 obese were genotyped for five SNPs located near 5' of *TNF* (rs1799964, rs1800630, rs1799724, rs1800629, rs361525) on the Sequenom MassARRAY platform and analysed for association with metabolic syndrome and its five components by logistic regression, assuming an additive model. rs1800629 allele G was associated with risk of metabolic syndrome in the obese (P=2.23 x 10^{-3}, OR=3.263, 95%CI=1.529 to 6.964). This association remained significant after adjusting for BMI (P=2.12 x 10^{-3}, OR=3.314, 95%CI=1.543 to 7.116). rs1800629 allele G was also associated with larger waist circumference (one of the components of metabolic syndrome) after adjusting for BMI (increase by 1.3cm per allele, P=3.19 x 10^{-2}, 95%CI=0.112 to 2.479). Both BMI and TNF rs1800629 allele G accounted for about 5.8% of the variation of the outcome of having metabolic syndrome in the obese. The results showed that genetic variation in TNF plays a part in explaining the susceptibility to having metabolic syndrome in young obese males.

Poster No: P075-F
From Research to Clinic : the clinician-patient relationship as a key of the high throughput technology transfer into clinic

[1]**Sophie Julia**, [1]Alexandra Soulier, [2]Anne Cambon-Thomsen
[1]INSERM, France; [2]CNRS, FRANCE

Genetic testing is becoming a growing part of healthcare services. Seeing that genetic diseases are often molecularly and clinically highly heterogeneous, their clinical investigation requires sophisticated diagnostic tools. Rapid developments of high throughput technology (HTT) provide such opportunities but sequencing genes generates complex and heterogeneous information with different levels of clinical relevance. Some ethical issues specifically challenge the transfer of these improved technologies from the research laboratory to their clinical applications. Dealing with unexpected health related information as well as communicating data that lack widely accepted interpretation among clinicians are especially of relevance.

Our concern is about the evaluation that estimates the clinical relevance of genetic information generated by new technologies. The process that allows validation of technologies and the interpretation of the results they provide is usually formalized in the ACCE framework – which considers successively Analytic validity; Clinical validity; Clinical utility; and Ethical, legal, and social implications of innovative technologies. The place allocated to the patient/clinician relation is not clear in this assessment. But, as emerging technologies provide complex data that deliver different levels of (un-)certainty and thus may affect the clinician's ability to communicate with patients, it seems that the patient/clinician relation should be thoroughly studied and become a relevant parameter in the evaluation process.

ACCE framework does not specify this dimension. This paper therefore aims to illustrate by some case studies how such a level of complexity affects the relation between patient and clinician and to provide some insights about the way this dimension could be integrated into the technology validation process. Using qualitative focus groups and questioning them about these emerging technologies, we will analyze results concerning how various publics perceive the increasing complexity of data and its impact over the communication between clinicians and patients.

Poster No: P076-W
Mutational Spectrum of CFTR gene in Korean Patients with Cystic Fibrosis

[1]**Haiyoung Jung**, [2]Won-Jung Koh, [3]Kang-Mo Ahn, [4]Sang-Il Lee, [5]Jeong-Ho Kim, [6]Jae Sung Ko, [6]Jeong Kee Seo, [7]Seung-Ick Cha, [8]Eun-Sil Lee, [1]Jong-Won Kim, and [1]Chang-Seok Ki
[1]Department of Laboratory Medicine & Genetics, [2]Division of Pulmonary and Critical Care Medicine, [3]Department of Pediatrics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 135-710, Korea; [4]Department Pediatrics, Patima Hospital, Daegu, Korea; [5]Department of Laboratory Medicine, Yongdong Severance Hospital, Yonsei University College of Medicine, Seoul; [6]Department of Pediatrics, Seoul National University College of Medicine, Seoul,

Korea; [7]Department of Internal Medicine, Kyungpook National University School of Medicine, [8]Daegu, Korea; Department of Pediatrics, College of Medicine, Yeungnam University, Daegu, Korea

Introduction: Cystic fibrosis (CF) is one of the most common inherited disorders among Caucasians. The common mutations of CFTR (Cystic Fibrosis Transmembrane conductance Regulator) gene has been well established among Caucasian populations. However, little has been reported on the spectrum of mutations in Koreans. In this study, we summarized the results of CFTR mutational spectrum from Korean CF patients.

Material and method: We have analyzed the complete coding regions and flanking intronic sequences of the CFTR gene among seven unrelated CF patients (2 males, 6 females). To rule-out large deletions or gene rearrangements, the multiplex ligation-dependent probe amplification (MLPA) analysis was performed in one patient detected with only one known mutation.

Results: Twelve CFTR mutations were identified from seven patients including Q98R, Q220X, D979A, Q1291X, Q1352H, 579+5G>A, IVS8-T5, 1766+2T>C, 2052delA, 2623-?_2751+?del, 3272-26A>G, and 3908insA. Q98R mutation was the only recurrently observed mutation with a frequency of 21.4% (3/14 alleles). No delta F508 mutation, which is known to be the most common mutation among Caucasians, was detected in Korean patients in this study.

Conclusion: This study is the first to collectively summarize the CFTR mutational spectrum of Korean CF patients. Detection of no delta F508 mutation and a heterogeneous spectrum of CFTR mutations differed from the common CFTR mutational spectrum in Caucasians. Therefore, the genetic analysis of CFTR gene among Korean CF patients necessitates the sequencing of the whole gene followed by the MLPA analysis.

Poster No: P077-T
ABCA13: The Story of a Gene Symbol
Ruth Seal, Susan Gordon, Michael Lush, Matt Wright, Elspeth Bruford
HUGO Gene Nomenclature Committee (HGNC), European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD, United Kingdom

The HUGO Gene Nomenclature Committee (HGNC) aims to approve a gene symbol and name for every human gene. Standardisation of gene symbols is important as it allows researchers to refer to the same gene without ambiguity and facilitates electronic data retrieval. The primary tenet of the HGNC is that every approved gene symbol must be unique. Gene symbols should also be acceptable to researchers to ensure their widespread use, and should be based on structure, function or homology wherever possible. The HGNC encourages the development of a common root symbol for members of a gene family, with a hierarchical numbering system to distinguish the individual members, as this is an efficient way to name large numbers of related genes and makes each family member instantly recognisable. We provide individual web pages on our site for many established gene families and have over one hundred specialist advisors that help us to accurately maintain these families.

Here we present the story of one particular gene symbol, ABCA13, to demonstrate the steps taken when naming a gene. We explain how and why this gene received its current symbol and name (ATP-binding cassette, sub-family A (ABC1), member 13) and how the symbol has subsequently been used. The story begins with the development of the ABC (ATP-binding cassette) gene superfamily and sub-family nomenclature scheme, following in depth discussions between the HGNC and the ABC research community. It then follows the identification and naming of ABCA13 as part of this family, its appearance in the biomedical literature and databases, its adoption for the mouse and rat orthologs as Abca13, and its subsequent breakthrough into the international media.

For further information on human gene nomenclature, please email us at hgnc@genenames.org, or visit http://www.genenames.org/. The work of the HGNC is supported by the NHGRI and the Wellcome Trust.

Poster No: P078-W
Prediction of potential miRNA function modulators through the analysis of miRNA target variability in large transcriptome datasets.

Josep M Mercader[1], Montserrat Puiggròs[1], David Torrents[1,2]
[1]*Life Sciences Department, Barcelona Supercomputing Center, Jordi Girona 31, 08034 Barcelona, Spain* [2]*Institució Catalana de Recerca i Estudis Avançats (ICREA) Pg. Lluís Companys 23, 08010 Barcelona, Spain*

miRNAs are small non-coding RNAs involved in the fine tuning of gene expression through base pairing with the 3'UTRs of their target genes. It has been shown that genes that contain target sites for miRNAs have higher brain expression variability among individuals than those that are not targeted by miRNAs (Zhang and Su 2008); and that high miRNA activity is correlated with lower expression of their targets (Sood et al 2006). We can therefore hypothesize that the expression of those genes involved in miRNA activity will also correlate with the expression of miRNA target genes. For example, mice with Dgcr8 haploinsuficiency, a key gene involved in miRNA biogenesis, have shown to have a higher than expected proportion of upregulated genes among all miRNA targets in brain (Stark et al 2008).

In order to identify novel candidate genes involved in the function of miRNA (such as DICER, DGCR8 and FMR1) we have examined the expression variability of miRNA target genes in three different large transcriptome datasets from 427 liver human samples (Schadt et al 2008), 210 human lymphoblastoid HapMap samples , and 199 cortex human samples (Myers et al 2007). Briefly, for each individual, we assessed the expression ratio between genes predicted to be miRNA target and those that are not. We next analyzed the correlation between this ratio, taken as a molecular phenotypic trait, and the expression levels for each of the genes tested within each transcriptome.

We found that: (i) the expression variability of targets compared to non-target genes observed among individuals also occurs among tissues. Particularly, in both HapMap and cortex samples, miRNA targets show higher expression than the non miRNA targets, whereas in liver, miRNA targets were downregulated when compared to the non targets. This suggests that miRNA repression might be higher in the liver, compared to the other tissues. (ii) We identified genes whose expression correlates with miRNA target/non target ratio in the three datasets used. which might be novel genes potentially involved in the silencing function of miRNA. In agreement with this, we have found among our best candidates, the FMR1 gene, which has been proved to be associated with miRNA activity (combined pvalue = 8.2 x 10-72). In addition, within the same set we also found STAU1 (Staufen), which has been shown to be involved in miRNA related p-body assembly (combined p value across three datasets = 5,2 x 10-66).

Our results strongly suggest that this approach might be useful to identify novel candidate genes involved in the silencing function of miRNA.

Poster No: P080-F
Alu motifs in exonised transcripts are probable targets for miRNA in response to heat shock

Rajesh Pandey[1], Amit K. Mandal[2], Vineet Jha[2], Mohit Swarnkar[1] and Mitali Mukerji[1,2].
[1]*Functional Genomics Unit,* [2]*G.N. Ramachandran Knowledge Centre for Genome Informatics, Institute of Genomics and Integrative Biology, CSIR, Mall Road, Delhi, INDIA.*

In recent times, Alu repeats, the most abundant repetitive elements in the human genome (~11%) and miRNAs (~1000 discovered miRNAs), small RNAs that alter gene expression at the post-transcriptional level have been the focus of research, owing to their multi-dimensional regulatory role at the genome wide level. miRNA mostly target 3' UTR regions which are also most preferred sites for Alu exonization, present in >5% of all human 3' UTRs. Role of Alu RNA in repressing transcriptional initiation has been demonstrated in response to heat shock. However, post transcriptional regulation mediated by miRNA targeting Alus has not been studied. We attempted to explore the role of miRNA targeting exonised Alus at genome wide scale in response to heat shock in HeLa cell line. Total RNA was isolated from HeLa cells following recovery for 2 hours, after heat shock treatment for 30 mins at 45°C. mRNA expression profiling was carried out using Illumina WG-6 v3.0 48K Beadchip and miRNA using Exiqon dual-channel arrays. The mRNA expression analysis was carried out using Illumina Beadstudio. For miRNA expression analysis we have used Gene-specific dye bias normalization, using GASSCO method from Bioconductor package. Validation of differentially expressed miRNAs and

mRNAs was carried out by Q-PCR using SYBR. The 3' UTR targets for miRNA was predicted using a consensus of 2 softwares – miRanda and TargetScan, and only those miRNAs were taken for validation which had free energy less than -20. Alu repeats were identified using repeat masker. Amongst the 3480 differentially expressed genes, 1341 were upregulated and 2439 were downregulated. 7 out of 12 differentially expressed genes were validated by Q-PCR. 15 miRNAs were observed to be differentially expressed, out of which 3 were validated by Q-PCR, using SNORD non-coding RNAs as internal control. We found 9 genes to be negatively correlated with the expression of the validated miRNAs. Interestingly, 6 out of 9 genes have miRNA targets only in the Alu repeat present in 3'UTR of Alu exonised transcripts. These results suggest that miRNA mediated regulation of Alu exonized transcripts may add to the repertoire of transcriptomic diversity being contributed by Alu repeats. This regulation may be cell line and condition specific. Further experimental validations of these results are in progress in different cell lines.

Poster No: P086-T

Genome-wide association study detects a novel profile of genetic variations associated with individual response to opioid therapy for cancer pain

Antonella Galvan [1], Pål Klepstad [2], Frank Skorpen [2], Stein Kaasa [2], Augusto Caraceni [1], Tommaso A. Dragani [1] in collaboration with the European Pharmacogenetic Opioid Study and the European Palliative Care Research Collaborative
[1] Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy [2] Medical Faculty, Norwegian University of Technology and Science, Trondheim

Patients treated for cancer pain experience different responses to opioid drugs and different side effects for reasons that are mostly unknown. We tested the hypothesis that genetic variations may be predictive of individual response to opioids by genome-wide association study (GWAS) using single nucleotide polymorphism (SNP)-arrays in patients who underwent opioid therapy for cancer pain. From the total study sample of 2294 patients we tested one million SNPs in patients selected for extremely poor (pain relief <40%; n=145) or good (pain relief >90%; n=293) responses to opioid therapy for cancer pain, using a DNA pooling approach. A first selection of SNPs based on allelic hybridization intensity detected 21,538 SNPs at a nominal p<1e-07. Among these SNPs, differences in chromosome counts between poor and good responders highlighted 72 SNPs (p<1e-05), which were further processed for genotyping in individual samples, resulting in 59 successfully genotyped SNPs. Almost all these SNPs (i.e., 56 of 59) showed statistically significant allelic association with opioid responder status (p<0.05), thus confirming the reliability of the DNA pooling approach in GWAS. Many of the associated SNPs mapped near or within genes that might be involved in the neural control of pain. Among putative candidate genes, gamma-aminobutyric acid (GABA) A receptor, beta 2 (GABRB2) gene encodes for the receptor of GABA, which mediates neuronal inhibition by binding to the GABA/benzodiazepine receptor; ankyrin 3, node of Ranvier (ankyrin G) (ANK3) gene is a membrane-cytoskeleton linker that may participate in the maintenance/targeting of ion channels and cell adhesion molecules at the nodes of Ranvier and axonal initial segments; spondin 1, extracellular matrix protein (SPON1) gene encodes for a cell adhesion protein that promotes the attachment of spinal cord and sensory neuron cells; solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 6 (VGLUT2) (SLC17A6) gene mediates the uptake of glutamate into synaptic vesicles at presynaptic nerve terminals of excitatory neural cells. Overall, our results that deserve validation in independent groups of patients represent new findings that can direct further research to genetic variations hitherto not known to be predictive of individual response to opioids.

Poster No: P087-W

Mapping of two large Tunisian families with retinitis pigmentosa to the RP25 locus; identification of mutations in EYS and characterization of the canine ortholog

Béatrice Bocquet[1*], Lorenne Robert[1], Estèle Lafont[1], Mohamed Ksantini[1], Saoussen Jilani[1], Wassim Kharrat[2], Saber Masmoudi[3], Emmanuelle Genin[4], Jamel Feki[2], Dominique Hillaire[1], Vasiliki Kalatzis[1], Christian P Hamel[1]
[1]INSERM U583, Institute for Neurosciences of Montpellier, Montpellier, France [2]Service d'Ophtalmologie, CHU Habib Bourguiba, Sfax, Tunisia [3]Laboratoire de Génétique Moléculaire Humaine, Faculté de Médecine, Sfax, Tunisia [4]INSERM U794, Fondation Jean Dausset/CEPH, Paris, France

Purpose: i) To identify the causative gene in two large Tunisian families affected with autosomal recessive retinitis pigmentosa (RP) and linked to the RP25 locus. ii) To characterize the canine ortholog of EYS, mapping within RP25.

Methods: Families were genotyped with microsatellite markers and SNP microarrays, and mapped using the Superlink and Merlin softwares. The 44 EYS exons were screened by PCR sequencing using primers designed in the intronic flanking sequences. Exonic rearrangements were searched by QM-PSF with fluorescent exon-specific primers. Total dog retinal RNA was extracted and subjected to RT-PCR to amplify the canine EYS cDNA.

Results: Both families were linked to the RP25 locus ($Z = +6.60$ at $\theta = 0$ for D6S257 for family A, $Z > +3.75$ for multiples SNPs for family B), which contains the recently identified EYS gene. In family A, 15/18 patients were homozygous for a frameshift mutation (c.5928-2A>G) leading to p.Arg1976ArgfsX13. In family B, 8/9 patients carried a homozygous exon 12 deletion leading to p.Cys590TyrfsX4. Both mutations cause a severe loss of the C-terminal part of the protein. Yet, 3/18 and 1/9 patients in families A and B, respectively, were heterozygous for the causative mutations. QM-PSF revealed an exon 34 duplication segregating in trans in the 4 heterozygous patients. However, the duplication was also present in 14% of the Tunisian population as well as in an unaffected carrier of family B, thus excluding it as a causative mutation. In parallel, as a murine Eys ortholog does not exist, we sequenced the cDNA of the canine EYS ortholog as a first step towards the identification of an animal model for EYS-linked RP. By sequence alignments we determined that canine EYS comprises 43 coding exons and encodes a 3208 aa protein. Furthermore, canine EYS contains 3 additional exons that are not present in the human ortholog. Overall, the human and canine proteins show 75% homology for the regions common to both species.

Conclusion: Heterozygous patients in the two families studied may carry an intronic EYS mutation causing exon skipping. We are currently RT-PCR-amplifying RNA from the lymphocytes of patients to identify aberrant transcripts. The identification of an EYS mutation in a dog affected with retinal progressive atrophy would provide a useful model for establishing future therapies for retinitis pigmentosa.
Acknowledgments: We thank the family members. This work was funded by private foundations (“Fondation des Aveugles et Handicapés Visuels de France”, “France Choroïdérémie”, “Information Recherche sur la Rétinite Pigmentaire”, Retina France, “SOS Rétinite” and UNADEV).

Poster No: P092-F
Chromosomal translocation in a patient with absence of breasts

[1,2,3]**Surasawadee Ausavarat,** [2,3]Siraprapa Tongkobpetch, [4]Verayuth Praphanphoj, [5]Charan Mahatumarat, [5]Nond Rojvachiranonda, [6]Thiti Snabboon, [7]Thomas C. Markello, [7]William A. Gahl, [2,3]Kanya Suphapeetiporn, [2,3]Vorasuk Shotelersuk
[1]Interdepartment of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University, Bangkok, 10330, Thailand [2]Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, 10330, Thailand [3]Molecular Genetics Diagnostic Center, King Chulalongkorn Memorial Hospital, Thai Red Cross, Bangkok 10330, Thailand [4]Center for Medical Genetics Research, Rajanukul Institute, Bangkok, 10400, Thailand [5]Department of Surgery, Faculty of Medicine, Chulalongkorn University, Bangkok, 10330, Thailand [6]Department of Internal Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand [7]Section on Human Biochemical Genetics, Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, 20892, USA

Although excellent progress has recently been made in defining the signaling pathways involved in the mouse mammary gland development, current knowledge about human mammary gland development is very restricted and still has to be elucidated. We identified an 18-year-old Thai woman who presented with absence of breasts (amastia), ectodermal dysplasia, and absence of the left kidney. The etiology of the disease is unknown. She was found to have a balanced chromosomal translocation of t(1;20)(p34.1;q13.13) inherited from her unaffected mother. We hypothesized that a gene was interrupted by one of the breakpoints which caused the phenotype. Using a combination of molecular genetics and cytogenetics experiments, we have demonstrated that a gene is disrupted on chromosome 1 of our patient. Measurements of its RNA and protein levels in her lymphoblastoid cell lines are severely deficient. Investigation of her DNA copy number using high-density SNP genotyping array showed no apparent deletions or duplications. In addition, sequencing of entire coding regions did not identify definite pathogenic mutation. However, microsatellite analysis showed that the proband inherited a paternal chromosome different from that of her unaffected siblings. In conclusion, we present

a new candidate involved in the development of breasts and nipples in humans. Its aberration may be responsible for amastia inherited in an autosomal recessive manner. This finding could provide new insights into developmental biology of human mammary glands and malignancy of breast cancer.

Poster No: P093-W
Systematic Building of Multiple Protein Alignments for Variant Interpretation

André Blavier, Audrey Deforche
Interactive Biosoftware, Rouen, France - European Union

The Alamut software is a decision-support system for mutation interpretation in medical molecular genetics. Alamut brings together molecular data and prediction methods for variant interpretation inside a graphical environment. A multiple protein alignment (MPA) of orthologues is included for each protein-coding human gene available in Alamut's database, bringing useful conservation information for missense variant interpretation.

Currently, the sets of orthologous sequences used to build the alignments are drawn from the Ensembl Compara database. However, since orthologue collection and alignment by Ensembl Compara are the result of a purely automatic process, a number of MPAs don't reach the quality needed for variant interpretation. On the other hand, manually building accurate alignments is very time-consuming. This is why we have sought to define a semi-automatic procedure for orthologue MPA construction, described here.

We have first sought how to quantify alignment quality. Based on work published by Tavtigian et al. (2008, 2009) and on recommendations published on the SIFT web site (Sorting Intolerant From Tolerant, a mutation classification system), we have chosen two criteria to assess alignment quality: average number of substitutions per position, and median of information content in alignments, a measure derived from Shannon's information theory. Correct alignments should contain on average three substitutions per position, and the median information content should be less than or equal to 3.25. We use the Protpars (Protein Sequence Parsimony Method) program from the PHYLIP package to compute the number of substitutions per position, and the Alpro program of the DELILA package to calculate information content.

New alignments are built as follows. Orthologous sequences are searched with the BlastP program, first against the Uniprot/Swissprot database. If sequences of distant species have not been found, BlastP is then run against the Refseq database, and finally against the NCBI non-redundant protein sequence database, if needed. The set of orthologues is then filtered manually, based on sequence length, identity with the human sequence, and available annotations. Selected sequences are then aligned with the M-Coffee program. M-coffee is a meta-aligner, combining the results of a number of multiple sequence alignment programs to construct a consensus alignment. If the alignment does not satisfy the above quality criteria, sequences creating large gaps are removed, and new sequences are added if needed to raise information content. As a final step, alignments are manually optimized with an alignment editor like JalView.

While still needing manual phases, this procedure allows us to create correct alignments for variant interpretation in a manageable time frame. We anticipate to quickly build new alignments for approximately 50 of the genes most frequently analyzed with Alamut. These alignments are expected to provide better interpretation quality, either through human analysis or through variant scoring systems like SIFT, PolyPhen, or AlignGVGD.

Poster No: P094-T
Genetic polymorphisms of the Paraoxonase 1 (PON1) gene: Association between T(-107)C or I102V polymorphisms and breast cancer risk among Malaysian women.

[1]**Rakesh Naidu,** [2]Yip Cheng Har, [2]Nur Aishah Mohd Taib
[1]*School of Medicine and Health Sciences, Monash University Sunway Campus, Jalan Lagoon Selatan, 46150 Bandar Sunway, Selangor Darul Ehsan, Malaysia,* [2]*Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia*

The PON1 gene, located on the long arm of chromosome 7q21.3-22.1, is polymorphic and the expression levels vary widely in human populations. Its ability to detoxify carcinogenic oxidative stress products has led investigators to hypothesize that PON1 polymorphisms might contribute to the increased risk of cancer. The purpose of the present study was to evaluate the association between the PON1 T(-107)C promoter or exon G) polymorphisms and breast cancer risk in a hospital-based Malaysian→4 I102V (A population. The distribution of genotype and allele frequencies of these polymorphic variants will be determined in breast cancer patients and normal healthy individuals. Genotyping of these polymorphisms was performed using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method. Peripheral blood samples were collected from 387 breast cancer patients and 252 normal and healthy women who had no history of any malignancy. The CC, CT and TT genotypes, and T allele of T(-107)C polymorphism were detected in 21.4%, 49.4%, 29.2% and 53.9% of breast cancer cases, respectively. The genotype (P<0.0001) and allele (P<0.0001) frequencies were significantly different between the breast cancer cases and normal individuals. The AA and AG genotypes of I102V polymorphism were detected in 96.1% and 3.9% of the breast cancer patients but GG genotype was not found in the population. AG and GG genotypes were not noted among the control subjects. The G allele was found in 2% of the patients and none in the normal women. The adjusted odds ratio was calculated using the multivariate logistic regression method with an adjustment for age and ethnicity to determine the correlation between the genotypes or alleles of T(-107)C or I102V polymorphisms and breast cancer risk. Women who were TT homozygotes (OR=2.39; 95% CI, 1.50-3.80), CT heterozygotes (OR=2.02; 95% CI, 1.36-3.00) or carriers of T allele (OR=1.61; 95%, 1.27-2.03) were associated with significant breast cancer risk. The risk of breast cancer associated with I102V was not determined due to the absence of the polymorphic variant in the population. In conclusion, our findings suggest that the polymorphic variant of T(-107)C polymorphism might be a useful genetic marker to identify women who are at greater risk of developing breast cancer in a hospital-based Malaysian population. Our data also showed that I102V polymorphism may not be common in our population.

Poster No: P095-T

Association between Cys311Ser or Ala148Gly polymorphisms of the Paraoxonase 2 (PON2) gene and breast cancer risk among Malaysian women.

[1]**Rakesh Naidu**, [2]Yip Cheng Har, [2]Nur Aishah Mohd Taib

[1]School of Medicine and Health Sciences, Monash University Sunway Campus, Jalan Lagoon Selatan, 46150 Bandar Sunway, Selangor Darul Ehsan, Malaysia, [2]Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

The PON2 gene is located adjacent to PON1 on the long arm of chromosome 7q21.3–22.1. Like PON1, PON2 is also polymorphic and it was hypothesized that the structurally-related PON2 has a similar function. The enzyme might be involved in the prevention of oxidative damage by detoxification of carcinogenic oxidative stress products. Two common polymorphisms have been identified at codon 311 (Cys311Ser) and 148 (Ala148Gly) of PON2 gene and these polymorphisms might contribute to increased risk of cancer. The present study was aimed to investigate whether Cys311Ser and Ala148Gly polymorphic variants of the PON2 gene are associated with breast cancer risk in a hospital-based Malaysian population. The genotype and allele frequencies of these polymorphims will be determined in breast cancer patients and normal healthy individuals. These polymorphisms were genotyped using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method. Peripheral blood samples were collected from 387 breast cancer patients and 252 normal and healthy women who had no history of any malignancy. The CC, CS and SS genotypes, and S allele of Cys311Ser polymorphism were detected in 11.1%, 33.3%, 55.6% and 72.2% of breast cancer cases, respectively. The genotype (P=0.023) and allele (P=0.01) frequencies of Cys311Ser polymorphism were significantly different between the breast cancer cases and normal individuals. However, the distribution of genotype (P=0.253) and allele (P=0.0930) frequencies of Ala148Gly polymorphism showed lack of statistical significance. The CC, CG and GG genotypes, and G allele of Ala148Gly polymorphism were detected in 53.5%, 36.2%, 10.3% and 28.4% of breast cancer cases, respectively. The adjusted odds ratio was calculated using the multivariate logistic regression method with an adjustment for age and ethnicity to determine the correlation between the genotypes or alleles of Cys311Ser or Ala148Gly polymorphisms and breast cancer risk. Women who were SS homozygotes (OR=1.56; 95% CI, 0.92-2.66) and CS heterozygotes (OR=1.00; 95% CI, 0.58-1.73) were not associated but carriers of S allele (OR=1.38; 95%, 1.08-1.77) were associated with increased risk of breast cancer. Women who were heterozygous (OR=1.22; 95% CI, 0.85-1.75) or homozygous (ORj=1.53; 95% CI, 0.82-2.86) for G allele, and carriers of G allele (OR=1.26; 95% CI, 0.96-

1.64) were not associated with breast cancer risk. In conclusion, our findings suggest that women who are carriers of S allele might be at higher risk of developing breast cancer but the polymorphic variant of Ala148Gly polymorphism might not be a suitable marker for breast cancer susceptibility.

Poster No: P096-F
Genome/chromosome instability during intrauterine development and cigarette smoking

[1,2]**Svetlana G. Vorsanova**, [1,2]Ivan Y.Iourov, [1,2]Irina A. Demidova, [1,2]Alphia K. Beresheva, [1,2]Oxana S. Kurinnaya, [1]Alexei D. Kolotii, [1]Elena A. Kirillova, [1,2]Victor S. Kravets, [1,2]Viktor V. Monakhov, [1]Ekaterina A Saprina, [1]Ilia V.Soloviev, [1,2]Yuri B. Yurov
[1]Institute of Pediatrics and Children Surgery, Rosmedtechnologii, Moscow, Russia; [2]Mental Health Research Center, RAMS, Moscow, Russia

Cigarette smoking has negative environmental impact on genome integrity and may lead to DNA damage, chromosome/genome rearrangements and aneuploidy. To test the influence of tobacco smoking on genome stability during human intrauterine development we have analyzed the rate of aneuploidy (chromosomal biomarker of genomic instability) in spontaneous abortions (SA). Chorionic villi of consecutive cases 600 SA of 5-15 weeks gestation from women aged from 16 to 47 years (mean 30 years) were processed for molecular cytogenetic analysis. Interphase multiprobe FISH with DNA probes for chromosomes 1, 9, 13/21, 14/22, 15, 16, 18, X and Y was applied. Chromosome abnormalities (autosomal and gonosomal aneuploidy, polyploidy, multiple aneuploidy and polyploidy and chimerism) were found in 303 cases from 600 analyzed (50.0%). Chromosomal mosaicism due mitotic errors was detected in 157 cases from 600 analyzed (26.2%). Among 463 non-smoking females chromosomal pathology was detected in 237 cases (51.2%) including 122 mosaic cases (26.3%). Among 137 smoking females (22.8% of analyzed cases) chromosomal pathology was detected in 65 cases (47.4%) including 35 cases with mosaicism (25.5%). Therefore, the incidence of regular and mosaic forms of chromosomal abnormalities was similar in SA affected and non-affected by tobacco smoke. However, the level of stochastic (or spontaneous) aneuploidy per individual chromosome pair not involved in aneuploidy was significantly increased in SA affected by tobacco smoking (1.4-2.1%) in comparison to non-affected (0.3-1.2%). We conclude that developmental mitotic instability frequently associates with mosaic chromosome/genome instability and abnormal intrauterine embryo development. Cigarette smoke has negative environmental impact on spontaneous level of mitotic chromosome instability. Supported by Philip Morris USA Inc.

Poster No: P097-W
Genome-wide association analysis of immune-related diseases in China

[1]**Xun Chu**, [1]Wei Huang, [2]Xuejun Zhang, [3]Furen Zhang, [1]Yi Wang, [1]Jinxiu Shi
[1]Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China [2]Institute of Dermatology and Department of Dermatology at No.1 Hospital, Anhui Medical University, Hefei, Anhui, China [3]Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Science, Jinan, Shandong, China

Immune-related diseases consist of a range of diseases, which include autoimmune and inflammatory disorders. Autoimmune disorders are initiated by damage to tissues and organs that arises from the response to self-antigens. By contrast, inflammatory disorders result from an excessive inflammatory response that is more harmful to host tissue than to exogenous antigens. Genetic predisposition is a major factor in susceptibility of immune-related diseases, and genome-wide association studies (GWAS) of cohort samples offer an opportunity to dissect the susceptibility genes (loci) of such diseases.

We have performed series of GWAS of immune-related diseases, including psoriasis, systemic lupus Erythematosus (SLE) and leprosy in Chinese populations, and obtained important results. We carried out the genome-wide association analysis using Illumina Human 610-Quad BeadChips in 1,139 psoriasis cases vs. 1,132 controls, 1,047 SLE cases vs. 1,205 controls and 706 leprosy patients vs. 1,225 controls.

In the replication stage of psoriasis, we took 64 top SNPS from GWAS for replication in two independent sample of 5,182 cases and 6,516 controls of Chinese Han ancestry, and 539 cases and 824 controls of Chinese Uygur ancestry. In addition to the strong replication for two known susceptibility loci MHC and IL12B, we indentified a new susceptibility locus within the LCE cluster on 1q21. Our result was published in Nature Genetics in February 2009.

In the replication stage of SLE, 78 SNPs were replicated in two additional Chinese Han cohorts (total 3,152 cases and 7,050 controls). In this study, eight genes were validated as SLE susceptibility genes in Chinese Han population (MHC, BLK, IRF5, STAT4, TNFAIP3, TNFSF4, 6q21 and 22q11.21), nine new ones were identified as susceptibility genes or loci for SLE for the first time (ETS1, IKZF1, RASGRP3, SLC15A4, TNIP1, 7q11.23, 10q11.22, 11q23.3 and 16p11.2). The result was reported in Nature Genetics in November 2009.

In the replication stage of leprosy, the top 93 SNPs were replicated in three independent samples consisting of 3,254 cases and 5,955 controls. We identified significant association (P<10-10) within six genes CCDC122 (13q14), C13orf31 (13q14), NOD2 (16q12), TNFSF15 (9q32), HLA-DR (6p21) and RIPK2 (8p21) and a suggestive association (P<10-7) within LRRK2 (12q12). Of interest, most of the identified genetic factors are in the NOD2-mediated signaling pathway (which regulates the innate immune response). This work was published in the journal The New England Journal of Medicine in December 2009.

The GWAS of immune-related diseases in Chinese Han population enhanced our understanding of the genetic basis of these disorders and were expected to reveal new pathways that contribute to immune-related diseases. In addition, our studies highlighted the genetic heterogeneity of disease susceptibility between different ethnic populations, and called for more GWAS in non-European populations.

Poster No: P098-T
The Role of Epigenetic Silencing in Docetaxel Resistance in MCF-7 Human Breast Cancer Cell Line

Nor Zarina Zainal Abidin[1], Lena Kastl[2], Iain Brown[2], Andrew C Schofield[2,3]
[1]Department of Diagnostic Laboratory Services, National University of Malaysia [2]Department of Medical Oncology, Grampian University Hospital NHS Trust, Aberdeen Royal Infirmary, Aberdeen, UK [3]Department of Molecular and Cell Biology, University of Aberdeen, Medical School, Aberdeen, UK

Introduction Docetaxel belongs to a class of taxane and have been recognized as the most chemotherapeutic agent in the treatment of advanced breast cancer. Docetaxel has a wide spectrum of antitumor activity and highly effective as monotherapy and combination therapy across a variety of tumour types including breast cancer. However, some of the breast cancer patient are either intrinsically resistance or develop resistance during the docetaxel regimen. One of the possible caused of drug resistance among this patient is DNA methylation which resulted in gene silencing and failed to chemotherapy treatment. Thus this study was used 5-azacytidine (azacytidine) as one of the DNA methylation inhibitor agent to investigate the role of epigenetic silencing in human breast cancer resistance to docetaxel cell line.

Methods MCF-7 human breast cancer cell line (docetaxel- resistance and docetaxel sensitive) were treated in single docetaxel treatment and in combination with docetaxel and 8μM 5-azacytidine (azacytidine) at 24 and 48 hours. Cytotoxic assay (MTT assay) was used to measure the cell viability after appropriate time of treatment. Result in single treatment was compared to in combination treatment at 24 and 48 hours incubations by comparing the docetaxel inhibition concentrations (IC50) which has killed 50% of the cell lines.

Results Statistical analysis does not show significant differences in IC50 values between both of the treatment at 24 and 48 hours.

Conclusions This study demonstrates that by adding 8μM azacytidine as DNA methylation inhibitors was not reversed docetaxel resistance in MCF-7 breast cancer resistance to docetaxel cell line. However result from this study could not totally conclude that DNA methylation does not occur in docetaxel resistance among breast cancer patient. Therefore future research needs to be done using another approach to investigate the involvement of DNA methylation on docetaxel resistance in breast cancer.

Poster No: P099-F
BlsI- and GlaI-PCR assays – A New Method of Human DNA Methylation Study

Danila A Gonchar, Alexander G Akishev, Sergey Kh Degtyarev
SibEnzyme Ltd., Novosibirsk, Russian Federation

BlsI- and GlaI-PCR assays have been developed to study DNA methylation. A new method includes DNA hydrolysis by unique methyl-directed site-specific DNA endonucleases GlaI or BlsI with subsequent PCR. Study of DNA methylation in regulation region of human tumor suppressor genes has been performed for a new method evaluation. BlsI- and GlaI-PCR assays have revealed different methylation patterns in promoter region of DAPK1, in promoter and first exon region of RARB and in first exon region of RASSF1A and SEPT9 tumor suppressor genes in malignant cell lines HeLa, Raji, U-937, Jurkat and control L-68 cells. GlaI-PCR assay has shown a methylation of RARB promoter and first exon region in DNA from all malignant cell lines, but not in control L-68 cells. GlaI- and BlsI-PCR assays have displayed DNA methylation of RASSF1A first exon region in Raji and Jurkat cells and SEPT9 first exon region in U-937, Raji and HeLa cells. BlsI-PCR assay of DAPK1 promoter region has demonstrated an additional DNA methylation in Raji cells only.

GlaI- and BlsI-PCR assays may be useful in determination of human cancer diseases and their discrimination.

Poster No: P100-W
Mutation analysis of the SOD1 gene mutations in the Korean Amyotrophic lateral sclerosis (ALS) population

[1]***Min-Jung Kwon***, M.D., [2]Won Ki Paek, M.D., [2]Hyun Young Kim, M.D., [1]Jong-Won Kim, M.D., [1]Chang-Seok Ki, M.D., [2]Seung H. Kim, M.D.
[1]Departments of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, [2]Department of Neurology, College of Medicine, Hanyang University, Seoul, Korea

Background: SOD1 gene mutations are the most commonly identified causes of Amyotrophic lateral sclerosis (ALS), accounting for approximately 20% of familial ALS (FALS) cases and around 4% of sporadic ALS (SALS) cases. However, the frequency of SOD1 gene mutations in ALS varies in different ethnic groups and no previous studies have been carried out in Korea. The aim of this study was to determine the frequency of SOD1 gene mutations in a Korean ALS population.

Methods: A total of 154 ALS patients (146 SALS and 8 FALS) were screened for SOD1 gene mutation. We applied polymerase chain reaction (PCR) and direct sequencing of all coding exons and flanking intronic sequences of the SOD1 gene. The sequences of the patients were compared with reference sequences to detect any sequence variations.

Results: The overall frequency (FALS and SALS) of SOD1 gene mutations in our series was 5.2%. Two novel mutations (A5F and P69S) and three known mutations (F21C, G38R and G142A) were identified in seven FALS patients (87.5%, 7/8). The F21C mutation was recurrently observed in three unrelated FALS patients. One novel mutation (P69S) was identified in a SALS patient (0.7%, 1/146).

Conclusion: The frequency of SOD1 gene mutations in Korean FALS (87.5%) is much higher than those reported in the other ethnic groups but only one SALS case had SOD1 gene mutation. Therefore, screening of SOD1 gene is highly recommended in Korean FALS but other genes should be further studied in order to reveal the genetic background of Korean SALS.

Poster No: P102-F
Unique spectrum of SPAST variants in Estonian HSP patients: presence of benign missense changes but lack of exonic rearrangements

[1*,9]**Riin Tamm**, [2*]Mark Braschinsky, [3]Christian Beetz, [3,4]Elena Sachez-Ferrero, [5]Elve Raukas, [2]Siiri-Merike Lüüs, [6]Katrin Gross-Paju, [7]Catherine Boillot, [8]Federico Canzian, [1,5,9]Andres Metspalu, [2]Sulev Haldre.
[1]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia,[2]Department of Neurology, University of Tartu, Tartu, Estonia, [3]Institute for Clinical Chemistry and Laboratory Medicine, University Hospital Jena, Jena, Germany, [4]Laboratorio de Genética Molecular, Hospital Central Universitario de Asturias, Oviedo, Spain, [5]Estonian Biocentre, Tartu, Estonia, [6]Centre for Neurological Diseases, West Tallinn Central Hospital, Tallinn, Estonia, [7]International Agency for Research on Cancer, Lyon, France, [8]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany, [9]Estonian Genome Center of University of Tartu, Tartu, Estonia

Background Hereditary spastic paraplegia (HSP) is a clinically and genetically heterogeneous disorder that can be an autosomal-dominant, autosomal-recessive, or X-linked disease. The most common autosomal-dominant form of the disease derives from mutations in the SPAST gene.

Methods The aim of this study was to analyze 49 patients diagnosed with HSP from the Estonian population for sequence variants of the SPAST gene and to describe the associated phenotypes. Healthy control individuals (n = 100) with no family history of HSP were also analyzed. All patient samples were screened using denaturing high performance liquid chromatography (DHPLC) and multiplex ligation-dependent probe amplification (MLPA) assay. Samples with abnormal DHPLC and MLPA profiles were sequenced, with the same regions sequenced in control samples.

Results Sequence variants of SPAST were identified in 19/49 HSP patients (38.8%), twelve among them had pathogenic mutations. Within the latter group there was one sporadic case. Eight patients had pure, and four – complex HSP. The twelve variants were identified: seven pathogenic (c.1174-1G>C, c.1185delA, c.1276C>T, c.1352_1356delGAGAA, c.1378C>A, c.1518_1519insTC, c.1841_1842insA) and five non-pathogenic (c.131C>T, c.484G>A, c.685A>G, c.1245+202delG, c.1245+215G>C). Only 2 of these mutations had previously been described (c.131C>T, c.1245+202delG). Three mutations, c.1174-1G>C, c.1276 C>T, c.1378C>A, showed intrafamilial segregation.

Conclusion This study identified new variants of the SPAST gene which included benign missense variants and short insertions/deletions. No large rearrangements were found. Based on these data, 7 new pathogenic variants of HSP are associated with clinical phenotypes.

Poster No: P103-W
A rapid method of point mutation screening in Duchenne muscular dystrophy (DMD) by high resolution melt (HRM) analysis for genetic treatment through translational readthroughs

[1]Lai PS, [1]**Yim OS**, [1]Xiong G, [1]Tay SKH, [1]Low PS and [2, 3] Yap EPH
[1]Dept of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, National University Hospital, Singapore 119074, [2]Defence Research & Technology Office, Ministry of Defence, Singapore 669645, [3]DSO National Laboratories, Defence Medical & Environmental Research Institute, Singapore 117510

Duchenne's Muscular Dystrophy (DMD) is a fatal X-linked disorder. There is no known cure and current management includes molecular diagnosis and carrier screening. Determination of the precise type of mutations present in patients will contribute towards identification of cohorts amenable for different types of future genetic treatment approaches such as antisense splicing or nonsense codon suppression. However, current control of disease transmission depends on genetic counseling and identification of female carriers. The most common and easily detectable mutations in DMD are deletions but point mutations account for the remaining 40% of patients. Conventional screening for non-deletion mutations and carrier analysis utilizes DNA sequencing as the method of choice, but faster, cheaper, and equally accurate alternatives may be possible. We report here a rapid and cheap mutation screening approach using High Resolution Melting (HRM) for identifying DMD mutations and carriers. This technique can be applied to amplify and melt products from all the coding exons in DMD gene. The different melt profiles generated are then analyzed to determine the mutant and wild-type homozygotes from the heterozygotes. A clinical case was evaluated from a family carrying an exon 60 nonsense mutation

(UGA stop codon) using this approach. A 187-bases long fragment of the mutation-containing exon 60 of the DMD gene was amplified by PCR and subjected to HRM. The patient exhibited different melt profiles from normals, and two female relatives were found to be carriers of the mutation. The HRM results were confirmed to be accurate by re-sequencing. Subsequent genetic correction assays were carried out using aminoglycoside-induced translational readthrough approach. A construct carrying this exon 60 nonsense mutation in a luciferase reporter gene vector was assayed for luciferase activity after treatment with different concentrations of gentamycin, paromomycin and tobramycin at 24, 48 and 72 hours. Highest readthroughs of about 18% was observed using 2.0 mg/ml gentamycin at 48 hours treatment indicating gene correction during translation. Our results show that HRM screening can identify point mutations, in particular nonsense mutations which are amenable for genetic correction by drug-induced translation readthough approach. It can thus be used for rapid molecular and prenatal diagnosis, carrier determination as well as for screening of patients for future customized therapeutics.

Poster No: P104-T
Four novel cis-acting elements of CYP1A1 and CYP1A2 identified by allelic expression imbalance in Chinese human livers

Dazhi Wang[1], Jinxiu Shi[1], Zhongyang Shen[2], Hui Wang[2] , Beilan Wang[1], Hong Zheng[2], Zhengwen Jiang[1, 3,*] and Wei Huang[1,*]

CYP1A1 and CYP1A2 are the major two P450 genes responsible for the metabolic activation of various polycyclic aromatic hydrocarbons (PAHs) and heterocyclic aromatic amines (HAAs) to electrophilic reactive intermediates, leading to toxicity and cancer. Cis-acting variants were observed to make up to 20% of inter-individual variability in mRNA level of CYP1A2. In this study, we used the relative expression levels of two SNP alleles in the same sample as an effective approach for identifying cis-acting elements of CYP1A1 and CYP1A2. We selected 16 tagSNPs representative of all common SNPs (>8%) within a ~37kb region including 6kb CYP1A1 gene, 7.76kb CYP1A2 gene and 23.3kb spacer region, and genotyped them in 96 Chinese liver samples. Four tagSNPs were used as marker SNPs for determining allelic expression ratio of CYP1A1 or CYP1A2. Association analysis was carried out on allelic expression ratio as quantitative phenotype and the heterozygosity of tagSNPs. Four cis-acting regulatory SNPs (rSNP) CYP1A1-940C>T (rs4646418), CYP1A1+2573C>T (rs4646422), CYP1A2-2217G>A (rs2069521) and CYP1A2+734A>C (rs762551) were identified. CYP1A1-940C>T (rs4646418), located in a G-rich domain upstream of CYP1A1, was affecting expression of both CYP1A1 and CYP1A2, while the other three were affecting expression level of either CYP1A1 (by +2573C>T) or CYP1A2 (by -2217G>A; +734A>C). Transient expression assay confirmed CYP1A1-940C>T (rs4646418) as a functional regulatory SNP.

Poster No: P105-F
Interactors of the cardiac ion channel, KCNE1, encoded by a Long-QT Syndrome-causative gene, identified

Carin de Villiers[1], Hanlie Moolman-Smook[1], Paul Brink[2], Valerie Corfield[1]
[1] *Department of Biomedical Sciences, Faculty of Health Sciences, University of Stellenbosch, Francie van Zyl Drive, South Africa,* [2] *Department of Internal Medicine, Faculty of Health Sciences, University of Stellenbosch, Francie van Zyl Drive, South Africa*

Long-QT Syndrome (LQTS), a hereditary cardiac disease characterised by syncope and tragic sudden death, is caused by functional disturbances in ion channels that orchestrate cardiac repolarisation. Disease-causing mutations in several ion channel-encoding genes are not sufficient to explain the clinical variability associated with LQTS. KCNE1 (encoded by a known LQTS-causative gene) regulates a number of ion channel subunits. We aimed to identify proteins that, by their interaction with KCNE1, alter disease severity.

Yeast 2 hybrid (Y2H) technology isolated C-terminal KCNE1-interacting proteins from a human cardiac c-DNA library. Standard selection protocols revealed the strength and specificity of interactions, while database mining provided supportive evidence for plausibility of proposed interactions and allowed prioritisation of putative interactors for verification

studies (co-immunoprecipitation and 3D fluorescent microscopy). 8.98 x105 cDNA library clones were screened. Following selection steps, 100 clones were sequenced and assessed further.

Interesting interactors identified include actin isoforms, Yotiao, Xin, heat shock protein beta-7 (HSPB7) and the cardiac isoform of myosin-binding protein C (MYBPC3). A plausible role for the first four in the potassium channel interactome is corroborated by existing data; although more tenuous, binding of KCNE1 with sarcomeric telethonin makes an interaction with other myofibrillar components, such as MYBPC3, plausible. The potential for involvement of sequence variants in the genes encoding the identified interactors, as either direct cause or modifiers of LQTS, will be assessed in unrelated LQTS subjects and an extended South African LQTS-affected family (Brink et al., Circulation [2005]), respectively.

These candidate interactors are excellent, novel KCNE1-binding partners which, with further investigation of their roles in ion channel regulation, will increase understanding of ion channel function. Furthermore, defining their role in the development and clinical course of LQTS may facilitate risk-stratification and, ultimately, disease interventions for affected patients.

Poster No: P106-W
Chromosome-Wide Haplotype Sharing: Integrating Recombination Information into Studies on Human Population Genetic Relationship

Shuhua Xu[1, 2] Wenfei Jin, [1] Ying Wang, [4] Li Jin [1, 2, 3, 5]
[1]*Chinese Academy of Sciences and Max Planck Society (CAS-MPG) Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China;* [2]*Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Chinese Academy of Sciences, Shanghai 200031, China;* [3] *State Key Laboratory of Genetic Engineering and Ministry of Education (MOE) Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China;* [4] *Chinese National Human Genome Center at Shanghai 201203, China;* [5] *China Medical City (CMC) Institute of Health Sciences, Taizhou, Jiangsu 225300, China.*

The vast recombination information in human genome is generally ignored or deliberately avoided in studies on human population genetic relationship, because estimation of recombination parameter from genotyping data is practically computationally challenging. Here we propose chromosome-wide haplotype sharing as a measure of genetic similarity between human populations, which is an indirect approach to integrate recombination information into human genetic relationship studies without involving estimation of recombination parameter. By analyzing 20,177 SNPs on chromosome 21 with density of 1.6 kb per SNP in 11 human populations representing African, European and East Asian, we found chromosome-wide recombination among populations is strongly correlated and this correlation further correlates the genetic difference between human populations, indicating recombination of human populations are evolutionarily related. We further demonstrated haplotype sharing can be used to reconstruct reliable phylogeny of human populations, where about 40% of variation in haplotype sharing matrix can be attributed to recombination, while only about 10% of the variation can be attributed to drift. With simulation studies, we found that haplotype sharing in single genome region of 100 kb is sufficient to study populations diverged 40~500 generations, while larger window size is necessary for populations separated recently (< 40 generations). However, for distantly related populations (> 500 generations), the utility of single region to reconstruct correct phylogeny is limited, suggesting that the linear correlation of haplotype sharing and population divergence could have been disturbed by recurrent recombination events at large time scale.

Poster No: P107-T
Expression and Somatic CNVs of GPx3 gene in breast cancer

Wang Zhimin[1], Fei Fei[2], Wang Haifeng[1], Fu Deyuan[2], Wang Beilan[1], Chen li[2], Xu Shijie[1], Shao Zhimin[2], Huang Wei[1*]
Chinese National Human Genome Center at Shanghai

Breast cancer is one of the most common malignant diseases in the world. However, the genetic events leading to the development of breast cancer are not clear yet. In this study, whole genome expression assay and copy number

variations (CNVs) analysis were performed in 20 primary breast carcinomas and paired normal adjacent tissue by using Illumina HumanCNV370-Duo BeadChip and Sentrix® Human-8 Expression BeadChip. We found that the expression of GPx3 was widely inactivated in breast cancers (p=0.0001), but the methylation status was no difference. Moreover somatic CNVs of GPx3 gene were detected in these breast cancer samples. It was found that higher DNA copy numbers caused increased expression levels (p=0.027). Direct sequencing and real time quantitative PCR (RT-qPCR) confirmed the somatic CNVs. And RT-qPCR results of more 48 paired samples showed CNVs of GPx3 related with estrogen receptor (ER) status (p=0.041). Thus, our results indicated that the CNVs of GPx3 may be a valuable biomarker for the carcinogenesis and progression of the breast cancer.

Poster No: P108-F
Comparative proteomic analysis of Neuro 2a cells in response to siRNA-mediated silencing of Dot1

Chunmei Wang, Sauna Tsai, Saiming Ngai
Department of Biology, The Chinese University of Hong Kong, Shatin, Hong Kong SAR. China

Dot1(disruptor of telomeric silencing) specifically catalyzes the methylation of H3K79. In budding yeast, Dot1 was involved in telomeric silencing, meiotic checkpoint control and DNA-damage responses. However, its biological function in mammals remains unknown. In this study, Dot1 silencing Neuro-2a (N2a) cell line was established via transfection with Dot1-siRNA. The roles of Dot1 in N2a cells was investigated using proteomic approach.The results indicated that Dot1 silencing resulted in the loss of cell normal neuronal shape. Whereas, Dot1 silencing showed no apparent effects on cell growth, cell cycle or induced cell death. Three proteins (stress-induced-phosphoprotein1, RAD23b homolog and peroxiredoxin 1) were up-regulated and five proteins (far upstream element-binding protein 1 (FBP-1), nucleosome assembly protein 1-like 1 (NAP1L1) isoform 2, eukaryotic translation initiation factor (eIFs) 3, nascent polypeptide-associated complex (NAC) alpha subunit and prohibitin) were down-regulated after Dot1 silencing. Moreover, decreased expression of p53 was also observed. This research provided us the evidences for the multi-functions of Dot1 at onaxon and dendrite outgrowth, transcriptional regulation, protein translation and folding, apoptosis and DNA breaks repair, which indicated the roles of Dot1 in brain dysfunction of neurodegenerative disease or aging.

Poster No: P110-T
Comparing genome-wide chromatin profiles using ChIP-chip or ChIP-seq

Wardenaar R.[1] Johannes F.[1], Colomé-Tatché M.[2], Mousson F.[3], de Graaf P.[3], Mokry M.[4], Guryev V.[4], Timmers H.T.[3], Cuppen E.[4] and Jansen R.C.[1,5]
[1]Groningen Bioinformatics Centre, University of Groningen, Kerklaan 30, Biologisch Centrum, 9751 NN Haren, The Netherlands. [2]Institute for Theoretical Physics, Leibniz Universität Hannover, Appelstr. 2, D-30167, Hannover, Germany. [3]Department of Physiological Chemistry, University Medical Center Utrecht, Universiteitsweg 100, 3508 AB Utrecht, The Netherlands. [4]Hubrecht Institute, KNAW and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands. [5]Department of Genetics, University Medical Centre Groningen, Hanzeplein1, 9713 GZ Groningen, The Netherlands.

ChIP-chip and ChIP-seq technologies provide genomewide measurements of various types of chromatin marks at an unprecedented resolution. With ChIP samples collected from different tissue types and/or individuals, we can now begin to characterize stochastic or systematic changes in epigenetic patterns during development (intra-individual) or at the population level (inter-individual). This requires statistical methods that permit a simultaneous comparison of multiple ChIP samples on a global as well as locus-specific scale. Current analytical approaches are mainly geared towards single sample investigations, and therefore have limited applicability in this comparative setting. This shortcoming presents a bottleneck in biological interpretations of multiple sample data.

To address this limitation, we introduce a parametric classification approach for the simultaneous analysis of two (or more) ChIP samples. We consider several competing models that reflect alternative biological assumptions about the global distribution of the data. Inferences about locus-specific and genomewide chromatin differences are reached through the estimation of multivariate mixtures. Parameter estimates are obtained using a version of the Incremental Expectation

Maximization algorithm (IEM). We demonstrate efficient scalability and application to three very diverse ChIP-chip and ChIP-seq experiments. The proposed approach is evaluated against several published ChIP-chip and ChIPseq software packages. We recommend its use as a first-pass algorithm to identify candidate regions in the epigenome, possibly followed by some type of second-pass algorithm to fine-tune detected peaks in accordance with biological or technological criteria.

Poster No: P111-T
Identification of Mutations in the ATP2A2 gene in Chinese patients with Darier's disease

ZM Niu[1] , ZH Zhang[2], B Chai[3], W Huang[1]
Chinese National Human Genome Center at Shanghai

Darier's disease (DD), also known as Darier-White disease or keratosis follicularis, is an autosomal dominant skin disorder characterized loss of adhesion between epidermal cells and abnormal keratinization. The relationship of ATP2A2 mutations and DD has been widely reported since 1999. In our study, we examined the ATP2A2 gene mutations of eight sporadic and two familial Chinese DD patients. The entire coding regions, about 500bp promoter and at least 50bp of the exon-intronic boundary sequence of ATP2A2 were directly sequenced from the genomic DNA extracted from the blood samples. A total of five mutations had been identified. In a familiar patient, we found a S920Y missense change which was previously described by Sakuntabhai et al. in 1999. The other four mutations we found in sporadic patients haven't been reported before. They are two frame shift changes, 2236insC and 3356-3357delTG, which cause codon changes; and two exon-intronic boundary region changes,1851-6A to G in intron10 and 3084+5G to C in intron16, which may cause the alternative splicing changes. Our results provide the new mutations of ATP2A2 gene for Chinese patients with DD.

Poster No: P112-F
Linkage analysis in Benign Adult Familial Myoclonic Epilepsy

[1,2,3]**Patra Yeetong,** [4]Chaichon Locharernkul, [5]Krisna Piravej, [4]Roongroj Bhidayasiri, [6]Tayard Desudchit, [4]Nath Pasutharnchat, [4]Jakrin Loplumlert, [4]Chusak Limotai, [2,3]Kanya Suphapeetiporn, [2,3]Vorasuk Shotelersuk
[1]Interdepartment of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University [2]Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University [3]Molecular Genetics Diagnostic Center, King Chulalongkorn Memorial Hospital, Thai Red Cross [4]Department of Internal Medicine, Faculty of Medicine, Chulalongkorn University [5]Division of Neurology, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University [6]Department of Rehabilitation, Faculty of Medicine, Chulalongkorn University, Bangkok, 10330, Thailand

Benign Adult Familial Myoclonic Epilepsy (BAFME) is an autosomal dominant disorder typically characterized by adult-onset cortical tremor and generalized seizures. BAFME diagnosis is based on clinical and electrophysiological criteria including irregular postural myoclonic tremor of the distal limbs, familial history of epilepsy, autosomal dominant inheritance and a rather benign outcome. The diagnosis is confirmed by electrophysiological features favoring cortical reflex myoclonus (enhanced C reflex at rest, giant somatosensory evoked potentials (SEPs), premyoclonus cortical spikes detected by jerk-locked back-averaging method) and good response to antiepileptic drugs. Although BAFME were previously mapped on chromosome 8q24 and 8q23.3–q24.1 in Japanese pedigrees and chromosome 2p11.1–2q12.2 in European pedigrees, the causative genes have not been identified. Here we studied a large BAFME family in Thailand consisting of approximately 13 affected members. We performed linkage analyses using six microsatellite markers covering chromosome 8q23.3-q24.1, and three covering chromosome 2p11.1-q12.2. Amplified fragments were resolved on an ABI PRISM 3100 Genetic Analyzer and analyzed with GeneScan Analysis and Genotyper software. Two-point linkage analysis was carried out by MLINK program. Our data demonstrate that two reported loci were excluded and the causative gene responsible for BAFME in the Thai pedigree may be located on a new region other than 8q23.3–q24.1 and 2p11.1–q12.2. In summary, we have excluded two previously linked loci for BAFME in Thai family. Currently, we are performing whole genome linkage analysis using 400 tandem repeat markers with average spacing of 10 cM, attempting to identify a new locus for this disease. Knowing the etiology of this disease will provide accurate genetic counseling for the affected families and better understanding the disease.

Poster No: P113-W
Human Alu repeats digestion with restriction endonucleases in vitro and in silico

Viktor Tomilov, Valery Chernukhin, Murat Abdurashitov, Sergey Degtyarev
SibEnzyme Ltd., Novosibirsk, Russian Federation

Earlier we have developed a simple method to carry out mammalian DNA digestion with restriction enzymes in silico based on the known DNA sequence of corresponding genomes [1]. This method allows to calculate lengths of all DNA fragments, which are formed after a whole genome digestion at recognition sites of restriction enzyme, and to construct the distribution diagrams of the calculated DNA fragments. These distribution diagrams display distinct peaks of DNA fragments of the definite lengths due to a presence of DNA repeats in eukaryotic genomes. Comparison of the obtained peaks in distribution diagrams and results of human chromosomal DNA hydrolysis by several restriction endonucleases has shown a good correspondence of theoretical and experimental data for 17 restriction enzymes [2]. Theoretical patterns of DNA cleavage with endonucleases are formed mostly as a mix of Alu and LINE1 repeats digestion products. As a rule DNA fragments of 300 bp and less are produced from Alu repeats [3]. The Alu repeats, which belongs to SINE group, is one of the most abundant and well characterized repetitive elements in human genome and include more than 1150 thousand copies (about 10% of human genome).

In this work we have suggested a method to carry out Alu repeats digestion with restriction endonucleases in silico. We have presented an advanced Alu repeats data base [4], which allows to construct the digestion diagrams in a simple way, and developed a software for a work on standard personal computer with providing the calculations results in several minutes. Human Alu repeats cleavage at definite nucleotide sequences (4-, 5-, 6- and 7-bp), which are the recognition sites of known restriction endonucleases, has been performed and corresponding diagrams of DNA fragments distribution have been constructed. The theoretical data have been compared to experimental patterns of human DNA hydrolysis with restriction endonucleases (AluI, AsuHPI, BpmI, Bpu10I, Bst2UI, BstSCI, BstDEI, BssECI, BstMAI, BstSFI, BstXI and HinfI) and a good correspondence for the most of DNA digestion diagrams has been observed.

The suggested method allows to simplify a study of human DNA cleavage with restriction endonucleases considering set of Alu repeats sequences instead of the whole human genome (about 440 million bps vs more than 3 billion bps).

References:
1. M.A. Abdurashitov, V.N. Tomilov, V.A. Chernukhin, D.A. Gonchar, S.Kh. Degtyarev Mammalian chromosomal DNA digestion with restriction endonucleases in silico // Ovchinnikov bulletin of biotechnology and physical and chemical biology V.2, No 3, pp 29-38, 2006 (online version: http://science.sibenzyme.com/article14_article_27_1.phtml)
2. Abdurashitov M.A., Tomilov V.N., Chernukhin V.A., Gonchar D. A., Degtyarev S. Kh. Comparative analysis of human chromosomal DNA digestion with restriction endonucleases in vitro and in silico. // Medical genetics V.6, No 8, pp 29-36, 2007 (online version: http://science.sibenzyme.com/article14_article_31_1.phtml)
3. Murat A Abdurashitov, Victor N Tomilov, Valery A Chernukhin, Sergey Kh. Degtyarev A physical map of human Alu repeats cleavage by restriction endonucleases// BMC Genomics 2008, 9:305
4. Aligned set of human Alu-repeats. Available online at http://science.sibenzyme.com/article4_article_37_1.phtml

Poster No: P115-F
Improved detection of cell-free methylated DNA in plasma

Inge Søkilde Pedersen[1], Henrik Krarup[1], Ole Thorlacius-Ussing[2], Poul Madsen[1]
[1] Section of Molecular Diagnostics, Department of Clinical Biochemistry, Aalborg Hospital, Denmark [2] Department of Surgical Gastroenterology, Aalborg Hospital, Denmark

Detection of cell-free methylated DNA in plasma is a promising tool for tumour diagnosis and monitoring. Due to the very low amount of cell-free DNA in plasma, sensitivity of the detection methods are of utmost importance. The vast majority of currently available methods for analysing DNA methylation are based on bisulphite-mediated deamination of cytosine. Cytosine is rapidly converted to uracil during bisulphite-treatment, whereas 5-methylcytosine is only slowly converted. Hence, bisulphite-treatment converts an epigenetic modification into a difference in sequence, amenable to analysis either by sequencing or PCR based methods. However, the recovery of bisulphite-converted DNA is very poor. Here we

introduce an alternative method for the crucial steps of bisulphite removal and desulphonation, vastly improving recovery, especially for specimens with low levels of methylated DNA.

The method is based on an accelerated deamination step and magnetic silica purification of DNA in combination with a first round of PCR amplifying 18 methylated markers concurrently, followed by individual detection of the 18 methylated markers by real-time PCR. Detection frequencies of methylated copies (calculated as the percentage of positive results of 18 markers analysed 6 times each)range from 31% when analysing 1.25 methylated copies to 86% when analysing 10 methylated copies.

The present method allows low levels of DNA to be easily and reliably analysed, a prerequisite for the clinical usefulness of cell-free methylated DNA detection in plasma.

Poster No: P116-W
A Survey on General Knowledge of Genetics among Medical Students in the Faculty of Medicine, University of Colombo

[1]**Vajira H. W. Dissanayake** [1]Fathu Mohammed, [1]Dilum Nisansala, [1]Rohan W Jayasekara
[1]*Human Genetics Unit, Faculty of Medicine, University of Colombo*

The demand for genetic services has increased due to the rapid development and increased public awareness of genetics. Therefore assessing the knowledge of medical students is very important to collect baseline information which may help in designing future genetic educational programs. To fulfill this requirement we carried out a survey of general knowledge of genetics among medical students of Faculty of Medicine, University of Colombo using a self administered questioner developed and validated for this purpose. The questioner was administered to all students of the faculty during April 2009. 634 (51.8%) responded. 288 (45.4%) were males. Analysis of the total marks showed that 56 (8.8%), 103 (16.2%), 107 (16.9%), 299 (47.2%) and 69 (10.9%) students had obtained 0-20%, 21-40%, 41-60%, 61-80% and 81-100% marks respectively. Comparison between students in different years showed that knowledge was highest among first and second year students. Regarding different aspects of genetics, it was found that the average marks for the nature of genetic martial, transmission, gene expression, gene regulation and genetics and society was 58.8%, 69.53%, 63.3%, 69.8% and 50.4% respectively. However, overall knowledge on evolution was very much less (27.8%) than the other aspects. Comparison of the groups of students regarding the different aspects showed that the knowledge on nature of genetic material, transmission, gene expression and gene regulation was more than 70% in first, second and fourth year students. However the knowledge on other aspects was low among all students. This information would be very useful in designing future genetic education programmes.

Poster No: P117-T
Rapid Genetic Diagnosis of Consanguineous Families with histological findings of FSGS by Homozygosity Mapping Coupled with Whole Exome Capture and Massively Parallel Sequencing

Chaker N Adra[1,10] Khaldoun I. Al-Romaih[1,4], Hamad Al-Mojalli[2], Giulio Genovese[4,5], Hadeel Al-Manea[3], Nourah Atallah[1], Maha Al-Rodayan[1], Mohammed Al-Jondubi[6], Mohammed Al-Suleiman[6], Astrid Weins[7], Richard P. Lifton[8] and Martin R. Pollak [4,9].
[1]Stem Cell Therapy Program, [2]Department of Pediatrics, [3]Department of Pathology and Laboratory Medicine, King Faisal Specialist Hospital & Research Centre, Riyadh, 11211. [4]Renal Division, Department of Medicine, Brigham and Women's Hospital and HarvardMedicalSchool, Boston, MA. [5]Department of Mathematics, Dartmouth College, Hanover, NH 03755. [6]Department of Nephrology, RiyadhKharjHospitals, Riyadh 11159. [7]Department of Pathology, Brigham and Women's Hospital, Boston, MA. [8]Department of Genetics, Howard Hughes Medical Institute, YaleUniversitySchool of Medicine, New Haven, CT06510. [9]Nephrology Division, Brigham and Women's Hospital, HarvardMedicalSchool, Boston, MA, USA [10]TransplantationCenter, Children's Hospital Boston, Brigham and Women's Hospital, HarvardMedicalSchool, Boston, MA.USA

Focal segmental glomerulosclerosis (FSGS) is a histological glomerular phenotype that can be familial, primary (idiopathic), or secondary to a multitude of pathological processes affecting the kidney, including such tubulointerstitial diseases as nephronophthisis. Mutations in a number of distinct nephronophthisis genes (NPHPs) have been described to date. We describe consanguineous unrelated Saudi Arabian families with sequence variants in one of the NPHP genes, namely NPHP1. Affected individuals in the two families presented with end-stage renal disease and clinical and histological features consistent with focal segmental glomerulosclerosis. Since FSGS patients may present atypical radiological findings, making the clinical diagnosis of the genetic syndrome difficult, we applied whole-genome single-nucleotide polymorphism analysis followed by state of the art sequence capture and exome sequencing on genomic DNA samples from these families. This analysis facilitated accurate diagnosis after isolation of homozygosity run of ~ 2 Mb. This homozygous run falls between rs6754115 (genomic position 109,328,776) and rs17464100 (genomic position 111,284,252), and is identical in affected subjects from the unrelated families. This provides evidence that this deletion is widely spread in the families' geographical regions, and implies its significant involvement in the development of chronic kidney failure in Saudi Arabia. This work emphasizes the importance performing genetic screening for this NPHP allele in CRF patients an outline an essay for this purpose.

Poster No: P118-F
Comparative analysis of (TG/CA)n repeats in the human and chimpanzee genomes

Vineet K. Sharma[+], Naveen Kumar[+], Todd D. Taylor[+], Srinivasan Ramachandran[++]
RIKEN

About 50% of the human genome consists of repetitive elements comprised of tandem and interspersed repeats. Among the various types of tandem repeat sequences, the dinucleotide (TG/CA)n repeats are one of the most abundant. These repeats also exhibit length polymorphism and in many cases have been shown to modulate gene expression and have the propensity to undergo conformational transitions under in vivo conditions. The genome sequence of chimpanzee (Pan troglodytes) offers the most relevant nonhuman dataset for comparison of these repeats in human. We carried out a whole genome analysis of these repeats in the human and chimpanzee genomes including the genic, intergenic, and promoter regions. We classified these repeats into three length categories (Type I ($6 \le n < 12$), Type II ($12 \le n < 23$), and Type III ($n \ge 23$)) based on their biological functions like length polymorphism, role in modulation of gene expression, propensity to undergo conformational transitions, recombination, etc. and examined their distribution in the two genomes. We then examined the conservation and polymorphism of the orthologous repeats in the selected regions. All three processes of expansion, contraction, and conservation of the repeats were observed. We also observed that a significantly higher proportion of (TG/CA)n repeats were expanded either in chimpanzee or human, while almost similar proportions of these repeats remained conserved or have contracted. However, no bias was evident with respect to the genomic location of these repeats.

Poster No: P119-T
The investigation of TAP1 and TAP2 polymorphisms in susceptibility to ankylosing spondylitis

Lei Rong[1], Fan Xiaoyun[2], Niu Zhenmin[1], Sun Weiwei[1], Wang Ying[1], Wang Yuan[2], Shen Nan[2], Huang Wei[1*]
Chinese National Human Genome Center at Shanghai

Ankylosing spondylitis (AS) is a chronic inflammatory joint disease that chiefly affects the sacroiliac joints and the spine. Genetic factors, in addition to the HLA-B27, play a major role in the pathogenesis of AS. In this study, we investigated the association of polymorphisms in transporter genes TAP1 and TAP2 with AS in Chinese Han population. TAP genes were genotyped in 194 AS patients and 475 randomly chosen unrelated controls by resequencing the entire gene regions directly, including 2000bp promoter, as well as all exons and at least 200bp of the exon-intronic boundary sequence. The genetic association with TAP1 and TAP2 was confirmed. There are 232 polymorphisms were identified and 105 of them were novel. Differences in genotypes and allele frequencies of 82 SNPs with MAF>5% were assessed by the use of contingency tables and the calculation of chi-square test. We found several associated polymorphisms and the most significant of which was rs41316548 located in intron_11 of TAP2 gene. The frequencies of the A allele is significantly increased in case compared with the control group (10.4% vs. 3.29%, P = 1.71E-07) with an odds ratio being 3.417 (95% confidence interval, 2.104-5.551), and the A allele showed significant susceptibility to AS (P = 1.80E-07, in dominant disease model) with an odds ratio being 3.618 (95% confidence interval, 2.181-6.002). TAP1 and TAP2 are involved in

antigen processing and presentation, and polymorphisms in these genes may be associated with altered antigen-peptide selection and the level of restoration of surface expression of MHC class I molecules. Our study suggested that TAP genes may act as a potential factor in the mechanism of AS pathogenesis.

Poster No: P120-W
Systematic mutation search in families with XLMR by next-generation sequencing

Haas, SA.[1], Kalscheuer, VM.[1], Chen, W.[1,2], Hu, H.[1], Emde AK.[1], Menzel C.[1], Bienek, M.[1], Zemojtel, T.[1], O'Keeffe, S.[1], Vingron, M.[1], Ullmann, R., Wrogemann, R.[1], Tzschach, A.,[1] M., Strobl-Wildemann, G.[3], Wieacker, P.[4], EURO MRX Consortium [5], Ropers, HH.[1]
1 Max Planck Institute for Molecular Genetics, Berlin, Germany 2 Max-Delbrück Center, Berlin, Germany 3 Praxis, Passau, Germany 4 Institute for Human Genetics, University Hospital Münster, Germany 5 Berlin Germany, Nijmegen The Netherlands, Tours France, Paris France, Leuven Belgium

X-linked mental retardation (XLMR) affects 1-2/1,000 males and accounts for approximately 10% of all mental retardation (MR). To date, almost 90 different XLMR genes have been identified. Mutations in these genes are seen in approximately 50% of the families. Thus, many more XLMR genes remain to be identified, especially for the non-syndromic forms of XLMR. We have set out to find the causative gene defect in a total of 200 families with this condition. To this end, we systematically enrich X-chromosomal exons employing several gene partitioning techniques followed by next generation sequencing, and depending on the outcome, by sequencing flow-sorted X-chromosomes. In contrast to previous studies (Tarpey et al., Nat. Genet. 2009) this strategy should enable us to detect mutations in the entire non-repetitive portion of the X-chromosome.

Poster No: P121-T
Gene Expression Profiling of Ovarian Granulosa Cell Tumour from Formalin-fixed Paraffin-embedded (FFPE) samples.

[1,5]**Norfilza Mohd Mokhtar**, [1] Koon Vui-Kee, [2] Ahmad Zailani Hatta, [3] Isa Mohammed Rose, [4] Razmin Ghazali, A Rahman A Jamal [5]
[1]Department of Physiology, Faculty of Medicine, Universiti Kebangsaan Malaysia, [2]Department of Obstetrics and Gynecology of UKM Medical Centre,[3]Department of Pathology, Faculty of Medicine, Universiti Kebangsaan Malaysia, [4]Unit Histopathology, Department of Pathology, Hospital Kuala Lumpur, [5]UKM Medical Molecular Biology Institute, Kuala Lumpur, Malaysia

Introduction: Ovarian granulosa cell tumours (GCTs) are uncommon gynaecological cancer that comprises around 2 to 5% of all ovarian neoplasm. The rarity of this tumour makes the early diagnosis is more crucial as it can easily misdiagnose with other malignant tumours. The purpose of this study was to identify the differences of expressed genes in ovarian GCTs and normal ovary tissues using the high-throughput microarray analysis.

Material and Methods: Total RNA was extracted from eight archived formalin-fixed paraffin embedded (FFPE) samples of ovarian GCTs and normal ovarian tissues. The extracted RNA were then analysed using Illumina Whole-Genome DASL assay (cDNA-mediated annealing, selection, extension, and ligation). We used GeneSpring GX10.0 and FlexArray software to determine the differentially expressed genes.

Results: Unsupervised clustering of gene expression correctly clusters the samples into two major groups. Of 24,526 known genes, 215 genes were differentially expressed with cut-off p-values < 0.05 and 2-fold changes in ovarian GCTs compared to normal ovaries. Among these, 47 genes were up-regulated and 168 genes were down-regulated in ovarian GCTs compared to normal ovaries. Biological processes for the genes were associated with cell cycle and growth, cell adhesion, anti-apoptosis and apoptosis.

Conclusion: Using the array data, we identified genes with their functional groups in ovarian GCTs. Further characterisation of the genes and their signalling pathways are essential to understand the pathogenesis of the tumour.

Poster No: P122-F
Whole Genome DNA Methylation Profiling of Schizophrenia

[1]**Benjamin Chanrion**,[1]Yurong Xin, [2]Anne O'Donnell, [1]Peter Graham, [2]Timothy H. Bestor, [1]Victoria Arango, [1]Andrew J. Dwork, [1]J. John Mann, and [1]Fatemeh Haghighi
[1] Department of Psychiatry, Columbia University and The New York State Psychiatric Institute, [2] Department of Genetics and Development, Columbia University, New York, NY, USA.

Background: DNA methylation may play a role in the etiology of neuropsychiatric disorders, possibly through abnormal genomic methylation patterns that regulate genes involved in brain development or physiology. Our aim is to investigate DNA methylation profiles of schizophrenia.

Methods: In order to better understand both the wild type genomic DNA methylation patterns and aberrant methylation events that occur in disease states, we have developed a cost-effective, unbiased, whole-genome methylation profiling technique, methylation mapping analysis by paired-end sequencing (MethylMAPS), that can assay the methylation state of more than 80% of the CpG sites in the human genome. The Methyl-MAPS method couples advances in next generation sequencing with enzymatic fractionation of DNA by methylation state, allowing for mapping of methylation at high genomic coverage in non-psychiatric controls and schizophrenia cases. Our Methyl-MAPS data have been validated by bisulfite sequencing using the Sequenom MassARRAY platform, with >75% correlation across 46 genomic regions that included gene promoters, internal exons and introns, and repeat sequences.

Results: We focused on the prefrontal cortex due to converging evidence from neuroimaging and functional studies implicating this region in schizophrenia. Secondarily, we also examined the auditory cortex within the schizophrenia samples, because schizophrenia disorder includes defects in sensory perception and processing. These data provide the first genome-wide DNA methylation profiling study of schizophrenia, allowing for identification of aberrant methylated regions including genes as well as repetitive elements that are typically ignored in existing DNA methylation profiling methods. These data are overlaid with whole genome gene expression data from the same sample subjects to determine the influence of DNA methylation on gene transcript activity. Differentially methylated regions with marked differences in transcriptional activity among our cases and controls will then be subject to further validation on the Pyrosequencing platform.

Discussion: This large-scale whole genome DNA methylation study has identified numerous methylation changes, that will serve as candidates for further investigations into the role of DNA methylation in schizophrenia disorder. DNA methylation abnormalities thus identified may have clinical utility as biomarkers, and evaluation of the frequency of these alterations may help identify etiologic factors involved in schizophrenia.

Poster No: P123-T
Detection, genetic impact and distribution of hypervariable region in human Na(+)-Ca(2+) exchanger (NCX1) intronic region among European, Asian and African populations.

[1]**Katrin Kepp,** [1]Laura Kasak, [1]Elin Org, [1]Siim Sõber, [1]Piret Kelgo, [2]Margus Viigimaa, [1,3]Gudrun Veldre, [1]Neeme Tõnisson, [1]Peeter Juhanson, [1]Margus Putku, [4]Andreas Kindmark, [5]Viktor Kozich and [1]Maris Laan
[1]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia; [2]Centre of Cardiology, North Estonia Medical Centre, Tallinn, Estonia; [3]Department of Cardiology, University of Tartu, Tartu, Estonia; [4]Department of Medical Sciences, UppsalaUniversityHospital;[5]Institute of Inherited Metabolic Diseases, CharlesUniversity – First Faculty of Medicine, Prague, Czech Republic

Although short indels (<100bp) are referred as the second most abundant genetic variation in the human genome, little attention is paid to their effect on gene regulation and disease susceptibility. We identified a 14bp indel (rs11274804) within the second intron of human cardiovascular disease (CVD) candidate gene, Na(+)-Ca(2+) exchanger (NCX1) and screened the locus in two Eastern-European sample sets: essential hypertension (Estonia; cases n=470/controls=652) and coronary artery disease, CAD (Czech; cases n=257/controls=413). The analyzed genomic segment (348bp) was hypervariable, represented by seven different indel alleles (minor allele frequencies <8.4%). An association was detected between the carrier status of 14bp indel allele and the diagnosis of CAD (P=0.0016, OR=2.02), but not with hypertension.

A suggestive evidence was also detected for the association with heart rate, serum triglyceride and LDL levels (P=0.04). The alignments of human-chimpanzee-macaque sequences revealed that the major human variant (allele frequency 90.45%) was actually a human-specific deletion compared to other primates. This derived variant was surrounded by an abundance of other short (5-43bp) deletions and duplication (40bp) polymorphisms. High variation occurrence indicates a potential indel hotspot area triggered by the initial deletion in human lineage. In order to characterize the human population diversity of hypervariable region in NCX1 intron 2, 12 populations from three continents Europe (n=150), Asia (n=71), and Africa (n=364) were analyzed. Among African populations (both North- and Sub-Saharan Africa) the carrier-frequency of the ancestral allele lacking the 14bp deletion was significantly higher compared to other populations (allele frequencies up to 35%).

Poster No: P124-F
Combined genomic and phenotype screening to discover novel therapeutic targets in breast cancer

Wendy W. Soon[1], Lance D. Miller[2], Cyril Dalmasso[3], Kartiki V. Desai[1] and Edison T. Liu[1]
[1]Department of Cancer Biology and Pharmacology, Genome Institute of Singapore, Singapore 138672 [2]Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC 27157 [3]Department of Computational and Mathematical Biology, Genome Institute of Singapore, Singapore 138672

Secretory factors and extracellular receptors that drive cancer progression make attractive immunotherapeutic targets. To discover such novel targetable oncogenes, we used a whole-genome data-mining approach that takes advantage of large independent microarray studies of breast tumors annotated for clinical outcomes. Genes without previous implication in breast cancer were identified based on their significant association with cancer recurrence in multiple cohorts, and further filtered for secretory factors and extracellular receptors based on gene ontology. A high-content screening strategy was then applied to the candidates to assess the phenotypic consequences of altering candidate gene expression. This screen suggested that serine protease inhibitor Kazal-type 1 (SPINK1), a secreted extracellular factor, was associated with breast cancer aggressiveness. We demonstrate that enforced SPINK1 expression and exposure to exogenous recombinant SPINK1 induce both invasiveness and survival of breast cancer cells and down-regulation of SPINK1 is associated with cancer cell death. Moreover, increased expression of SPINK1 resulted in an increased resistance to drug-induced apoptosis. Immunohistochemical analysis of breast cancer tissue microarrays suggested that high SPINK1 expression was not seen in normal breast but is apparent in advanced disease. Moreover, SPINK1 shows nuclear localization primarily in higher grade breast cancers. This nuclear localization was recapitulated when tagged exogenous SPINK1 was applied to cells in vitro. SPINK1 is therefore a marker for cellular robustness in breast cancer and maybe a potential therapeutic target in this disease.

Poster No: P125-W
HGVbaseG2P: an advanced database for the integration and interrogation of genetic association datasets

Hastings R, Thorisson GA, Gollapudi VLS, Beck T, Lancaster O, Free R, Brookes AJ*
Department of Genetics, University of Leicester, UK
University of Leicester, UK

Genetic association study data are rarely published in any comprehensive fashion, neither in journals nor in databases, and in the case of negative findings such data are often not reported at all. Consequently, it is difficult to compare and contrast the results of different studies, and it is completely impossible to examine a full and unbiased picture of all the data that exist.

To address this deficit, the Human Genome Variation Genotype to Phenotype database (HGVbaseG2P: http://www.hgvbaseg2p.org) has been constructed. This free and open access resource provides a new publication medium for summary-level genetic association data, combining the best features of a database and a scientific journal. In doing so, it employs powerful graphical and text based data presentation methods which enable the simultaneous visualisation and co-examination of many studies, without any risk of re-identification of study subjects.

The latest release of HGVbaseG2P (v3) improves upon the foundations of previous versions by providing a simpler interface, and enhanced search capabilities. These allow researchers to identify studies of interest based on chromosomal regions/genes and markers. Studies of interest can be selected for comparison at both genome-wide and region-specific levels.

We constantly and actively search for new data available in publications and multiple online public resources, not least the NHGRI GWAS catalog (http://www.genome.gov/gwastudies). Consequently, the database now hosts >18,000,000 p-values and 352 studies (vs 2,000 p-values and 373 studies in the NHGRI GWAS catalog), representing ~5% of all such data yet produced.

We will soon release new software that will allow researchers to directly submit studies and association data into HGVbaseG2P. This will be accompanied by a data access control system enabling data submitters to decide what studies and which data elements are to be shared and with whom (e.g. collaborators, own research group).

By summer 2010 the complete HGVbaseG2P code base will be available for labs, consortia, institutes, etc, to install on their own servers so they can run their own genetic association database. Such installations will be interoperable and can be searched as a federated network, whilst leaving data curation and data sharing issues fully under the auspices of each database team whose efforts will thereby be more easily and more fully recognized.

Acknowledgements:
The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 200754 – the GEN2PHEN project.

References:
1. Thorisson GA, Lancaster O, Free RC, Hastings RK, Sarmah P, Dash D, Brahmachari SK, Brookes AJ. Nucleic Acids Res. 2009 Jan;37(Database issue):D797-802.

Poster No: P126-T
Computational analysis of polymorphisms in gene regulatory regions predicts their functional impact

[1]**Patrick Beaulieu**, [1]Manon Ouimet, [1]Vincent Gagné, [3]Tomi Pastinen, [1,2]Daniel Sinnett
[1]Centre de recherche, CHU Sainte-Justine, Montréal, QC, Canada, [2]Département de Pédiatrie, Université de Montréal, Montréal, QC, Canada,[3]McGill University and Genome Quebec Innovation Center, Montreal, QC, Canada

Previous efforts to identify functional disease-causing DNA variants were focussed on the coding regions of candidate genes, as coding variants (cSNPs) might have a direct impact on the structure and function of the affected proteins. However, abnormal expression of finely regulated genes can also influence different biological processes. Thus, analysis of the evolutionary conservation in the regulatory regions of candidate genes as well as investigation of the functional impact of polymorphisms in these regions (rSNPs) should improve our knowledge of complex disease etiology.

In an effort to study regulatory polymorphisms in the genome level, we have previously studied several layers of information in the context of regulatory regions including gene structure, SNP content, genomic patterns (i.e. CpG islands, conserved regions) and in silico analysis together with experimental results from electrophoretic mobility shift assays (EMSA) and promoter activity assays. These data has been integrated into a web-based environment, which is freely available and can be used to combine and analyze regulatory genomics data (see www.regulatorygenomics.org).

We have now combined the different layers of regulatory genomics data that has been stored in this database and we present a detailed genome wide computational analysis of regulatory regions content This analysis allowed us to establish criteria and parameters for better predictions of active regulatory regions as well as improved prediction of rSNPs with impact on gene expression. A better annotation should provide a more effective selection of regulatory targets that might have an impact on human disease and thereby facilitates our efforts of understanding the genetic determinants of complex human diseases.

Poster No: P127-T
A Modular Pipeline for Detecting Genetic Variations from Next-Generation Sequencing Data at NCBI

[1]**Chunlin Xiao**, [2]Tom Blackwell, [3]Al Ward, [1]Anatoly Mnev, [2]Wei Chen, [2]Paul Anderson, [2]Hyun Min Kang, [2]Bingshan Li, [2]Xiaowei Zhan, [3]Amit Indap, [3]Donald Stewart, [3]Wan-Ping Lee, [3]Michael Stromberg, [3]Wen Fung Leong, [3]Derek Barnett, [1]Richa Agarwala, [1]Justin Paschall, [1]Tam Sneddon, [1]Wratko Hlavina, [1]Avi Kimchi, [1]Mike DiCuccio, [1]Deanna Church, [1]Lon Phan, [1]Martin Shumway, [1]Don Preuss, [1]Eugene Yaschenko, [2]Gonçalo Abecasis, [3]Gabor Marth, [1]Stephen T. Sherry
[1]National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 45 Center Drive, Bethesda, MD 20892, USA; [2]Department of Biostatistics, University of Michigan at Ann Arbor, 1420 Washington Heights, Ann Arbor, MI 48109-2029, USA; [3]Boston College Biology Department, Higgins Hall, Chestnut Hill, Massachusetts 02467, USA

Next-generation sequencing (NGS) technologies have revolutionized genome sequencing by coupling extremely high throughput with low cost, thereby providing researchers with unprecedented opportunities to address many important biomedical problems efficiently. Large-scale resequencing projects, e.g. 1000 Genomes, TCGA, and TSP, have been initiated to extend our knowledge of single nucleotide polymorphisms (SNPs), short insertions/deletions (INDELs) and structural variations (SVs) and relate these variants to human diseases. The amount of NGS data submitted to public repositories such as the Short Read Archives (SRA) at the NIH National Center of Biotechnology Information (NCBI) is growing exponentially and submissions represent a wide array of technology platforms and sequence collection strategies. To process and analyze these data for variation detection in a uniform manner is a challenge requiring a standard modular pipeline. In collaboration with investigators at Boston College and the University of Michigan, NCBI is developing a framework Variation Discovery and Annotation Pipeline (Gpipe). The pipeline generates quality input sequence data from the Short Read Archives, checks sample identities, aligns the read data with the human reference genome sequences, refines the mapping of placed reads (duplicate removal and base quality recalibration etc), and calls SNPs, INDELs, and SVs according to data availability and project-specific policies. A centrally implemented pipeline streamlines the data processing workflow for the data generated by next-generation sequencing technologies.

Poster No: P128-F
Interplay between epigenetic changes and transacting factors regulating the CFTR promoter

[1]**Albertina De Sario**, [1,2,3]Celine René, [4]Marie-Elisabeth Brun, [1]Ahmed Squalli, [1,2,3]Anne Bergougnoux, [1,2]Marie-Catherine Romey, [1,2,3]Mireille Claustres and [1,2]Magali Taulan.
[1]INSERM U827, [2]CHU, [3]UM1, [4]IGH CNRS 1142, Montpellier, France

Mutations in the CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) gene are responsible for cystic fibrosis, an autosomic recessive lethal disease that is characterized by frequent pulmonary infections, a permanent inflammation, pancreatic insufficiency and male sterility. The CFTR gene is weakly expressed in epithelial tissues derived from the affected organs (lung, colon, pancreas, epidydime, sweat glands,…) and almost absent in other tissues. CFTR is not only regulated in a tissue-specific manner but also during development. Epigenetic changes have been suggested to be important for this tightly regulated expression (Blackledge et al., 2007 ; Paul et al., 2007) but no clear mechanism has been identified. Our aim is to investigate epigenetic factors responsible for the spatio-temporal regulation of CFTR expression.

Previously, we identified YY1 (Yin Yang 1) as a strong transcriptional repressor, which binds to the CFTR minimal promoter. YY1 could act by recruiting chromatin modifiers and DNA methyltransferases (Ko et al., 2008 ; Yao et al., 2001). By co-immunoprecipitation assays and chromatin immunoprecipitation experiments, we demonstrated that YY1 is able to recruit chromatin modifiers, such as CBP and HDAC, in the CFTR minimal promoter. This mechanism is accompanied by histone acetylation changes at the CFTR promoter region.

Moreover, to determine whether DNA methylation in the promoter region correlates with gene expression, we analyzed DNA methylation at high resolution using bisulphite and genomic sequencing. We found that DNA methylation in the minimal promoter of the CFTR gene ranges from 12% to 98% in various human cell lines. Currently, we are analyzing DNA methylation by pyrosequencing in 20 normal tissues including those that are affected in cystic fibrosis. To decipher the molecular mechanism governing this epigenetic change, it will be interesting to evaluate the YY1 role in

the methylation of the CFTR promoter. All together, these results provide further inside into the mechanisms responsible for the regulation of CFTR expression.

Poster Np: P129-T
BRCA1 and BRCA2 Germline Mutations in Korean Breast Cancer Patients: Spectrum of Mutations and Importance of Whole Gene Sequencing

Ja-Hyun Jang[1], Sollip Kim[1], Min-Jung Kwon[1], Seung-Tae Lee[1], Jeong Eon Lee[2], Seok Jin Nam[2], Jung-Hyun Yang[2], Chang-Seok Ki[1], Jong-Won Kim[1]
[1]Department of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea; [2]Division of Breast and Endocrine Surgery, Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

The frequency and spectrum of germline mutations in BRCA1 and BRCA2 genes vary according to ethnic groups and founder mutations were identified in some ethnic groups such as Ashkenazi Jews, Icelander, Norwegians, and European groups. In Korean population, although a few reports on the BRCA1/BRCA2 germline mutation frequencies are available, most of them used screening methods such as F-CSGE or DHPLC so that it is difficult to know real frequency of BRCA mutations. Here we performed a comprehensive mutation analysis for 265 Korean breast cancer patients by using whole gene sequencing covering all exons and their flaking intronic regions of BRCA1 and BRCA2 genes. Overall, 50 patients (18.9%) carried 37 deleterious mutations including 15 novel mutations. Among them, four BRCA1 mutations (c.3627dupA, c.5445G>A, c.1511dupG, c.5496_5506del11insA) and 3 BRCA2 mutations (c.97G>T, c.7480C>T, c.6952C>T) accounted for 20/50 (40%) of all mutations detected in this group of patients. It is notable that the estimated mutation detection rate in our study is significantly higher than those of previous studies using screening methods (2.5~11%) and similar to that of recent study using whole gene sequencing (18.4%). Therefore, whole gene sequencing seems the method of choice for detecting BRCA1/BRCA2 gene mutations in Korean population considering that significant number of mutations would be missed by using screening method. And, a two-tiered approach could be suggested for cost-effectiveness by screening of above recurrent mutations followed by whole gene sequencing in cases without these mutations.

Poster No: P130-F
PLaS-GWAS and functional analysis identified PDE4B as a novel susceptible gene for high-grade myopia

Changqing Zeng[1], Chen Wei[1], Xiangtian Zhou[2], Chaohua Li[1], Jia Qu[2]
[1] Beijing Institute of Genomics, Chinese Academy of Sciences; [2] Wenzhou Medical College

Myopia is the most common form of vision disorder worldwide. High-grade myopia (HM), defined as an axial eye length of >26 mm or a refractive error of <-5.00 diopters, is also designated as ‘pathologic’ myopia. Previous pedigree analysis had identified several susceptible loci for high grade myopia. However, with its high heterogeneity, the genetic mechanism underlying HM still remains largely uncertain.

To identify the susceptible genes at a population level, we conducted a cost efficient association analysis and named our strategy as PLaS-GWAS (Pooling Large Samples and Genome Wide Association Study). Briefly, DNA samples from 407 cases with <-8.00 diopter in both eyes and 294 controls with >-0.50 diopter were accurately measured with a NanoDrop 3300 fluorospectrometer. Equal amount of DNA from each sample of case and control groups was pooled together respectively. Genotyping was then performed using Illumina 1M duo chip with 3 replicate arrays for each pool. The allele frequency of each locus was obtained by computing the averaged ratio of signal intensity of each allele. After association analysis using Z-test, promising SNPs (p values ~10-5) were genotyped individually by MassArray in most samples to verify PLaS-GWAS discovery. These two steps resulted in several loci that appear to be associated with HM, including MYP3 locus, PDE4B, PHF10, 11q14.1, and PODN. Among these, PDE4B is a novel candidate involved in HM pathogenesis (PLaS-GWAS, p=1.3*10-6; individually genotyping, p=0.02). Interestingly, PDE4B is a member of type IV cyclic nucleotide phosphodiesterase family. The gene product of PDE4B specifically hydrolyzes cAMP which is actively involved in eye development. Next, animal model was utilized to further explore the possible role of PDE4B in HM formation. When one eye of Guinea Pigs was induced to myopia by form deprivation, PDE4B expression was reduced in sclera in comparison with the fellow eye. Furthermore, after the subconjunctival injection of pde4 inhibitor for 4 weeks, the normal guinea pigs developed a significant amount of myopia.

Our results strongly suggest that the PDE4B gene plays an important role in HM pathogenesis. In addition, PLaS-GWAS showed as an effective and cost efficient way to perform large scale genome wide association study.

Poster No: P131-W
An antioxidant cocktail rescues neurodegeneration in an X-adrenoleukodystrophy mouse model: therapeutic implications

Aurora Pujol[1,2,3,6,] Stéphane Fourcade,[1,2,3] Jone López-Erauskin [1,2,3] , Montse Ruiz[1,2,3], Agatha Schlüter[1,2,3], Nathalie Launay [1,2,3] ,Isidre Ferrer,[3,4] Reinald Pamplona,[5] Manuel Portero-Otín[5]
1 Neurometabolic Disease Lab, Institut d'Investigació Biomédica de Bellvitge (IDIBELL), Hospitalet de Llobregat, Barcelona, Catalonia, Spain, 2 CIBER de Enfermedades Raras (CIBERER) Spain, 3 Institut de Neuropatología de Bellvitge, Universitat de Barcelona, L'Hospitalet de Llobregat, Barcelona, Catalonia, Spain, , 4 CIBERNED Spain, 5 Departament de Medicina Experimental, IRB Lleida, Catalonia, Spain, 6 Catalan Institution of Research and Advanced Studies (ICREA) Barcelona, Spain.

X-linked adrenoleukodystrophy (X-ALD) is the most frequent inherited monogenic demyelinating disease (minimal incidence 1:17,000, OMIM 300100). X-ALD leads to neuroinflammation and demyelination leading to death in boys or to axonal degeneration in spinal cords in adults (adrenomyeloneuropathy or AMN). The disease is caused by loss of function of the ABCD1 gene, a peroxisomal ATP-binding cassette transporter that imports very long-chain fatty acids (VLCFA) into peroxisomes for degradation. The mouse model for X-ALD (Ald knock-out) exhibits a late-onset phenotype closely related to adrenomyeloneuropathy, with neurodegenerative features beginning at 15 months of age (1,2). Recently, we have found that oxidative damage might be a major contributor to disease pathogenesis in the mouse. Indeed, Abcd1 deficient mice present accumulation of MDAL (malonaldehyde-lysine), a consequence of lipoxidative damage to proteins, in spinal cords as early as 3.5 months of age. At 12 months, Abcd1- mice accumulate additional oxidative damage products arising from metal-catalyzed oxidation and glycoxidation/lipoxidation to proteins. Importantly, we observed that VLCFA excess is able to generate ROS and decrease reduced glutathione levels. Using gas chromatography/mass spectrometry - tocopherol analog Trolox is able to reverse oxidative?we have found that the lesions in vitro, thus providing therapeutic hope (3). In addition, we have treated human fibroblasts with a cocktail of 3 different antioxidants, which have been proven to prevent C26:0 dependent-ROS production in vitro. A preclinical test to treat an ALD mouse model for 5 months with a cocktail of these 3 antioxidants has normalized oxidative lesions in spinal cords, and has reversed immunohistochemical signs of axonal degeneration, such as synaptophysin, and APP accumulation in axonal swellings, together with microgliosis and astrocytosis. Further, the cocktail prevents and reverts the impairment of locomotor and neurobehavioural abilities that characterize the disease, thus providing the conceptual proof for oxidative stress as the major causative factor leading to neurodegeneration in X-adrenoleukodystrophy. Results will be translated in a clinical trial for X-adrenoleukodystrophy patients.

(1) Pujol et al, Hum Mol Genet. 2002 Mar 1;11(5):499-505; (2) Pujol et al, Hum Mol Genet. 2004 Dec1;13(23):2997-3006 ; (3) Fourcade et al, Hum Mol Genet 2008. Jun 15;17(12):1762-73

Poster No: P132-T
SAQC: SNP Array Quality Control

[1]**Hsin-Chou Yang** and [1]Hsin-Chi Lin
[1]*Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan*

[Background:] Genome-wide single-nucleotide-polymorphism (SNP) arrays providing hundred of thousands of SNPs on human genome have been broadly applied to study important human genetic/genomic topics. Data quality of SNP arrays plays an important role on the accuracy and precision of down-stream data analyses. However, good quality indices still await development.

[Results:] We propose new quality indices to measure data quality of SNP arrays, and statistical properties of the quality indices are investigated. The indices quantify a departure of the estimated individual-level allele frequencies from the expected ones via Mahalanobis distances. The proposed quality indices are shown, through empirical studies of several

large genomic projects, to follow log-normal distributions with different parameters characterized by ethnic populations. Databases of allele frequency references and quality index references for different SNP array platforms based on samples in various reference populations are established. Furthermore, a confidence interval method, based on the underlying empirical distributions, is developed to identify SNP arrays or DNA samples with poor quality. Interestingly, the method can also be applied to identify chromosome aneuploidies. Analyses of authentic biological data and simulated data show that our method has high sensitivity and specificity for detections of poor SNP arrays and aberrant chromosomes.

[Conclusions:] This paper develops computationally efficient and practically useful measures of data quality and detectors of poor SNP arrays. SNP Array Quality Control (SAQC) written in R and R-GUI is developed as a user-friendly tool for evaluation and visualization of data quality of genome-wide SNP genotyping studies.

Poster No: P133-F
DYNAMICS AND SUPRANUCLEOSOMAL ORGANIZATION OF THE MAMMALIAN CHROMATIN FIBER : A CONTRIBUTION OF THE 3C-qPCR METHOD

Thierry FORNE, Franck COURT, Julie MIRO, Caroline BRAEM, Marie-Noëlle LE LAY-TAHA, Audrey BRISEBARRE, Florian ATGER, Thierry GOSTAN, Michael WEBER and Guy CATHALA
Institut de Génétique Moléculaire de Montpellier, UMR 5535 CNRS-UMII, 1919 Route de Mende, 34293 Montpellier cedex 5, France.

In interphasic cells, the mammalian genome, packed into the chromatin fiber, is spatially restrained to specific chromosomal territories. However, beyond the simple nucleosomal array, very little is known about the organization and the dynamics of the chromatin fiber within chromosomal territories. Although it is largely admitted that one essential determinant is chromatin-looping in relation with gene expression and other chromosomal activities, the basic structural landscape of the chromatin remains largely unknown at the supranucleosomal level. Indeed, at that scale (~10-350 kb), the organization of the chromatin fiber and its dynamics in living cells are difficult to access by cell imaging techniques due to intrinsic limitations of light microscopes. The Chromosome Conformation Capture (3C) assay, and derived technologies, represents a real technological advance. Indeed, such techniques allow direct quantification of the average interaction frequency between two distant genomic regions, at the supranucleosomal scale, in their native genomic context.

Our laboratory contributed to the recent development of the 3C technologies by improving the quantification of the interaction frequencies (3C-qPCR method) (Hagège et al., 2007 Nat. Protocols 2, 1722).

- In a first approach, we applied this method to analyze long-range chromatin interactions at the Dlk1/Gtl2 (Braem et al., 2008 J. Biol. Chem. 283, 18612) and Igf2/H19 imprinted loci. We identified several locus-specific looping interactions, thus bringing original insights into the complex mechanisms of gene regulation at both loci.

- In a second approach, we have determined random collision frequencies occurring between genomic sites separated by increasing genomic distances and sought to establish whether some fundamental intrinsic constraints apply to the supranucleosomal chromatin fiber. We demonstrate that, in the absence of long-range locus-specific interactions, gene-rich domains of the mammalian chromatin fiber fold into a basic statistical helix. Using bioinformatics analyses, we show that conserved sequences at co-regulated gene loci are highly overrepresented at genomic distances corresponding to one and two helix turns relative to transcription start sites. Therefore, the basic statistical helix organization of the chromatin affects the dynamics of long-range genomic interactions and may contribute to mammalian genome evolution.

Poster No: P134-W
An analysis of CFTR mutations in Korean cystic fibrosis patients: Summary of eight known and five novel mutations

Haiyoung Jung, and Chang-Seok Ki
Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 135-710, Korea

Introduction: Cystic fibrosis (CF) is the most common fatal inherited disorder among Caucasians. CF has been recently recognized more commonly within many nonwhite populations. However, little has been reported on its occurrence in Southeast Asians including Koreans. In this study, we have identified 8 known and 5 novel cystic fibrosis transmembrane conductance regulator (CFTR) mutations from 8 Korean CF patients.

Material and method: We have analyzed the complete coding regions and flanking intronic sequences of the CFTR gene among 8 unrelated CF patients (2 males, 6 females). To rule-out large deletions or gene rearrangements, the multiplex ligation-dependent probe amplification (MLPA) analysis was performed in one patient detected with only one known mutation.

Results: We have identified 8 previously known mutations, namely Q98R, Q220X, Q1291X, Q1352H, 579+5G>A, IVS8-T5, 2623-?_2751+?del, and 3272-26A>G. Q98R was the most commonly identified known mutations with frequency of 37.5% (3/8 alleles). No delta F508 mutation, which is known to be the most common mutation among Caucasians, was detected in Korean patients of this study. The newly identifed mutations, some of which has already been published as case reports, include L441P, D979A, 1766+2T>C, 2052delA, and 3908insA.

Conclusion: This study is the first to collectively summarize the CFTR mutational spectrum of Korean CF patients. Detection of no delta F508 mutation, and five novel and seven rare mutations reflects a heterogeneous spectrum of CFTR mutations in Korean CF patients, which might be different from that of Caucasian populations. Therefore, more specialized practice guidelines and management of data-base for molecular diagnosis of CF are necessary for Southeast Asians including Korean population.

Poster No: P135-T
B-lymphocytes Genomic Analysis in Chronic Lymphocytic Leukemia Patients

Sarka Pospisilova, Karla Plevova, Katerina Stano Kozubik, Sarka Pavlova, Marek Mraz, Jitka Malcikova, Hana Jelinkova, Marek Borsky, Boris Tichy, Michael Doubek, Jiri Mayer
Center of Molecular Biology and Gene Therapy, Department of Hematooncology, University Hospital Brno and Medical Faculty, Masaryk University, Brno, Czech Republic

Chronic lymphocytic leukemia (CLL) is a lymphoproliferative disorder characterized by a clonal expansion of mature CD5+ B-lymphocytes and represents the most frequent leukemia in the Western world. Its clinical course is very variable depending on several prognostic markers, including genomic complexity, mutational status of immunoglobulin heavy chain variable region (IgVH) and p53 functionality. Chromosomal abnormalities are present in majority of CLL cases (>80%) and are routinely analyzed using CLL-specific FISH panel (probes for 13q14, 17p13, 11q22 deletions and chromosome 12 gains). Deletion of 13q14 region, coding DLEU genes, miR-15a, miR-16-1 and RB1, is the most frequent (in approx. 60% of cases) and its prognostic impact depends on the range of deletion and number of involved alleles. Recently, a novel deletion of 22q11 was detected in CLL patients, this locus encodes Lambda immunoglobulin light chain subgenes (IgL-Lambda). Interestingly, several protein-coding genes (PRAME, ZNF280A, ZNF280B, GGTLC2) and microRNA are localized within V subgenes of variable IgL-Lambda region, in contrast to IgL-Kappa and IgH locus. These protein-coding genes and microRNA are alternatively deleted in process of immunoglobulin gene splicing during B-lymphocyte development and could play a role in CLL pathogenesis.

We analyzed genomes of 40 CLL patients using array-CGH (Human Genome CGH Microarray 4x44K, Agilent); FISH data were available for all these patients. Additional aberrations not detected by FISH were found in the majority of samples (87,5 %). Deletion of 13q14 was observed in 60% of patients (24/40) and in all those cases included miR-15a and miR-16-1. Monoalelic 13q14 deletion as a sole aberration suggesting a good prognosis occurred in 9 cases. Detail characterization of 22q11 locus with deletions of approx. 0.77 Mb in 7 cases, demonstrated IgL-Lambda gene recombination, which was confirmed by PCR analysis (BIOMED-2 protocol). Additional IgL-Lambda rearrangements, not detected by array-CGH, were assessed in 9 patients using the PCR method (in total, 16 patients with IgL-Lambda deletions were detected).

Interestingly, one of IgL-Lambda variable segment (V2-8) includes an annotated microRNA gene miR-650. Also other subgenes of V2 family include homologues for miR-650, which overlay the leader exon of V2 subgenes. IgL-Lambda rearrangement led to miR-650 deletion in 87,5 % of cases (14/16). We detected high expression of this miRNA (RT-qPCR, TaqMan Assay, ABI) in two cases, where IgL rearrangement uses a subgen from V2 family. It implies that miR-650

expression depends on the presence and the type of IgL-Lambda rearrangement and miR-650 could be involved in B-lymphocyte development.

To conclude, whole genome array-CGH has proved wide application in clinical diagnostics and brought more precise characterization of CLL cells than routinely used FISH. These methods are useful for assessment of complex genomic changes during course of disease and bring new insight into CLL pathogenesis and biology.

This work was supported by research grants MSMT MSM0021622430 and IGA MZ CR NS10439-3/2009.

Poster No: P136-F
Gene-based approach for genome-wide association studies

[1]**Nam Yunsun**
[1]KNIH, Korea, Republic

In most genome-wide association studies (GWAS), single-locus case-control comparisons are used to identify single nucleotide polymorphisms associated with a disease. However, there are several limitations in these methods. For example, these methods are underpowered to detect small risk effects and it has difficulty in understanding and interpreting significant SNPs with a biological meaning. To overcome the limitations, many new methods have been suggested [1,2,3]. Biological functions of genes are much better investigated than SNPs, so it will be easier to understand significant genes with a biological theme.

Here we suggest two method for measuring the significance value of a gene. One is called minP method and the other is called Fisher's combination method. We have applied our methods to the data set consisting of 1,042 T2DM (Type 2 diabetes mellitus) patients and 2,943 healthy controls from KoGES study, and present the results of two methods.

Poster No: P137-W
Population-prevalence of Finnish desmosomal mutations predisposing to arrhythmogenic right ventricular cardiomyopathy

[1,2]**Annukka M. Lahtinen**, [1,2]Annukka Marjamaa, [3]Maija Kaartinen, [3]Tiina Heliö, [3]Lauri Toivonen, [3]Heikki Swan, [4]Veikko Salomaa, [5,6]Eero Lehtonen, [1,2]Kimmo Kontula
[1] Research Program for Molecular Medicine, Biomedicum Helsinki, University of Helsinki, Helsinki, Finland [2] Department of Medicine, University of Helsinki, Helsinki, Finland [3] Department of Cardiology, University of Helsinki, Helsinki, Finland [4] National Institute for Health and Welfare, Helsinki, Finland [5] Department of Pathology, University of Helsinki, Helsinki, Finland [6] Research Animal Centre, University of Helsinki, Helsinki, Finland

Arrhythmogenic right ventricular cardiomyopathy (ARVC) is a severe cardiac disorder caused mainly by dominant mutations of the desmosomal cell adhesion proteins plakophilin-2 (PKP2), desmoplakin (DSP), desmoglein-2 (DSG2), and desmocollin-2 (DSC2). Upon search of 29 Finnish ARVC probands, we have previously identified PKP2 missense mutations in 10% of the patients. In order to further characterize the genetic basis of ARVC in Finland, we screened the probands for three other desmosomal genes and conducted a search for ARVC-associated mutations in a large Finnish population sample (Health 2000 Study, N = 6300). A novel frameshift mutation (E1020fsX1037) of DSG2 appeared in one ARVC proband, and another patient carried a DSP variant, T1373A, whereas DSC2 was not mutated in any of the 29 patients. Multiplex ligation-dependent probe amplification (MLPA) did not reveal any deletions or duplications in the desmosomal genes. Immunohistochemical analyses showed a reduced amount of both desmoglein-2 and plakophilin-2 in the endomyocardial samples of the DSG2 mutation carrier, indicating disturbance of the desmosomal protein interactions. In the population sample, a total of 31 subjects (0.5% [95% CI 0.33-0.71%]) carried a desmosomal gene mutation, PKP2 Q59L being the most prevalent form (n = 19). The penetrance of ARVC was substantially reduced as only a subset of these mutation carriers showed ARVC-related arrhythmic and electrocardiographic features. In conclusion, up to 1 in 200 Finns carry a mutation with a proposed risk of ARVC. However, our findings suggest that penetrance of a typical ARVC phenotype may require additional (genetic or nongenetic) factors in addition to the mutant desmosomal gene itself.

Poster No: P138-T
High resolution genomic profiling of xenografted human Gliomas to delineate non-angiogenic and highly angiogenic phenotypes in a clinically relevant model system.

Daniel Stieber[1], Per-Øestein Sakariassen[2], Rolf Bjerkvig[1,2] and Simone P. Niclou[1]
[1]*NorLux Neuro-Oncology Laboratory, CRP-Santé L-1526 Luxembourg and* [2]*NorLux Neuro-Oncology, Department of Biomedicine University of Bergen, N-5020 Bergen, Norway*

Glioblastoma multiforme (GBM) is the most common form of malignant brain tumor in adults. Patients with GBM have a uniformly poor prognosis, with a median survival of one year thus, advances on all scientific and clinical fronts are needed. We have developed a human glioblastoma xenograft model in immunodeficient rodents that is characterised by a highly infiltrative non-angiogenic phenotype. Upon serial transplantation this phenotype will develop into a highly angiogenic tumor. Thus, we have developed an animal model where we are able to establish two characteristic tumor phenotypes that define human glioblastoma (i.e. diffuse infiltration and high neovascularization). It is well established that the cancer genome is moulded by the dual processes of somatic mutation and selection. In order to assess whether the observed phenotypic shift is due to clonal selection in vivo, we have performed high-resolution aCGH on primary tumors and xenografts derived thereof. We show that although the overall genomic pattern is highly conserved, additional genomic events may be involved in the angiogenic switch observed in vivo. We are currently further analysing our findings by applying whole exome resequencing to the analysed samples in order to delineate the mutational evolution of xenografted gliomas at single nucleotide resolution and identify new mutational events linked to the phenotypic switch. `This molecular dissection of the two hallmarks of GBM could lead to the identification of potential biomarkers and will facilitate the elucidation of the molecular pathways involved in the switch from invasive to angiogenic growth, thereby potentially opening new diagnostic and treatment avenues in the clinic.

Poster No: P139-F
A Novel Gene Mapping Approach for Integration of GWA and eQTL Studies

[1]**Koichiro Higasa**, [1]Natsuhiko Kumasaka, [2]Kazuharu Misawa, [1,3]Yukinori Okada, [4]Michiaki Kubo, [5,6]Yusuke Nakamura, and [1]Naoyuki Kamatani
[1]*Laboratory for Statistical Analysis, Center for Genomic Medicine,* [2]*Research Program for Computational Science, Research and Development Group for Next-Generation Integrated Living Matter Simulation, Fusion of Data and Analysis Research and Development Team,* [3]*Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo Japan,* [4]*Laboratory for Genotyping Development,* [5]*Center for Genomic Medicine, RIKEN, Kanagawa, 230-0045, Japan,* [6]*Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo, Tokyo 1088639, Japan*

Genome-wide association (GWA) studies are a powerful and efficient approach to identify genetic variants associated with complex human traits. GWA studies became feasible as a result of several key advances in genetic knowledge, genotyping technologies, statistical analysis algorithms, and the availability of large collections of samples. Although GWA studies have successfully identified important genetic variants associated with many human traits, including diseases, the identification of functional connections between variants and phenotypes remains the major barrier to interpretation of GWA study results. Here we developed a novel gene mapping approach that integrates GWA and genome-wide expression-QTL (eQTL) data to simultaneously interpret both datasets under a single statistical framework. We applied our approach to several human diseases and their hematological, biochemical, and physical traits using publicly available expression data from lymphoblastoid cell lines. Our approach identified genetically and functionally associated variants in these datasets. The applicability of our approach is supported by the enrichment of previously identified genes involved in these phenotypes in our analysis. Additionally, we identified several new candidate genes with notable functional association to the phenotypes examined. Our method not only dramatically shortens the time required to discover functional SNPs but also has the potential to aid our understanding of the normal variation in gene expression, resulting from yet to be identified combinations of genetic polymorphisms involved in complex diseases and traits.

Poster No: P140-W
Association of genetic polymorphisms in human NR1I3 with hyperlipidemia in elderly Japanese population.

Shinobu Ikeda[1], Tomio Arai[2], Makiko Mieno[3], Noriko Tanaka[4], Motoji Sawabe[2], and Masaaki Muramatsu[1]
[1]Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan., [2]Department of Pathology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan., [3]Department of Medical Informatics, Center of Information, Jichi Medical University, Tochigi, Japan., [4]Department of Biostatisticas, Harvard School of Public Health, Boston, Ma, USA.

[Background]
The nuclear receptor constitutive androstane receptor (CAR; NR1I3) is involved in detoxification and regulation of xenobiotics and endogenous molecules. CAR has also been shown to be involved in the metabolism of hepatic bile acids and cholesterol. Such function of CAR gene has been established in mice, however the role of human CAR gene is not well known yet.

In this study, we performed search for genotype-phenotype associations in consecutive autopsies of elderly Japanese population, focusing on human CAR polymorphisms and metabolic phenotypes.

[Methods]
Six polymorphisms in human NR1I3 gene were genotyped using melting curve analysis or TagMan Assay. The study subjects (n=1536) were consecutive autopsies of elderly Japanese population registered in the Japanese SNPs for geriatric research (JG-SNP) database. Multiple logistic regression analyses were performed to examine the genetic effects and interaction effects on the metabolic disease. The estimates were adjusted for gender and age at death of the subjects; the presence of hypertension, hyperlipidemia, and diabetes mellitus; and the smoking and drinking status. The P-value under 0.05 was considered statistically significant. The SAS system for Windows ver.9.1.3 was used for all statistical analyses.

[Results]
We found that NR1I3 rs2307424 C-carrier (CC+CT) was associated with reduced risk of hyperlipidemia (p=0.0115). The odds ratio (OR) was 0.38 (95% confidence interval [CI], 0.182-0.806) after adjustment for significant risk factors such as age, gender, prevalence of hypertension and diabetes mellitus. Moreover, rs55802895 A-carrier (AA+AG) had increased risk of hyperlipidemia (OR=2.25, 95%CI 1.009-4.995, p=0.0474), after adjustment.

[Conclusion]
We provide data that CAR polymorphism is association with lipid profile in human. This is in agreement with the fact that mice CAR regulates cholesterol metabolism.

Poster No: P141-T
Integrative genomic analysis in ovarian cancer

[1]**Shakeela Banu**, [1]Cyril Dalmasso, [1]Kartiki V Desai, [2]Euan A Stronach, [2]Nona Rama, [2]Claudia Hayford, [2]Sadaf Ghaem-Maghami, [2]Mona El-Bahrawy, [1]Philippe Broët, [2]Hani Gabra, [1]Edison T Liu
[1]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138672, Singapore, [2]Ovarian Cancer Action (HHMT) Research Centre, Imperial College London Hammersmith Campus, Du Cane Road, London W12 ONN, UK

Chromosomal abnormalities such as rearrangements, copy number gains and losses are characteristic attributes of cancer often resulting in activation of oncogenes and inactivation of tumour suppressor genes. Epithelial ovarian cancer (EOC) is the leading cause of death from gynecological malignancy in women and is characterized by a high degree of chromosomal instability. In this study, we performed genomic and transcriptomic profiling of 125 benign, borderline and malignant ovarian tumours using Affymetrix SNP 6.0 and Affymetrix GeneChip U133 Plus 2 arrays to investigate the molecular changes associated with such instability. CGHmix was used to detect copy number alterations (CNA) and determine the frequencies of amplifications and deletions in all the samples. We then jointly evaluated the propensities of amplifications and deletions for each genomic locus to distinguish the exclusively amplified or deleted regions that are

most likely to be non-random chromosomal events and hypothesized that such events harbored driver genes. Correlation of copy number and gene expression changes in these regions enabled us to prioritize functionally relevant aberrations. We found that 43% of 52,000 probesets analyzed had a significant correlation between CNA and gene expression changes, suggesting that a large number of expression differences in tumours resulted from local copy number changes. Moreover, 27% of the genome was significantly predisposed to amplifications while 20% of the genome was significantly predisposed to deletions. In addition we mapped the occurrence of loss of heterozygosity (LOH) in the tumours and correlated LOH with gene expression changes to identify potential tumour suppressor genes that may be silenced by LOH. We found 37% of the genome bearing LOH in more than 10% of tumours. Subsequent overlap with clinicopathological parameters such as tumour grade and disease-free survival has identified our best candidate novel oncogenes and/or tumour suppressor genes. Gene-by-gene screening of these candidates using cell-based assays is ongoing which will reveal new insights into the role of genetic aberrations in ovarian cancer.

Poster No: P142-F
Gene-environment interactions influencing elevated serum C-reactive protein levels in a genome-wide association study of 8,837 Koreans

[1]**Ji Wan Park**, [2]Nam Han Cho, [3]Chol Shin, [4]Jong-Young Lee
[1]Department of Medical Genetics, College of Medicine, Hallym University Chuncheon, Korea, [2]Department of Preventive Medicine, College of Medicine, Ajou University, Suwon, Korea, [3]Department of Internal Medicine, Korea University Ansan Hospital, Ansan, Korea, [4]Center for Genome Science, Korea National Institute for Health, Seoul, Korea

Elevated C-reactive protein (CRP) level, an inflammatory marker, is considered to be a predictor of a variety of aging-related diseases such as cardiovascular disease. Serum CRP level is known to be influenced by a complex interaction of genetic and environmental factors. We conducted a genome-wide association study among 8,837 unrelated Koreans and tested for influence of gene-environment interaction on serum CRP level. We initially analyzed the association of serum CRP with various environmental and clinical variables to identify possible confounding variables. We analyzed natural log transformed CRP for investigating interactions between 352,228 single nucleotide polymorphism (SNP) markers and nongenetic factors including education, physical activity, cigarette smoking, alcohol drinking, history of disease, and history of medication. We included a term for the additive effect of genotype, a term for covariate, and a term for the interaction between the nongenetic risk factor in multiple linear regression models. Subsequently, all analyses were stratified by sex with adjustments for age, residence, body mass index, nine biochemical variables, and five lifestyle variables. We found significant evidence for interaction between multiple loci and nongenetic factors (e.g. rs7553007 of CRP gene and positive history of hepatitis (p=0.026), cigarette smoking, or insulin medication). While we consider these interaction analyses exploratory and the results must be confirmed in future studies, accounting for genotype by environment interaction might explain, in part, the problem of missing heritability that have been identified in previous genome wide association studies.

Poster No: P143-W
Loci controlling specific IgEs in Czech and Russian Populations

[1]**Elena S.Gusareva** , [1]Jana Badalová, [1]Helena Havelková, [2]Elena Ju. Bragina, [1]Hanna Blažková, [3]Svetlana N. Buinova, [3]Boris A. Chernyak, [2,4]Valery P. Puzyrev, [5]Evgenija V. Deeva, [5]Olga S. Fedorova, [5]Ludmila M. Ogorodova, [6,7]Petr Kučera, [8]Vlastimil Král, [1,6]Marie Lipoldová
[1] Institute of Molecular Genetics, Academy of Sciences of the Czech Republic v.v.i., Prague, Czech Republic [2] Research Institute of Medical Genetics, Siberian Branch of Russian Academy of Medical Sciences, Tomsk, Russia [3] Department of Allergology and Pulmonology, Irkutsk State Institute for Post-graduate Medical Education, Irkutsk, Russia [4] Siberian State Medical University, Department of Medical Genetics, Tomsk, Russia [5] Faculty of Pediatrics, Siberian State Medical University, Tomsk, Russia. [6] Third Faculty of Medicine, Charles University, Prague, Czech Republic [7] Allergology and Clinical Immunology, University Hospital KV, Prague, Czech Republic [8] Department of Immunology and Microbiology, Institute of Public Health, Ústí nad Labem, Czech Republic

Elevated production of IgE (atopy) is often associated with development of allergic diseases and is known to be a complex trait influenced by both genetic and environmental factors.

Aim of the project was to define genetic loci responsible for atopy in human. We combined the genome-wide study in mouse with the candidate locus approach in human.

In the genome-wide search for IgE-controlling loci in mouse recombinant congenic strains, several regions on chromosomes 1, 2, 3, 4, 5, 8, 10, 16 and 18 contributed to IgE regulation. From the conserved synteny between mouse and human genomes, we identified corresponding orthologous regions on human chromosomes. Majority of the orthologous regions have been already reported as atopy loci in humans, thus supporting the precision and the predictive power of our approach. However, the human region 8q12 orthologous for the shortest IgE locus Lmr9 (chromosome 4) in mouse has not been described in human studies of atopy. Therefore, in the 8q12 region, we selected three microsatellite markers D8S1828, D8S285, and D8S1816, and tested them for non-parametric linkage and association (QTDT) with levels of total IgE and specific IgEs to twenty inhalant and five food allergens in Czech atopic nuclear families. In the position marked by D8S285 (57.22 Mb, 71cM) we demonstrated the novel human IgE-controlling locus exhibiting suggestive linkage to composite inhalant allergic sensitization and to nine specific IgEs.

We also tested the effects of several the most promising candidate chromosomal regions reported by other scientific groups on atopy and associated traits in Czech and Russian atopic families. Linkage and association (QTDT) to plant-specific IgEs were identified at loci 5q33, 7p14, 12q13 and 13q14 in the Czech population. In Russian family group, cat-specific IgE (which is the most abundant in Russian patients with asthma from Siberia) showed suggestive linkage with the 12q24.3 region.

Poster No: P146-T
An Enhanced International Fanconi Anemia Registry (IFAR)

[1,2]**Arleen D. Auerbach**, [2]Agata Smogorzewska, [2]Francis Lach , [3] Matthew J. Wrobel, [3]Mousumi Sengupta, [3]Edward Barbour
[1]Human Genetics and Hematology, [2] Laboratory of Genome Maintenance, [3]Hospital Informatics, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

Fanconi anemia (FA) is a rare recessively inherited disorder characterized by genome instability, DNA crosslink hypersensitivity, congenital malformations, bone marrow failure, and predisposition to malignancy. At least 13 FA genes have been identified. The International Fanconi Anemia Registry (IFAR) was established at The Rockefeller University in 1982 in order to collect data regarding patients with FA and their families. In addition, the Fanconi Anemia Mutation Database, http://www.rockefeller.edu/fanconi/mutate/, was established in 1998 to accelerate the availability of information on mutations in these 13 important cancer-predisposing genes.

The new IFAR project proposes to develop a comprehensive, ontology-driven Phenotype Recording Instrument (PRI) and database for FA. The PRI will integrate the existing IFAR data held in multiple disparate data sources for a unique look at the FA patient population, which includes the patient data available in the current IFAR database, the polymorphic (SNP) and mutation data available for patients representing the known FA genes, as well as pedigree information showing the inheritance patterns from multiple generations. Furthermore, the NIH's Genomics Center (NHGRI) is currently sequencing certain patient tissue samples using Next Generation technology, which will further enhance the current IFAR patient data. The new development effort will also include "deep phenotyping" from new patient history questionnaires. We are hypothesizing the addition of these two data stores, full gene sequencing and deep phenotyping, will add significantly to the understanding of FA while providing a deeper basic scientific understanding of DNA repair mechanisms, and more prognostic ability for the current patient population.

The enhanced data model resulting from the aforementioned data sources will all be modeled using an OWL ontology defined with Protégé. We are committed to a more elaborate semantic query mechanism for data mining of this comprehensive and "state of the art" data warehouse for FA patients. We will build data mining and visualization tools that will be required to maximize the research view of this integrated patient data. As the name of the existing IFAR system implies, this registry is an International collaboration. The new system will enable worldwide international collaboration between FA researchers. To accomplish this, we will build a web-based system utilizing the Java JSP front-end coupled with an Oracle 10g database backend for performance and data integrity. This is truly a unique medical informatics opportunity made possible because of the 28 year collection of FA data combined with next generation sequencing and deep phenotyping data.

Poster No: P147-T
Genomic Sequence of a mutant strain of Caenorhabditis elegans with an altered recombination pattern

Ann M. Rose[1]*, Nigel J. O'Neil[1], Mikhail Bilenky[2], Yaron S. Butterfield[2], Nawar Malhis[2], Stephane Flibotte[2], Martin R. Jones[1], Marco Marra[2], David L. Baillie[3] and Steven J.M. Jones[2]
[1]Department of Medical Genetics, University of British Columbia, 419 – 2125 East Mall, Vancouver, BC, V6T 1Z4, Canada [2]Genome Sciences Centre, British Columbia Cancer Research Centre, 600 West 10th Avenue, Vancouver, British Columbia, Canada V5Z 4E6 [3]Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, B.C. V5A 1S6

The original sequencing and annotation of the Caenorhabditis elegans genome along with recent advances in sequencing technology provide an exceptional opportunity for the genomic analysis of wild-type and mutant strains. Using the Illumina Genome Analyzer, we have sequenced the entire genome of Rec-1, a strain that alters the distribution of meiotic crossovers without changing the overall frequency. Rec-1 was derived from ethylmethane sulfonate (EMS)-treated strains, one of which had a high level of transposable element mobility. Sequencing of this strain provides an opportunity to examine the consequences on the genome of altering the distribution of meiotic recombination events. Our analysis of high-throughput sequencing was able to detect regions of direct repeat sequences, deletions, insertions of transposable elements, and base pair differences. A subset of sequence alterations affecting coding regions were confirmed by an independent approach using oligo array comparative genome hybridization. The major phenotype of the Rec-1 strain is an alteration in the preferred position of the meiotic recombination event with no other significant phenotypic consequences. In this study, we observed no evidence of a mutator effect at the nucleotide level attributable to the Rec-1 mutation.

Poster No: P148-F
Mitochondrial genomics in Indigenous Asians

ME Phipps, LC Hong and MA Abdulla
Monash University

The unique properties of mitochondrial DNA lie in the fact that their inheritance is uniparental. Thus, its analyses provide critical information regarding on the origin of species, population genetics and forensic studies. Malaysia with its distinct multiracial population is one of the most important regions for studying genetic diversity of human populations, in terms of evolution and diseases. However, a significant number of Malaysia's distinct ethnicities especially the indigenous groups have been underrepresented in many scientific reports. The efforts to collect samples and generate data especially for mtDNA markers in these groups are lacking. This study represents a pioneering effort to investigate mtDNA variants for four Malaysia indigenous groups and has resulted in new knowledge. Mitochondrial DNA HVS-1 and -2 sequences and RFLP polymorphisms were genotyped and analyzed in 188 individuals from four indigenous groups, Jehai, Kensiu, Temuan and Bidayuh. Mitochondrial haplotypes, and haplogroup frequency distributions were determined in order to characterize the mtDNA structure and diversity in Malaysia indigenous group senario. A total of 64 mtDNA haplotypes were observed. The indigenous groups exhibited low diversity reflecting enhanced genetic drift, bottle neck or founder effects in these small, isolated populations. The majority of mtDNA haplogroups were found to be exclusive and distinct to them. Most mtDNA variants belonged to three major haplogroups M, N and R. The most frequent were haplogroups included M21a and R21, which both reflect very old lineages, deeply rooted in South East Asian population's lineages (~60,000 years ago). This was evident in about 80% of the Negritos (subgroups of Jehai and Kensiu) that were genotyped. The mtDNA data indicated very old variants suggesting that the Negritos in Malaysia are descendants of the earliest inhabitants in South East Asia. This invetigation also indicated some admixture of our indigenous groups with other Southeast Asian ethnic groups. The low level of gene flow among these groups supports the Out of Africa hypothesis. In addition, 7 new mutations were discovered and have been reported to the international Mitomap database.

POSTERS INDEX

AUTHORS INDEX

POSTER KEYWORD INDEX

Register at **www.ispmontreal2010.com**

The **6th International Congress of Pathophysiology** will enlarge the scope of the meeting to reflect current trends in the evolution of pathophysiology and biological science such as the progress in genomics. The **ISP 2010** will present novel therapeutic targets based on this new knowledge, future tools in prevention, and optimized therapy based on predictive genomics.

The **14th International SHR Symposium** will offer the best opportunity for researchers and clinical to meet and discuss critical issues on the ecogenomics of cardiometabolic diseases.

Pavel Hamet, OQ, MD, PhD, FRCPC, FAHA, FRSM, FCAHS
President of the 6th International Congress of Pathophysiology
and of the 14th International SHR Symposium

A balanced and dynamic scientific program: Lectures, Symposia and Workshops

"Gene-environment interaction in health and disease" and "Ecogenomic models of cardiometabolic diseases"

More than 80 renowned speakers, including:

- **Allen ATTIE** (USA)
- **Alexei BAGROV** (USA)
- **Albert L. BARABASI** (USA)
- **Giuseppe BIANCHI** (Italy)
- **Michel BOUVIER** (Canada)
- **Allen W. COWLEY** (USA)
- **Alan DENG** (Canada)
- **Anna DOMINICZAK** (UK)
- **Victor DZAU** (USA)
- **Noboru FUKUDA** (Japan)
- **Aron M. GEURTS**(USA)
- **Richard E. GILBERT** (Canada)
- **Qide HAN** (China)
- **Osmo HANNINEN** (Finland)
- **Norbert HÜBNER** (Germany
- **Hiroshi ITOH** (Japan)
- **Shokei KIM-MITSUYAMA** (Japan)
- **Jaroslav KUNES** (Czech republic)
- **Bernard LÉVY** (France)
- **Peter LIU** (Canada)
- **Tak W. MAK** (Canada)
- **Tomoji MASHIMO** (Japan)
- **Bruce MCMANUS** (Canada)
- **John MULLINS** (UK)
- **Monica NERI** (Italy)
- **Tommy NILSSON** (Canada)
- **Marju MELANDER** (Sweden)
- **Alexander MONGIN** (USA)
- **Michal PRAVENEC**(Czech republic)
- **Mark PRENTKI** (Canada)
- **Allen D. ROSES** (USA)
- **Laszlo ROSIVALL** (Hungary)
- **Guy ROULEAU** (Canada)
- **Rafik SÉKALY** (Canada)
- **Kazuaki SHIMAMOTO** (Japan)
- **Vladimir STERBAK** (Slovakia)
- **Moshe SZYF** (Canada)
- **Miyako TAKAKI** (Japan)
- **Yoram YAGIL** (Israel)
- **Barry ZIRKIN** (USA)

ISP & SHR 2010
22-25 September 2010, Palais des Congrès de Montréal
For more information visit **www.ispmontreal2010.com**

N O T E S