Genomic Med. (2008) 2:401–413
DOI 10.1007/s11568-009-9108-3

ABSTRACTS

Micro RNA and non-coding DNA and repeats

513: In silico approach towards discovering novel small molecules as potential inhibitors of oncogenic microRNA processing

[1]**Saurabh Agrawal**, [2]Vinod Scaria, [1,2]Souvik Maiti

[1]Proteomics and Structural Biology Unit, Institute of Genomics and Integrative Biology, Mall Road, Delhi, India, [2]GN Ramachandran Knowledge Center for Genome Informatics, Institute of Genomics and Integrative Biology, Mall Road, Delhi, India

MicroRNAs are small (~22 nucleotide) noncoding functional RNAs that has been recently implicated in a wide variety of disease processes and critical biological processes including development and host-pathogen interaction. The regulatory role of microRNAs in health and disease is just emerging. Oncogenesis is one of the major biological processes now known to be mediated by microRNAs. MicroRNAs are processed from imperfect palindromic sequences in the genome and undergo processing in association with a series of proteins before finally getting incorporated into a ribonucleoprotein complex (RISC) which mediates post-transcriptional repression of gene expression. We propose that modulation of microRNA processing would be a novel strategy to inhibit the function of a microRNA, especially for specific microRNAs implicated in oncogenesis. We compiled a list of microRNAs differentially expressed in various cancers. Many of these microRNAs were found not to be exclusively dysregulated in cancer. To find microRNAs exclusively dysregulated in cancers, we re-analyzed human small RNA expression data in public domain derived from high-throughput sequencing technologies. We found hsa-mir-372 is an oncogenic microRNA exclusively expressed in germ-cell tumors. Literature evidence also supports the direct functional role of this microRNA in the pathophysiology of cancer. We compiled a library of known RNA-binding ligands from published literature and used in silico screening of hsa-mir-372. We compiled a list of small molecules that potentially bind specifically to microRNA hairpin bulges, which are important for microRNA processing from hairpin forming pre-miRNAs. We used a consensus approach using Sybyl and Dock6 and used stringent cutoff criteria for the screening of the ligands. Further the binding interaction of the hsa-mir-372-ligand complexes were determined through molecular dynamics simulations followed by energetic calculations.

514: piRNABank: a tool to annotate and analyze functional relevance of piwi interacting RNAs at genome wide scale

Shipra Agrawal, Prakher Singhal, Sailakhsmi Subramanyam

Institute of Bioinformatics and Applied Biotechnology, ITPL, Whitefield Road, Bangalore, India

Piwi-interacting RNAs (piRNAs) are a class of small interfering RNA, expressed in germline cells and have been identified as key players in germline development. These molecules, typically of length 25–33 nucleotides, associate with Piwi proteins of the Argonaute family to form the Piwi-interacting RNA complex (similar to RISC complex), which regulates a number of processes like spermatogenesis, repression of retrotransposon transposition in germline cells, epigenetic regulation and positive regulation of translation and mRNA stability. Like miRNAs, piRNAs also exist as clusters and probably act on a target in a synergistic way as there is several more piRNAs than miRNAs. Their biological significance and biogenesis pathway is currently the subject of intensive study and several research groups are exploring and other potential functions. piRNABank is a highly user-friendly resource which stores empirically known sequences and other related information on piRNAs reported in human, mouse, rat and *Drosophila*. The database supports organism and chromosome wise comprehensive search features including accession numbers, localization on chromosomes, gene name or symbol, sequence homology based search, clusters and corresponding genes and repeat elements. It also displays each piRNA or piRNA cluster on a graphical genome-wide map along with all associated elements like retrotransposons, genes, CDS, exons, intron, ORF, GC region etc. Since miRNAs have been hypothesized to perform a combinatorial regulation with piRNAs, the tool also facilitates tracking of miRNAs with aforesaid features. Currently, piRNABank is being updated for piRNA sequence and cluster data from platypus and zebra fish. Newer tools are being integrated to extract the sequence and structural motifs in user desired set of piRNA data. piRNABank with these features will surely be very useful in deciphering the piRNA associated disease biology (recently it has also been associated with autism). Availability: http://pirnabank.ibab.ac.in/.

515: Human encoded microRNA interferes with HIV-1 replication and nef expression

[1]**Jasmine K. Ahluwalia**, [2]Sohrab Z. Khan, [1]Kartik Soni, [1]Ankit Gupta, [1]Manoj Hariharan, [1]Vinod Scaria, [1]Ashok Patowary, [1]Mukesh Lalwani, [1]Sridhar Sivasubbu, [1]Beena Pillai, [2]Debashis Mitra, [1]Samir K. Brahmachari

[1]Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi, India, [2]National Centre for Cell Science, Ganeshkhind, Pune, India

MicroRNA (miRNA) are endogenous, small non-coding RNA that regulate gene expression at the post-transcriptional level. miRNA may bind to 3′UTR of protein coding genes with partial complementarity and repress their translation whereas complete complementarity may lead to transcript degradation. miRNA are involved in various processes like, cell proliferation, differentiation, apoptosis, cancer, development etc. Human immunodeficiency virus type 1 (HIV-1) is a single-stranded RNA virus of the lentivirus family of retroviruses. It primarily invades CD4+ T-cells of the host immune system leading to one of the most adverse clinical conditions, acquired immune deficiency syndrome (AIDS). Besides the number of genes imperative for virus infection and/or replication, encoded by the HIV genome, the virus also depends on host cell machinery for the synthesis of viral RNA and proteins. Using in silico analysis, we have reported five cellular microRNAs with potential binding sites in four HIV-1 encoded genes (Biochem. Biophys. Res. Commun., 2005). Amongst them, hsa-mir-29a and hsa-mir-29b that target HIV-1 nef gene, showed maximum microRNA target prediction scores. Nef is well known as a key player in HIV pathogenesis and importantly, long-term non-progressors were found to be infected with virus carrying deletions in the nef gene. Thus, it was interesting to experimentally investigate the regulation of HIV-1 nef gene by these cellular microRNAs. First, we found that hsa-mir-29a and hsa-mir-29b were abundantly expressed in peripheral blood mononuclear cells (PBMCs) and jurkat cells. Next, using various molecular approaches we have verified that hsa-mir-29a can regulate HIV-1 nef expression as well as replication. Our results underscore the importance of cellular miRNA levels in HIV-1 infection.

516: A novel non-coding RNA in developing gonads of *Crocodylus palustris*: a candidate gene having a role in sex determination

Amit Anand, Minarbha Patel, Albert Lalremruata, Lalji Singh, R. K. Aggarwal

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad, 500007, India

Non-coding RNA have a Lion's share of transcriptome. The ncRNA have critical role in regulation of gene expression and other cellular processes, which is well established for the small non-coding RNA. In addition another class of non-coding RNA, mRNA like non-coding RNA also appear to play important regulatory roles, for example in heat shock response, regulation of Hox genes, X chromosome inactivation etc. We report here one such novel mRNA like ncRNA having a probable role in temperature dependent sex determination (TSD) in *Crocodylus palustris*. In TSD, the incubation temperature determines the sex of the developing embryo during mid-late stages of embryonic development, which is called temperature sensitive period (TSP). We identified an ncRNA like transcript in a global screen for differentially expressed genes responsive to temperature in the bipotential gonads during early TSP. The mRNA like nc-RNA is 3kb long and holds no significant similarity to any database sequence except 160 bp region to a solute carrier gene (*SLC35F5*). This nc-RNA was ruled out to be of intronic origin from *SC35F5* coding locus by genome walking. Most interestingly, the detailed expression analysis of the ncRNA showed that: (1) both strands are transcribed; (2) one of the strands is expressed at 35–40-fold than the other one only at male promoting temperature (MPT), and (3) the ncRNA at MPT specifically localizes in the developing testicular cords (TCs) and not even in the interstitium. Thus, the study provides the first evidence of an mRNA like ncRNA having a putative role in male-sex determination in TSD species.

517: Integrated analysis of gene expression by correlating the transcriptome and miRNome

Rhishikesh Bargaje, Mahantappa Halimani, Kartik Soni, Beena Pillai

Institute of Genomics and Integrative Biology, Mall Road, Delhi, India

Information flow in biological systems, involves the expression of genes through the sequential steps of transcription and translation in a regulated manner. Although the regulatory role has been widely attributed to proteins, changes in the transcriptome are also used to assess the effect of mutations, chemical and physical agents. Recently, hundreds of 21–23 nt small, naturally occurring regulatory RNA molecules have been discovered in almost all eukaryotes. These microRNA (miRNA) molecules can pair with untranslated regions of messenger RNA molecules with imperfect complementarity and negatively regulate gene expression. The transcriptome, consisting of thousands of relatively long, functional transcripts, a majority of which code for proteins, is therefore a transient output of gene expression. This output is further regulated by the miRNome, necessitating an integrated view of the transcriptome and miRNnome to understand gene expression. The technological improvements in high-throughput data collection methods in recent years have accelerated collection of transcriptome and miRNome data using microarrays. However, integrated analysis of gene-expression by overlapping the regulatory information contained in the miRNome, over the transcriptome remains a challenge. We have used publicly available data and experimental data generated by our group to analyze expression data in an integrated manner. Firstly, we used publicly available microRNA data from microarray and bead based platforms to develop methods of cross-platform, inter-laboratory comparison. We also developed a freely available database of the meta-analyzed expression profiles. We are currently developing methods of correlating mRNA and microRNA expression, in cell. Secondly, our group generated couple datasets to use as test sets for integration of transcriptome, miRnome and proteome. The first dataset pertains to changes in neuronal cells associated with polyglutamine expansion in TATA binding protein, a general transcription factor. We collected data on the genome-wide expression profile of mRNAs and miRNAs in neuronal cells expressing normal and expanded polyglutamine variants of TBP. New data on the transcriptome, miRnome and proteome following transient expression of microRNAs, will be presented.

518: Incomplete penetrance and variable expressivity: the microRNA connection

[1]**Vani Brahmachari**, [1]Jasmine Ahluwalia, [2]Manoj Hariharan, [2]Rhishikesh Bargaje, [2]Beena Pillai

[1]ACBR, North Campus, Delhi University, India, [2]Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi, India

Incomplete penetrance and variable expressivity (IP-VE) are non-Mendelian phenomena resulting in the lack of correlation between

genotype and phenotype, leading to variations in gene expression in addition to that brought about by mutations. Notwithstanding the differences in mechanisms, variations in penetrance and expressivity of mutations can be viewed as a manifestation of differential expression of homologous alleles. In this framework and in the light of the recent advances in understanding the role of micro RNA in fine-tuning gene expression, we propose that miRNA-mediated regulation brings about non-equivalence in expression of homologous alleles at the translational level. To corroborate our hypothesis we carried out in silico analysis for the presence of miRNA binding sites in 3′UTR using a consensus of three independent prediction tools and find that a IP-VE and modifiers genes are significantly targeted by miRNA in comparison to housekeeping genes (P value less than 0.05). Though, co-expression of target gene with the miRNA is a debatable criterion for filtering functionally relevant interactions, we consider that for the miRNA to function as a post-transcriptional regulator, co-occurrence with target mRNA is required. This is based on the fact that the description of IP-VE implies modulation rather than complete silencing of bi-allelic expression. Consequently, subtle changes in the levels of the co-occurring target gene-miRNA pair, under different cellular conditions in different individuals may lead to variable expressivity. Our results of analysis of co-expression of target gene-miRNA pair, and their co-occurrence in different human tissues including the affected tissue, along with the inverse relationship of protein expression between target mRNA and miRNA levels and the occurrence of SNP in the miRNA target site implies an important role for miRNA in the epigenetic regulation of penetrance and expressivity of human genetic disorders.

519: Involvement of repeats in non-random genomic divergence in 5′ upstream regions of human and chimpanzee genes of different biological processes

[1]**Amit Chaurasia**, [2]Ravi Shankar, [1]Mitali Mukerji

[1]Functional Genomics Unit, IGIB, Mall Road, Delhi, India,
[2]Department of Bioinformatics and Structural Biology, IIAR, Gandhinagar, Gujarat, India

Non-coding DNA which comprises a major fraction of the human genome is increasingly being implicated in diverse functions. In an earlier exploratory study we demonstrated that the genes involved in neurodevelopment, housekeeping and neurophysiology processes show significant differences in divergence between human and chimpanzee. Whereas neurophysiology shows high divergence in the 5′ upstream regions of the genes, those of neurodevelopment show significant divergence in the non-coding sequences of downstream genic regions. Housekeeping genes demonstrate high divergence both in the downstream as well as upstream regions that is completely diminished if the repetitive sequences are removed. We also demonstrated that repetitive sequences in the 5′ upstream regions provide one third of novel transcription factor binding sites which might contribute to the functional divergence in human. In this study we have extended this analysis to the entire genome to probe for enrichment of specific processes or functions which have diverged maximally between the two species. We used Macaque as an out-group in this study to infer human lineage-specific events. From the entire human genome set we retrieved sequences for 11,012 annotated genes (with SwissProt ids) which share considerable amount of homology in the 5′ upstream regions of chimpanzee and human. We also mapped the repeat coordinates (simple repeats, SINEs and LINEs) in these sequences using RepeatMasker. The sequences were aligned and Juke's Cantor (JC) distance was used to estimate divergence between the species, with and without repeat-masking. The upper and lower 10% diverged sets (Set-1- with repeats and Set-2- after masking repeats) were identified from both these groups and were analyzed for GO (biological process and cellular component) enrichments after applying Bonferroni correction. In Set-1 the low diverged set had significant enrichment in genes involved in regulation of transcription and nucleobase, nucleotide and nucleic acid metabolic processes and highly diverged in gene involved in multicellular organismal processes related to response to stimulus, neurological process, sensory perception, inflammatory process, cellular and lipid metabolic processes etc. On an average, nearly 20% of the divergence is due to accumulation of repeats in the highly diverged set. Contribution of the repeats to the transcriptional repertoire is being studied.

520: Determination of relative telomere length in normal, obese and hypertensive individuals: a quantitative real time PCR approach

Birajalaxmi Das, Divyalakshmi Saini, M. Seshadri

Radiation Biology and Health Sciences Division, Bio-medical Group, Bhabha Atomic Research Center, Trombay, Mumbai 400085, India

Telomeres are specialized nucleoprotein complexes that serve as protective caps of human chromosomes. It consists of tandem repeats of DNA sequences (TTAGGG)n at the end of the chromosome and plays an important role in maintaining chromosome stability. They are also implicated in regulating the replication and senescence of cells. The gradual loss of telomere repeats in cells has been linked to aging and tumor development. Telomere length is considered as a biomarker for aging, stress and survival. Experimental methods have been developed to determine telomere length including southern blot, fluorescence in situ hybridization (FISH) and Flow-FISH analysis. Recently developed quantitative real time PCR method provides an easy and faster method to measure the relative telomere length, where the telomere length is obtained by calculating the ratio of a PCR product from the same sample using specific primers for telomeres and single copy gene (Cawthon 2002). In the present study, we have made an attempt to measure the relative telomere length of random normal, obese and hypertensive individuals using quantitative real time PCR approach. Genomic DNA was isolated from human peripheral blood lymphocytes of 105 normal, 27 obese and 27 hypertensive using a non-enzymatic salt precipitation method (Lahiri and Nurnburger 1991). PCR amplification was achieved in LC480 (Roche Diagnostics) using SYBR Green based assay. Both telomere specific (T) and single copy gene primers (S) were used for calculating the relative ratio of telomeric and single gene. All the samples were analyzed in triplicate. For single copy gene, a 36B4 gene specific primer was used and was validated using beta-globin gene. The relative ratio 36B4 and beta-globin genes for all experimental DNAs in comparison with the reference DNA was approximately 1.0. The ratio of telomeric and single gene obtained is proportional to the telomere length. The mean ratio of T and S was 1.22, 1.21 and 1.11 for normal, obese and hypertensive individuals, respectively. There is a statistically significant difference ($P < 0.001$) in the mean ratio of T and S of hypertensive individuals as compared to control and obese individuals. The relative ratio of telomere and single gene is also correlated with age and sex. There is no significant difference between the male and female individuals. A marginal decrease was observed in the mean ratios in the elder age group. However, more samples are required to correlate the mean telomere length with age.

521: Signatures of recent positive selection in human miRNA genes based on a genome-wide analysis

Navonil De Sarkar, Nidhan Kumar Biswas, Sayantan Datta

Human Genetics Unit, Indian Statistical Institute, 203, B.T. Road, Kolkata 700108, West Bengal, India

Introduction: MicroRNAs (miRNAs) regulate gene function post-transcriptionally and their role in modulating disease outcome is becoming increasingly clear. Though only the mature miRNA acts as an effector molecule, pri-miRNA (pre-miRNA and its $\sim$100 bp flanks) is important for miRNA biogenesis. Nucleotide (nt) changes in miRNA genes are known to affect its function. Thus, such nt changes may be impacted upon by natural selection. Under recent positive selection (RPS), alleles are expected to be found in polymorphic frequencies in multiple populations and the homozygosity around the site with a favored allele should extend over a large genomic distance.
Aim: We have sought to test whether recent positive selection has structured the nature and extent of variability within and around miRNA genes.
Method: SNPs genotyped in the HapMap Phase II in all miRNA genes ($n = 678$) identified in the human genome (http://www.microrna.sanger.ac.uk) and their 100 bp flanks (comprising the pri-miRNA region) were included if these were polymorphic in at least two of the three HAPMAP populations. Standardized integrated haplotype score (iHS) was calculated for each included SNP using Haplotter (http://www.hg-wen.uchicago.edu/selection/haplotter.htm) in 200 kb region around core SNP. We tested whether the proportion of SNPs with high iHS in miRNA genes (based on recommended thresholds) was significantly greater than that of the genome wide proportion.
Results: Our inclusion criteria for SNPs yielded 89 SNPs in 83 pri-miRNA genes. Ten of these SNPs showed significant sign of RPS ($|iHS| \geq 2$) in any one population. These SNPs fall within the following miRNA genes: *MIRN148A*, *MIRN197*, *MIRN211*, *MIRN219-1*, *MIRN423*, *MIRN608*, *MIRN611*, *MIRN628*, *MIRN640*. The proportion of SNPs with high iHS was significantly greater in the miRNA genes compared to the genome wide proportion ($P < 0.005$). Among the 10 SNPs, 8 were present in the 100 bp flanks (comprising the pri-miRNA region) and one each in the pre-miRNA and mature miRNA regions. Seven SNPs showed signs of RPS in African(YRI) population and 5 of these had $iHS > 2$ indicating that positive selection has favored the ancestral allele. The remaining three SNPs showed RPS in the Asian (CHB + JPT) population with the indication of positive selection favoring the ancestral allele.
Conclusions: Compared to the entire human genome, a disproportionately higher impact of recent positive selection was observed in miRNA genes. Ancestral alleles have usually been favored.

522: MicroRNAs and non-micro-short RNAs (nmsRNAs) in cell quiescence and cancer

Anindya Dutta

Department of Biochemistry, University of Virginia, 1300 Jefferson Park Ave, Box 800733, Charlottesville, VA 22908, USA

We are using three approaches to discover microRNAs that regulate cell proliferation. In the first approach, we focused on miRNAs induced during differentiation of C2C12 myoblasts into myotubes in vitro. One of the microRNAs induced during differentiation, miR-206, suppresses cell proliferation. By surveying mRNAs that are downregulated by miR-206 for genes predicted to be targets of the microRNA, we discovered the mRNA of DNA polymerase alpha catalytic subunit (POLA1) is a direct target that is destabilized upon miR-206 transfection. We now report that additional microRNAs induced during muscle differentiation target other cell-cycle regulators. After noting that targeted mRNAs are often destabilized by microRNAs, in the second approach, we tested whether global downregulation of microRNAs by knockdown of Dicer or Drosha can help identify microRNA-repressed mRNAs. The principle was proved by the discovery that HMGA2 oncogene is repressed by the growth-suppressive let-7 microRNA. Indeed, chromosomal translocations in lipomas and leiomyomas de-repress HMGA2 by deleting the 3′UTR that is normally repressed by let-7. Conversely, downregulation of let-7 has been noted in lung cancers and large leiomyomas and is correlated with over-expression of the HMGA2 oncogene. In the third approach we have cloned short RNAs from androgen-dependent prostate cancer cells grown in the presence or absence of androgens and subjected them to ultra-high-throughput sequencing. The frequency of a number of clones present in the libraries is changed by androgens and we have identified several microRNAs that change upon androgen depletion. Over 30–40% of the short RNAs cloned, however, do not correspond to known microRNAs and are produced by cleavage of known mRNAs and noncoding RNAs. The diversity and abundance of the non-micro-short RNAs (nmsRNAs) contrast with how little is known of their function and suggest that much remains to be done before we understand the biological functions of many short RNAs present in the cell.

523: Mir-29 is up-regulated during skeletal muscle differentiation

Fude Fang, Liuluan Zhu, Aibin He, Jin Zuo, Yongsheng Chang

Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, 5# Dong Dan San Tiao, China

MicroRNAs(miRNAs) are endogenous small noncoding RNAs which suppress gene expression by binding to the complementary sequence in the 3′-UTR of target mRNAs. MicroRNAs regulate their targets by direct cleavage of the mRNA or by inhibition of translation, according to the degree of complementarity with their targets. Lots of microRNAs have been identified to play critical roles in many aspects, such as cell proliferation, apoptosis, development, differentiation and metabolism, the functions of most remain unknown. Here we showed that the miR-29 family was up-regulated during skeletal muscle differentiation. In the process of L6 differentiation from myoblast to myotube, all of the miR-29 family was up-regulated, expression assayed by northern blotting analysis. There are three paralogs (29a, 29b and 29c) of miR-29 family with the same seed sequence in two separate clusters in rat. To further study, we predicted a potential binding site of MyoD upstream of miR-29c. MyoD is a key prerequisite to myogenesis by opening and closing some genes' expression. We proposed that miR-29 was transcriptionally up-regulated by MyoD, therefore suppressed some targets during L6 differentiation from myoblast to myotube.

524: Involvement of a nuclear restricted, Drosha processed non-coding RNA in gene regulation

Gayatri Ganesan, M. R. S. Rao

Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India

Non-coding RNAs constitute a huge repertoire of gene regulatory molecules. Our previous, fine resolution characterization of a mouse

meiotic recombination hotspot from chromosome 8 resulted in identification of 2.4 kb unspliced and polyadenylated non coding mrhl RNA(1). The gene is expressed in multiple tissues and is also present in rat but absent in humans. The mrhl RNA gets processed to a small 80 nt RNA species and is mediated by Drosha complex. We also observe that the 80 nt Drosha product could be processed further to a 22 nt small RNA by Dicer in an in vitro reaction. However, this 22 nt product was not detected in vivo. The 80 nucleotide as well as the 2.4 kb full length RNA are nuclear localized showing distinct punctate nuclear signal(2). Silencing of mrhl RNA showed a significant difference in the level of expression of some of the genes associated with signaling pathways. Our results point that the non coding mrhl RNA could be a member of nuclear regulatory RNA.

1. Nishant KT, Ravishankar H, Rao MRS (2004) Characterization of a mouse recombination hot spot locus encoding a novel non protein coding RNA. Mol Cell Biol 24:5620–5634
2. Gayatri G, Rao MRS (2008) A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. RNA. doi:10.1261/rna838308

525: The human satellite III non-coding RNA is involved in general stress response and modulates the cellular toxicity of misfolded proteins associated with neurodegenerative disorders

S. Ganesh, S. Sengupta

Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur, India

Nuclear stress granules (NSG) are sub-nuclear compartments formed in the human cell under the condition of heat shock. Under thermal stress, the heat shock transcription factor-1 activates the transcription of satellite III repeats (S3R) into stable RNAs that eventually form the NSG. S3R transcripts associate with several splicing factor and other RNA processing proteins, and thereby possibly protect them during stress conditions. In an effort to understand the cellular functions of this non-coding RNA, we have blocked or over-expressed S3R transcripts under various stress conditions in a cellular model and evaluated the effect on cell survival and functions. Our studies show that, besides the heat shock, the S3R transcription is induced by proteasomal blockade or by an increase in the cellular levels of misfolded proteins. While S3R knockdown leads to apoptosis in heat-shocked cells, its overexpression does not appear to have any adverse effect on the cell survival. Curiously, S3R overexpression elevates the toxicity caused by huntingtin protein with expanded polyglutamine repeats. S3R show intriguing similarity with hsr omega transcripts of *Drosophila*. Both are non-coding transcripts with repeat sequences, and are strongly induced under the condition of heat shock, suggesting that they are functional homologues. Consistent with this notion, overexpression of hsr omega transcripts rescued the loss of S3R under heat shock in human cells, and increased the toxicity of mutant huntingtin protein when expressed together. Intriguingly, hsr omega is known to increase the cytotoxicity of expanded huntingtin protein in fly models. Our findings, besides strengthening the notion that SatIII and hsr omega are functional equivalents, provide novel functions for noncoding RNAs in stress response and in disorders associated with toxic gain-of-function proteins. We will discuss the cellular pathways that are regulated by S3R and the factors that regulate the expression of S3R.

526: Genome wide analysis and discovery of MicroRNA: associated regulatory motifs in two mammalian genomes

Sherin C. George, Ramkumar Hariharan, M. Radhakrishna Pillai

Translational Cancer Research Laboratory, Rajiv Gandhi Centre for Biotechnology, Thycaud PO, Poojappura, Thiruvananthapuram, India

Previous experiments have shown that human and mouse embryonic stem cells express a specific set of microRNAs that are believed to be essential for the maintenance of the stemness phenotype. Although the core promoter elements (proximal) in the upstream regions of human MicroRNAs have been identified and reported, we wished to investigate whether there are specific cis-regulatory elements upstream of co-regulated microRNAs that might function in a cell type specific (here, cell type = embryonic stem cell) manner. Since expression data for microRNAs in the context of 'stemness regulators' were available in the case of human and mouse, we decided to use a hybrid approach to discover motifs. An integrated approach, making use of both phylogenetic footprinting (across human and mouse) as well as co-regulation data (from human and mouse ES cells) was used to discover new motifs. The discovered motifs comprised of previously reported ones and new ones. Comparisons with known TS binding motifs in the TRANSFAC database revealed that three of the discovered motifs were known transcription factor sites. While this analysis can be extended to analyse more genomes, in general, we present a powerful new approach, combining phylogenetic footprinting and co-regulation analysis to discover regulatory motifs.

527: MicroRNAs in cervical carcinogenesis

[1]**Reshmi Girijadevi**, [2]Vinod Chandra, [1]M. Radhakrishna Pillai

[1]Translational Cancer Research Laboratory, Rajiv Gandhi Centre for Biotechnology, Thycaud PO, Poojappura, Thiruvananthapuram, India, [2]Department of Computer Applications, College of Engineering, Thiruvananthapuram, India

Intensive research efforts aimed at unveiling the molecular mechanisms of cancer, has demonstrated existence of at least a subset of tumors characterized by over-dependence on certain key oncogenic pathways. The components of such pathways therefore constitute attractive molecular targets for therapeutic intervention. MicroRNAs (miRNAs) are non-coding small RNAs that play important roles in a variety of biological pathways including cellular proliferation and apoptosis. Recent studies have linked the expression of selected miRNAs to carcinogenesis and viral pathogenesis. Bioinformatics approaches have proved to be very useful toward this goal by guiding experimental investigations. To investigate the possible involvement of miRs in human cancers on a genome-wide basis, we have mapped genomic instability regions in cervical cancer cells. Various chromosomal and segmental loss or amplification regions have been detected in primary cervical tumors and cervical cancer cell lines. We hypothesized that microRNAs on common genomic instability region in cervical cancer cells may contribute to cervical tumorigenesis. First, we localized more than ten common fragile chromosomal regions with the help of bioinformatics resources and used novel algorithms to detect potential miRNAs from same sites, as well as in minimal regions of loss of heterozygosity. In-silico studies are also being carried out on our candidate microRNAs to check the kinetics and target predictions. These data provide a catalog of miR genes that may have roles in cancer and argue that the full complement of miRs in a genome may be extensively involved in cancers. Functional assays will be performed for these candidates to better understand their potential roles in cervical tumorigenesis pathways.

528: More to intronic miRNAs: 'Double Locks' and feedback loops

Manoj Hariharan, Vinod Scaria, Samir Brahmachari

Institute of Genomics and Integrative Biology, Mall Road, Near Jubilee Hall, Delhi, India

MicroRNA (miRNA) are small noncoding RNAs that have the potential to interact with specific sites in the 3′ untranslated regions (UTRs) of mature transcripts and generally reduce their translational efficiency. Almost half of the currently known miRNAs are encoded in the introns of protein coding genes. Other than the understanding of their location, no functional significance of these miRNAs is studied yet. We hypothesize that the location has a precise functional advantage of being able to coexpress the miRNA with the source gene. This would enable the inhibitory RNA to interact with functionally antagonistic genes thereby creating a synchronized downregulation of proteins while maintaining the transcript levels of the targets. Such processes are most relevant in processes like cell cycle where activity of a certain set of genes is possible only while certain proteins are not actively functional. A whole genome analysis for target sites using all known conserved human intronic miRNAs (CIMs) to all 3′UTRs was performed using a message passing interface (MPI) suite of parallel programs. Four ubiquitously expressed source genes functional in cell cycle, viz. DLEU2, SMC4, NFYC and MCM7, encode seven CIMs which target 1,397 genes. We sought evidence of post transcriptional gene silencing (PTGS) by comparing the transcript and protein levels obtained from UniGene and Human Protein Atlas which provides data for 45 tissues. Out of the 78 genes which had both EST and protein values known, 13 show maximum PTGS. We focus on the target genes CD82, FN1 and MCAM since they are targeted by two CIMs hsa-miR-106b and hsa-miR-93, encoded in intron of MCM7. All the three genes are tumor suppressors which have the tendency to hamper cell cycle progression. Hence, while MCM7 (a key gene involved in cell cycle progression) is active, the miRNAs which can reduce CD82, FN1 and MCAM genes are also expressed so that they do not interfere with the normal cell cycle. We also notice that the transcriptional activators of the target genes are also targeted by the same miRNAs. This indicates a 'double-lock' where inhibition operates at the transcriptional and post-transcriptional levels of the target genes. Another observation is that the transcriptional activator of MCM7 is regulated by its own intronic miRNAs, which is a feedback regulation. The current study opens up a new area of understanding of regulatory networks mediated by intronic miRNA. The details of the workflow and the key findings would be discussed.

529: Application of artificial microRNA in antiviral therapeutics

Vaibhav Jadhav, Vinod Scaria, Souvik Maiti, Beena Pillai

Institute of Genomics and Integrative Biology, Near Jubilee Hall, Mall Road, Delhi 110 007, India

MicroRNAs are recently discovered regulatory molecules that generally mediate repression of gene regulation at post-transcriptional stages. MicroRNAs play a critical role in a wide variety of biological processes including host-virus interaction. We have earlier shown that human microRNAs can target critical HIV genes including nef, thought to be a critical determinant of long-term progression of disease. Our studies and other reports of anti-viral miRNAs suggest an attractive approach for designing therapeutics based on the principles of microRNA binding and action. Artificial microRNA, designed against critical viral genes are expected to down regulate specific viral target transcripts, individually and in combination with each other. We have developed an algorithm to design artificial microRNAs using structural and functional principles of natural microRNA, against ultra-conserved sequence elements in gag, env and pol transcripts of Human Immunodeficiency Virus 1. Targets were sorted out using a consensus of five different microRNA target prediction tools. Further, targets were also evaluated for accessibility and thermodynamic feasibility based on their RNA secondary structure. We developed a cell culture reporter assay for proof of concept experiments to validate our design. Our data demonstrate that expression of virus-specific artificial miRNAs is an effective and predictable approach to engineer viral gene expression. To the best of our knowledge, this is the first report of computational design and validation of artificial microRNAs as a therapeutic strategy against human viruses. This algorithm can be used for designing artificial microRNAs against viral transcripts, as well as against specific cellular transcripts. All the three artificial miRNAs we designed using the algorithm showed more than 70% reduction in target reporter expression. The high efficacy of all the designed miRNAs suggests that the algorithm could be useful in designing artificial miRNAs as parts of synthetic modules in gene regulation. This strategy finds applications in design of synthetic regulatory RNA molecules for fundamental as well as in applied research.

530: Novel non-coding RNA from human Yq12 heterochromatic block provides a testis-specific 5′UTR to CDC2L2 mRNA by trans-splicing

Rachel A. Jesudasan, S. Vallinayagam, Zeenath Jehan, Shrish Tiwari, Suman Pradhan, Lalji Singh, Hemakumar M. Reddy

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad, A.P., India

Y chromosome, enriched in repetitive DNA, has been the most intractable to genetic and molecular analyses. Recently, a comprehensive reference sequence of the human male-specific Y (MSY) was built up. Gene prediction programs and RT-PCR experiments mapped 156 transcription units to the human Y chromosome, all of which localized to MSY euchromatin. No evidence for transcription was found in the major distal, heterochromatic block (Yq12) of MSY. Besides reports of localization of various satellite sequences like male-specific DYZ1, DYZ2 and DYZ18, there was practically no insight into the region. With increasing evidence on transcription from various repeats, the 40 Mb heterochromatic block, enigmatically poised at the rear end of the long arm of human Y, was examined for gene expression. Here, we present evidence for the first time, for developmental stage-specific and testis-specific transcription from MSY distal heterochromatic block. We have isolated two cDNA sequences, which localize to Yq12 in multiple copies, show testis-specific expression and lack active X-homologues. Thus, by definition and by empirical evidence, two partial cDNAs conforming to the class 2 genes of Y span this region of MSY. One of the cDNAs identified a 67 nt region with 100% homology within the 5′UTR of CDC2L2 gene, a ubiquitously expressed cell division control kinase. Experimental evidence showed that one of these Yq12 noncoding RNAs (ncRNAs) trans-spliced with CDC2L2 mRNA from chromosome 1p36.3 locus to generate a testis-specific chimeric _sv13 isoform. This 67-nt 5_UTR provided by the Yq12 transcript contained within it a Y box protein-binding CCAAT motif, indicating translational regulation of the _sv13 isoform in testis. In vitro transfection experiments using the Luciferase assay in HeLa cells showed upregulation of translation on incorporation of the UTR in front of the reporter gene. The study presents for the first time definitive evidence for testis-specific and development stage-specific transcripts from the MSY (male-specific Y) distal heterochromatic block, classically considered junk. This is also the first report of trans-splicing between a Y chromosomal non-coding RNA and an autosomal transcript. This is therefore, the first report of Y-autosomal interaction as well.

531: Mechanistic insights into Kcnq1ot1 non-coding antisense RNA-mediated transcriptional silencing of the chromosomal domains

Chandrasekhar Kanduri, Radha Raman Pandey, Faizaan Mohammad, Tanmoy Mondal

Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Dag Hammarskjölds Väg 20, 75185, Sweden

Transcriptome studies over the past few years have revealed that the genomes of all studied eukaryotes are almost entirely transcribed and that a significant part of the mammalian genome codes for non-translated RNA (ncRNA) while only 1.2% of the mammalian genome codes for protein-coding RNA. It is increasingly apparent that many of these ncRNAs possess regulatory functions. NcRNAs can be classified into two groups; small ncRNAs (21–31 nucleotide long, and include siRNAs, miRNAs and PiRNAs) and long ncRNAs (range in size from 100 bp to several hundred kb). Interestingly, both small and long ncRNAs seem to regulate diverse functions through regulating chromatin architecture and epigenetic memory, thus underlining the functional role of ncRNA in the maintenance of chromosome structure and function. Of particular interest is that gene clusters, with parent-of-origin-specific expression patterns, often contain one or more ncRNAs as their partners. Some of these ncRNAs are very large, ranging in size from 50 kb to several hundred kb, and transcribed in antisense direction from an intronic portion of protein coding gene. The expression of long antisense ncRNAs has been shown to be correlated with the repression in cis of protein coding genes, spread over several hundred kb on either side of the antisense transcription unit, indicating a link between ncRNA expression and silencing of neighboring protein-coding genes. Over the last few years my laboratory has been working to understand the functional role of long antisense ncRNAs in transcriptional silencing using Kcnq1ot1 as an example. The Kcnq1ot1 promoter maps to a 3.6 kb imprinting control region (Kcnq1 ICR) in intron 10 of the Kcnq1 gene and is encoded only on the paternal chromosome, as the Kcnq1ot1 promoter is methylated on the maternal chromosome. It is transcribed in antisense orientation to the Kcnq1 gene. The production of Kcnq1ot1 RNA on the paternal chromosome has been linked to the silencing of nearly 10–12 genes spread over a one mega-base Kcnq1 domain. Here we show that Kcnq1ot1 is 91 kb long and contains a 0.9 kb silencing domain at its 5′ end, which mediates bidirectional silencing through positioning the flanking sequences in the perinucleolar compartment, during mid S phase. More importantly, it carries out lineage-specific silencing through regulating the chromatin by exploiting lineage-specific silencing mechanisms. Our data, therefore, provide mechanistic insights into epigenetic regulation by long antisense ncRNAs.

532: Endogenous XIST complementary small RNA and RNAi machinery in human X dosage compensation

[1]**Satya Keerthi Kota**, [2]Lalji Singh, [1]Utpal Bhadra

[1]Functional Genomics and Gene Silencing Group, Centre for Cellular and Molecular Biology, Hyderabad 500007, India, [2]Centre for Cellular and Molecular Biology, Hyderabad 500007, India

Equalization of X-linked gene expression between two sexes despite their differences in copy number is mediated by transcriptional silencing of one X chromosome in female cells. Small regulatory RNA from the non-coding repetitive sequences plays crucial roles in transcriptional silencing and heterochromatin formation. To investigate the involvement of RNAi and small non-coding RNA in the maintenance of human X inactivation, we have performed strand-specific northern blots using antisense/sense probes derived from XIST exon-1 with RNA extracted from different human male and female cell lineages. Small non-coding RNA species ranging from ~18 to 25 nt in length with predominant class of 21–24 nt RNA were observed. These RNA referred as sXiRNA (small XIST RNAs) were found in RNA from the different human cells carrying at least one active X chromosome. These small regulatory RNAs are independent to RNA polymerase II transcription and Xi localized XIST or TSIX transcripts. Further, preferential localization of sXiRNA in the nuclear fraction indicates their functional role in the nuclear transcriptional control rather than post transcriptional gene regulation. These small RNA are complementary to Repeat C region but not to other regions of XIST transcript. Formation of sXi RNA requires DICER1, an RNase III class enzyme, central mediator of RNAi pathway in human cells. DICER1 depletion in somatic cells leads to significant reduction in sXiRNA levels revealing the double stranded RNA (dsRNA) presence within the XIST locus of somatic cells. DICER1 depletion also leads to global and X chromosome specific increase in histone modifications indicative of transcriptionally active chromatin and increase in transcription of selected X linked genes from distinct X chromosomal loci. These effects were partially rescued by exogenous addition of sXiRNA indicating specific role of RNAi pathway and sXiRNA in the process of the Xi maintenance. Therefore, these repeat derived sXiRNA of XIST locus might be crucial players in mammalian RNAi mediated Xi transcriptional silencing complex by altering histone code and organization of silent X chromatin.

533: The predicted microRNA regulation and target identification on candidate genes involved in congestive heart failure

Anita Krishnan, Sri Lakshmi Dingari, Manika Pal Bhadra

Indian Institute of Chemical Technology, Department of Chemical Biology, Hyderabad 500007, India

Congestive heart failure is characterized by the left ventricular remodeling and the reactivation of the fetal gene program in which both environment and genes effectively contributes. However, the evaluation mechanism of the congestive heart failure is very poorly understood. microRNA has recently been discovered as a most important regulatory molecule consisting of endogenous tiny RNA with only 22–25 bp in length. It regulates gene expression either by complementary hybridization to messenger RNA. With the consequence of mRNA degradation or translational inhibition of mature transcripts. Recent evidence suggests that microRNA also may participate in control of cardiac hypertrophy of humans. The accumulation of microRNA in the heart impairs pool of proliferating ventricular myocytes. On the contrary, mir-133 increases the skeletal myoblast proliferation by inhibiting the serum response factor. In the present study, we describe a computational approach in identifying complete set of cardiac microRNA that plays the crucial role in key regulation of congestive heart failure. The presence of microRNA binding sites on the coding regions and 3′UTR of five selected genes was predicted. Using different algorithm with rigid statistical parameter and higher confidence values we identified 46 microRNA targets. Among them mir-612 and mir-493, mir-432 is most prevalent related to heart failure because they identified multiple target sites in the selected genes. Precisely, mir-939 identified multiple target sites in TNFalpha genes. Further validation of miRNA network and regulation will substantially improve our understanding of cardiovascular biology.

534: GATA repeat associated regions of human Y chromosome mark chromatin domain boundaries and facilitate long-range interactions

Ram Kumar, Lalji Singh, Rakesh Mishra

Center for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India

Repetitive sequences are the major component of the genomes of higher eukaryotes. Although simple sequence repeats account for ~3% of the human genome, their functional significance remains unclear. Abundance and distribution of these repeats across the genome, however, shows tremendous repeat specific variation. This non random pattern is an indication of their possible functional relevance. Earlier studies showed an unusual enrichment of longer GATA repeats in the genomic context that could coincide with boundaries of chromatin domains. These GATA repeats on the human Y chromosome are associated with the genes that are expressed during early development, further indicating that GATA region may be marking loci for a coordinated regulation of gene expression. Here, we show that such GATA repeat elements have chromatin domain boundary activity in *Drosophila melanogaster* as well as in human cells. This boundary function is positively influenced by the presence of additional transgenic copies indicating trans interaction among such elements. We also show by FISH that GATA repeat elements cluster in *Drosophila* indicating the long-range interaction mediated by these elements. Our studies demonstrate that GATA repeat elements have a role in packaging of genome and regulatory mechanisms involving large regions of chromosomes. These observations suggest that repetitive elements in higher eukaryotes may function as part of genomic 'packaging code' with unique regulatory properties of regulating many loci and provide an explanation why higher eukaryotes have accumulated greater proportions of repetitive elements.

535: Identification of long non-coding RNAs regulated by retinoic acid during neuronal differentiation of the human neuroblastoma cells

[1]**I. V. Kurochkin**, [1]Y. Nishida, [1]R. Ozawa, [2]N. Adati, [3]K. Takamatsu, [1]T. Takeda

[1]MetaSystems Research Team, RIKEN Advanced Science Institute (ASI), Yokohama, 230-0045, Japan, [2]Cellular System Modeling Team, RIKEN Advanced Science Institute (ASI), Yokohama, 230-0045, Japan, [3]Kobe Tokiwa University, Kobe, 653-0838, Japan

It has recently become clear that the human genome is almost entirely transcribed and leads to generation of many thousands of non-protein coding RNAs (ncRNAs). Diverse ncRNA species include short miRNAs, piRNAs and much longer ncRNAs (lncRNAs). Although the involvement of miRNAs in various biological processes including cell differentiation is increasingly evident, the biological role of much more abundant class of lncRNAs is currently unknown. In this study, we identify lncRNAs potentially involved in retinoic acid (RA)-induced neuronal differentiation of human neuroblastoma cell line SH-SY5Y. A custom-built 60-mer oligonucleotide microarray platform, which comprises human cDNA sequences covering known, similar to known, putative lncRNAs and unclassifiable transcripts, and those encoding hypothetical short proteins (less than 80 amino acids ORF), was used for this purpose. Our array profiling in this system revealed that 874 (6.7%) of 12,974 lncRNAs were regulated significantly across the time course of RA response. A smaller fraction (1.8%) of protein coding genes was responsive to RA as determined using Human Genome U133 Plus 2.0 Array (Affymetrix). Furthermore, RA-regulated lncRNAs could be grouped into seven clusters based on their different temporal patterns of expression similar to the coregulated groups of protein-encoding transcripts. Genome mapping analysis revealed that 358 (41%) of 874 significantly regulated lncRNAs were located away from any gene loci (intergenic lncRNAs) with the rest of those overlapping with exonic/intronic regions (intragenic lncRNAs). About 20% (111/516) of the intragenic lncRNAs were mapped to the opposite (antisense) strand of protein-coding genes. Possible regulatory interactions between sense/antisense pairs were assessed by monitoring correlation of expression with time. We found that 27 of 111 sense/antisense pairs showed various patterns of reciprocal regulation, where one transcript was induced while the other declined in response to RA. Some of these antisense lncRNAs may have biological functions relating to a classical antisense mechanism. Interestingly, a number of these antisense lncRNAs overlap with the sense strand of protein-coding genes implicated in the processes of cell growth and differentiation. This study represents an important early step in appreciating the potential importance of lncRNAs in neuronal development and identifies the set of lncRNAs for future functional studies.

536: Alu repeat mediated transcriptome diversity through antisense, editing and exonization

Amit K. Mandal, Mitali Mukerji

Functional Genomics Unit, Institute of Genomics and Integrative Biology, Mall Road, D.U. North Campus, Delhi 110007, India

Primate specific Alu repeats that are non-randomly distributed in the human genome are increasingly being implicated in every aspect of genome organization from transcriptome, proteome to epigenome. Genome wide studies have revealed that a major fraction of alternatively spliced, edited and antisense transcriptome comprises Alu repeats. We hypothesize that tissue-specific exonization, antisense formation and editing of genes could increase the transcriptional repertoire of the genome. For instance Alu exonization is favored when they are present in an anti-sense orientation. These exonized Alus can be anti-sense targets of transcribing Alus with *trans* or *cis* origin from the negative strand. Editing of Alus within exonized transcripts could protect them from anti-sense. Thus, the same transcripts could have different fates depending on the relative dynamics of exonization, editing and antisense. As a first step towards addressing the hypothesis we created a dataset of Alu edited, exonized and antisense transcripts through data mining diverse genome wide studies. This was followed by mapping the coordinates of each of these repeats onto the corresponding genes. For identifying exonized transcripts we screened the RefSeq, H-Invitational and dbEST databases and identified a total of 421,683 transcripts which map to 13,757 genes. We identified antisense Alus from 790,442 antisense tags retrieved from a dataset of >9900000 SAGE tags compiled from NCBI ftp. These were from 6,546 unique UniGene clusters which mapped to 1,611 unique genes. About 6,947 Alu edited transcripts were retrieved from supplementary data of three genome-wide studies and mapped to 1,435 genes. Comparison of these three datasets revealed a common set of 337 genes. Further GO analysis revealed distinct enrichments of overlapping biological processes and pathways e.g., genes involved in nucleotide metabolism, cell cycle, and inflammation pathways. The common set of genes is significantly enriched for KRAB box transcription factor, methyl transferases and hydrolases. Using a novel bioinformatic method we also studied the tissue-specific expression of Alu exonized transcripts with antisense tags. We observed large number of cases for possible Alu sequence mediated transcript regulation and tissue-enrichment. This comprehensive framework for studying regulatory networks arising from interplay of Alu

exonization, antisense and editing would provide insights into the functional significance of a major inhabitant of our genome.

537: Probing the involvement of HSF elements within Alu repeats in global heat shock response

Rajesh Pandey, Amit Chaurasia, Amit K. Mandal, Mitali Mukerji

Institute of Genomics and Integrative Biology, CSIR, India, Mall Road, Delhi, India

Alu repeats, which are retro-transpositional elements harbor a large number of active transcriptional factor binding sites, including heat shock factor (HSF) binding sites that are distributed non-randomly in the human genome. These elements can also transcribe Alu RNA whose levels increase in response to heat shock stress, cancerous conditions etc. Recently it has also been shown that Alu RNA formed during heat shock adopt a tertiary structure which can act as a transcriptional repressor. Besides Alu repeats are exonised in a large number of transcripts when present in the anti-sense orientation. These Alu exonised transcripts can form dsRNA with transcribing non coding Alu RNAs formed in cis or trans and can thus be targets for editing, anti-sense etc. We have analyzed global expression profiles to delineate the diverse roles of Alu elements in response to heat shock. We exposed the Hela cells to heat shock at 450°C for 30 min followed by 2 h incubation at 370°C for recovery from stress, followed by RNA isolation. Three biological replicates of treated and untreated RNA were hybridized to Illumina HumanWG-6 v2.0 48K bead chip for global expression analysis. Data was analyzed using Illumina Bead studio. Amongst the differentially expressed genes 2,439 were upregulated and 1,341 were downregulated. These included the well reported HSF genes. There was no significant difference in the density of Alus in the upregulated or downregulated genes either in the upstream or genic regions. We next analyzed consensus HSF binding sites within the Alus using an in house developed program 'Promo' a probabilistic regulatory element finding program, which uses prior information and Hidden Markov Model(HMM), in its core. We observed significantly higher number of genes ($P < 0.0000$) having Alus harboring HSF sites in anti-sense orientation both in the down regulated as well as up-regulated genes. Interestingly, amongst the genes harboring HSF in the anti sense Alus, the representation of down regulated genes was significantly ($P < 0.03$) higher and enriched (after Bonferroni correction) in pathways related to p53, Ras, EGF, PDGF signaling and angiogenesis. These results suggests that Alu containing regulatory sites (HSF in this case) could modulate response to heat shock through down regulating core processes related to cell proliferation through antisense mechanism. Further experimental validations of these results are in progress.

538: Role of transcriptional dysregulation in polyglutamine mediated neuronal cell death

Beena Pillai, Tanay Ghosh, Reema Roshan, Abdul Hannan

Institute of Genomics and Integrative Biology (IGIB), Mall Road, Delhi 110007, India

Expansion of CAG repeats in coding and non-coding regions of genes is associated with neurodegenerative disorders. In coding regions, the CAG repeat corresponds to a polyglutamine stretch in the protein. Polyglutamine expansion in the protein leads to formation of intranuclear protein aggregates and is correlated to neurodegeneration in specific regions of the brain. However, the molecular effects that lead to neurodegeneration have remained elusive. We have explored the role of transcriptional dysregulation by TATA-box binding protein (TBP) containing an expanded polyglutamine stretch in a mouse neuronal cell culture based model. Polyglutamine expansion in the N-terminal region of TBP is associated with a neurodegenerative disorder called Spinocerebellar ataxia 17 (SCA17). We find that mouse neuronal cells expressing a variant of human TBP harboring an abnormally expanded polyQ tract not only form intranuclear aggregates, but also show global transcription dysregulation. We identified differentially expressed genes from cDNA and high-density oligonucleotide array platforms and validated the differential expression at different time-points using independent methods like real-time PCR and western. We have earlier shown that the differentially expressed genes include the Voltage dependent Anion Channel1 (VDAC1), whose overexpression has been shown to induce apoptotic cell death in other cell types. We monitored cytochrome *c* release overtime and found concomitant upregulation of VDAC1 and increase in cytoplasmic cytochrome *c*, followed by apoptosis. Overexpression of VDAC1 could also induce the same effects in neuronal cells. Further, we found that interferon induced genes and the signal transducer and activator of transcription (STAT1) gene were differentially expressed in cells expressing an expanded polyglutamine allele of TBP, suggesting the putative role of STAT1 and subsequent involvement of inflammation in SCA17. The recent discovery of large number of small cellular RNAs, called microRNAs suggest that the transcriptome and proteome are significantly influenced by their post-transcriptional regulation. We explored changes in the expression profile of microRNAs in cells expressing an expanded polyglutamine allele of TBP, compared to a normal allele. The differentially expressed microRNAs and mRNAs fall into similar cellular pathways and provide a comprehensive view of mutually reinforcing pathways of cellular response in polyglutamine diseases.

539: Dissecting RNA mediated gene silencing pathways in murine embryonic stem cells

[1]**Michael Rossbach**, [1]Prasanna Kolatkar, [2]Kurt Zänker

[1]Genome Institute of Singapore, 60 Biopolis Street, #02-01 Genome, Singapore, [2]University of Witten/Herdecke, Stockumer Strasse 10-12, Witten, Germany

In this study, Dicer-deficient mouse embryonic stem (ES) cells (A3 cells) were rescued with the two Dicer proteins from *Drosophila melanogaster*, dcr-1 and -2. This resulted in a structure and function analysis of Dicer with the potential of specifically rescuing either the miRNA or the siRNA pathway. To dissect these two RNA mediated gene silencing pathways, dcr-1 and dcr-2 from *Drosophila* were separately targeted into dcrΔ/Δ (A3) cells. The phenotypical rescue of these cells through gene targeting experiments with hDicer (mammals have only one dicer gene) served as a control and was accomplished. The successful gene targeting of the rosa locus was confirmed by Southern analysis; protein expression of DCR-1 and -2 could be shown by Western blot analysis using DCR-1 or DCR-2 specific antibodies, respectively. Since dcr-1 is believed to play an important role in the processing of microRNAs, the processing of microRNAs precursors into mature miRNAs was observed after successful gene targeting; the processing of microRNAs was shown by Northern analysis. In contrast, cells that had been targeted with *Drosophila* dcr-2 did not show any processing of microRNAs but siRNA processing. DCR-2 is believed to mediate the processing of siRNA in cells. Since such important processes like V(D)J recombination in B cells are regulated by (antisense) siRNA, it is of high value to have cells on hand that either express the miRNA or siRNA pathway to study these regulatory processes and especially their malfunctions in tumor genesis.

540: Quality control of miRNA in biological extractions

R. Salowsky, M. Greiner, M. Gassmann

Agilent Technologies, Hewlett Packard Str. 8, 76337 Waldbronn, Germany

In recent years, hundreds of microRNAs have been discovered. Although their functional roles remain to be classified in detail, accumulating experimental data suggest a fundamental involvement in cancer development and progression. The expression level of individual miRNAs is typically quantified by methods like Northern Blot, RT-PCR or microarray hybridization. One of the major drawbacks in miRNA research is the lack of adequate analytical methods for small RNA extractions before going in the subsequent analysis of the miRNA fraction. As different studies showed that microRNA profiles could be of value in cancer diagnosis, the quality control and quantification of miRNA samples becomes more and more important. Here we present a highly sensitive microfluidic 'high resolution' assay for the analysis of small nucleic acids. The assay utilizes a novel sieving matrix formulation and a special dye system which allows detecting miRNA fraction down to a concentration of 50 pg. It measures integrity, size and concentration of small RNA and is especially calibrated for miRNA. In conclusion, our approach allows the implementation of miRNA quality control in a miRNA expression workflow by monitoring miRNA fractions derived from different tissues or extraction methods.

541: In silico models of microRNA regulatory networks in host-pathogen crosstalk

Vinod Scaria, Manoj Hariharan, Beena Pillai, Souvik Maiti, Samir Brahmachari

G.N. Ramachandran Knowledge Center for Genome Informatics, Institute of Genomics and Integrative Biology (IGIB-CSIR), Mall Road, Delhi 110007, India

MicroRNAs are a recently discovered class of endogenous regulatory RNAs. These small RNAs are encoded by genomic regions which transcribe for transcripts harboring imperfect palindromes. The hairpin precursors (pre-miRNA) arising from these regions are then further processed to from $\sim$22 nucleotide RNAs, which bind to near complementary regions in the 3′UTR of target transcripts in association with a protein complex. This results in either degradation of the target transcript or reduction of its translational ability in a sequence specific manner. Recent high throughput experimental discovery of microRNAs in many eukaryotes and viruses has revealed a huge and hitherto unknown regulatory network mediated by microRNAs. We used a consensus computational approach for microRNA target prediction. The pipeline includes highly accurate tools developed in-house for prediction of thermodynamically feasible microRNA target interactions based on target site accessibility and de novo prediction of viral microRNAs from genomic sequences. We identified several cases of microRNAs mediating crosstalk between the host and pathogen which suggest novel and critical roles for these small regulators in host-pathogen interaction. Analyses revealed that five Human microRNAs could target four critical genes in Human Immunodeficiency Virus including Nef which is involved in slow progression of HIV, and two human microRNAs targeted the Hemagglutinin and Neuraminidase genes in Influenza virus suggesting its involvement in disease progression and host-specificity, respectively. Similarly twenty of the thirty-two microRNAs encoded by oncogenic virus EBV could potentially targeted transcripts of human genes involved in apoptosis and tumor suppression, suggesting a novel mechanism of viral oncogenesis. Apart from this, viral integration in the genome and genetic instabilities induced by viruses can significantly influence microRNA regulatory networks in the host cell. Based on the data, we suggest a comprehensive in silico model for the role of microRNAs in host-pathogen interaction and viral oncogenesis. This evidence would help in elucidating novel biomarkers for pathogen susceptibility and disease progression. Understanding the intricacies of microRNA regulation in host–virus crosstalk has enabled us to design synthetic regulators with potential therapeutic applications.

References

Hariharan M et al (2005) BBRC
Scaria V, Jadhav V (2007) Retrovirology
Scaria V et al (2006) Retrovirology
Scaria V et al (2007) Cell Microbiol

542: Argonaute 2 is dispensable in self-renewal and differentiation of murine ES cells

P. Chandra Shekar, Adnan Naim, D. Partha Sarathi, Satish Kumar

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007 (Andhra Pradesh), India

One of the important conserved gene regulation mechanisms in eukaryotes is effected by small RNAs through RNA-interference (RNAi). Many components and pathways of RNAi play pivotal role in self-renewal and differentiation of variety of stem cells. Argonaute 2 (Ago2), also known as slicer, is one of the key components of RNA induced silencing complex (RISC), the catalytic machinery in RNAi. In *Drosophila* Ago2 is involved in microRNA mediated germ line stem cells proliferation. Mammalian Ago2, encoded by the EIF2C2 gene, has been shown to be essential in siRNA and miRNA pathways but its role in maintenance and differentiation of stem cells is not understood. To study the role of Ago2 in murine embryonic stem (ES) cells, we have generated Ago2 deficient ES cells by homologous recombination followed by loss of heterozygosity. Ago2 deficient ES cells are defective in siRNA mediated gene silencing as was expected. These cells can self-renew and can be passaged for more than 30 passages. The mutant cells give rise to all the three germ layers when differentiated into embryoid bodies and these cells can also form teratocarcinomas when injected subcutaneously into the nude mice. Interestingly, we observed a significant reduction in the mature miRNA levels in Ago2 deficient ES cells in comparison to the wild type ES cells. The reduced miRNA levels was observed both in ES cell specific and non ES cell specific miRNAs. Thus, absence of Ago2 leads to significant global reduction in biogenesis of miRNA. Unlike the Dicer deficient ES cells, which are defective in differentiation, Ago2 deficient ES cells are capable of both self-renewal and differentiation.

543: Sequence based method of Riboswitch detection

Payal Singh

Centre of Computational Biology and Bioinformatics (CCBB), Jawaharlal Nehru University, New Delhi, India

RNAs not only play conventional role during protein synthesis but are also key players in gene regulation. They can perform a variety of functions ranging from showing catalytic activity to regulating gene expression. A large number of cis-acting as well as trans-acting RNAs have been found to play an important role in regulating gene expression. Riboswitches are one such class of regulatory RNAs. They are involved in post-transcription cis-acting gene regulation and act as high affinity receptors for small molecules. They are found in

the 5′ untranslated region and consist of two domains—the aptamer domain and the expression platform. The aptamer domain is highly conserved both at sequence and structure level and act as metabolite sensor. The binding of the target ligand to the aptamer domain causes the conformational changes in the expression platform, which lies just downstream of the aptamer domain. This conformational change modulates the gene expression by either premature transcription termination or by preventing the translation initiation. A number of tools have been developed to detect Riboswitches across different genomes. They are either extremely slow, as they incorporate both the sequence as well as structure level information or very specifically designed for certain riboswitch families. The major drawback of sequence as well as structure based methods is that they are computationally intensive and require large CPU time. So there is a need for fast and sensitive method of riboswitch detection. The aim of our study is to combine the strengths of various profile based searching programs available, so as to develop efficient program which is both sensitive and requires less CPU time. We applied our approach to detect FMN riboswitch across various bacterial genomes. When compared with the existing approach of non-coding RNAs detection, based on Covariance Model, our method has been found to have comparable results in terms of sensitivity as well as specificity, and is about hundred times faster. Using this approach, we have searched all microbial genomes for different riboswitch families and now extending search to the large fungal and eukaryotic genomes.

544: Maternal inheritance of a microRNA with pleiotropic tissue specific effects

[1]**Kartik Soni**, [1]Mukesh Lalwani, [1]Ashok Patowary, [1]Manoj Hariharan, [1]Tanay Ghosh, [1]Vinod Scaria, [2]Shipra Bhatia, [2]Vani Brahmachari, [1]Sridhar Sivasubbu, [1]Beena Pillai

[1]Institute of Genomics and Integrative Biology, Mall Road, New Delhi, India, [2]Ambedkar Center for Biomedical Research, University of Delhi, Delhi, India

MicroRNAs are short 21–23 nt non coding RNAs present endogenously that regulate gene expression by inhibition of translation or degradation of messenger RNA molecules. We used a bioinformatics approach to predict miRNAs that are maternally inherited and experimentally studied the role of these miRNAs during early development. We analyzed publicly available gene expression profiling data from various stages of *Drosophila* development since *Drosophila* zygotic transcription is delayed and the time at which maternal transcription starts is well defined. We classified transcripts into two mutually exclusive classes—maternal and zygotic. Target prediction by a consensus approach was used to identify miRNAs whose targets are over represented in the 3′UTR of mRNAs. To validate our computational predictions we isolated RNA from *Drosophila* embryos at three time points. Using a primer extension based assay we studied the endogenous expression levels of seven predicted maternally inherited miRNAs. We observed that dme-mir-34 showed the highest expression level in pre-zygotic transcription stages. We did a similar experiment in Zebrafish as it is easy to obtain unfertilized eggs in this system, thus avoiding any zygotic RNA. Our experiments led us to the conclusion that mir-34 is maternally inherited in *Drosophila* and in Zebrafish. Next, we used a comparative genomics approach to look for conservation of mir-34 and its targets in Human, Zebrafish and *Drosophila*. In lower organisms like *Drosophila* the miR-34 family consists of a single miRNA whereas in higher organisms it is represented by several family members e.g., hsa-mir-34a (MIRN34A), hsa-mir-34b (MIRN34B), hsa-mir-34c (MIRN34C) in humans. Amongst the members of the mir-34 family, hsa-mir-34a (MIRN34A) was found to be closest to dme-mir-34 and dre-mir-34. Experimental evidence from literature suggests that hsa-mir-34a (MIRN34A), under the regulation of p53, a tumor suppressor gene, leads to suppression of cell proliferation and promotes apoptosis. The targets of hsa-mir-34a (MIRN34A), include NOTCH1, DLL1 (ligand of NOTCH1), E2F3, and BCL2. We are validating Zebrafish homologs of these targets to assess the effect of maternally inherited mir-34 in Zebrafish. In mouse, we found two alternatively processed transcripts from the β-actin (ACTB) gene, one of which harbors a mir-34 target. These studies suggest that mir-34, a potential key regulator of cell cycle regulation, early embryonic development and neural development, is maternally inherited.

545: Screening for human miRNA targets on the genetic pathway in type II diabetes mellitus

[1]**Venkateswara Rao Sripathi**, [1]Utpal Bhadra, [2]Manika Pal-Bhadra

[1]Functional Genomics and Gene Silencing Group, Centre for Cellular and Molecular Biology, Hyderabad 500007, India, [2]Indian Institute of Chemical Technology, Department of Chemical Biology, Hyderabad 500007, India

Type II diabetes is a metabolic disorder, characterized by insulin resistance, relative insulin deficiency and hyperglycemia. It maintains constant high levels of blood glucose and associated with many other traits including obesity, renal and cardiovascular failures. A complex interaction between the physio-chemical environment and nearly 50 genes is involved in the genetic pathways of Type II diabetes. It is well established that small regulatory RNA commonly referred as microRNA controls post-transcriptional mRNA degradation and translational inhibition in animals. They play an important role in various periodical processes including cell development, cellular differentiation, apoptosis and many more other multi-factorial diseases. It is recently been reported that few candidate of 540 predicted microRNAs in human plays an important role in insulin secretion and glucose homeostasis. Several in silico approaches for screening of microRNA targets predicts several hot spot sequence, complementary to several microRNAs in the coding region and the 3′UTR of thirteen genes related to type II diabetes. Several algorithms identify same microRNA targets by very rigid statistical parameters (score range 150–200 and energy level <-25). Nearly ten potential target sites have been identified in each gene. The location of the clustering spots is quite variable. The miRNA cluster is located on the coding region of the ten selected genes while cluster of targets forms in 3′UTR in only two genes. Therefore majority of the genes undergo degradation post-transcriptionally by degrading specific mRNA while few candidate shows translational inhibition. We are currently approaching for verifying such microRNA targets, which can be further used as a natural biomarker for early clinical diagnostics and to identify an anti-microRNA based effective therapy for Type II diabetes.

546: A SOLiD™ approach to small RNA discovery and profiling using ligase enhanced genome detection

[3]**Raimo Tanzi**, [2]Robert Nutter, [1]Scott Kuersten, [2]Catalin Barbacioru, [2]Melissa Barker, [1]Jian Gu, [2]David Wang, [2]Roland Wicki, [2]Francisco de la Vega, [1]Bob Setterquist

[1]Ambion Inc., Austin Texas, USA, [2]Applied Biosystems, Foster City California, USA, [3]Applied Biosystems, Milan, Italy

The Applied Biosystems next-generation sequencing platform, SOLiD™, produces 10–100's millions of short reads (25–50 bp) in a single run. Sequencing is carried out via dibase sequential rounds of ligation with high fidelity and high read quality. A new RNA library

protocol, Legend Small RNA Expression Kit, has been developed and used on SOLiD™ system, requiring low amounts of total RNA and resulting in sequence ready samples in less than a day. In this study we describe some of the bioinformatics challenges and solutions produced by this specific type of data generated by SOLiD™ system from multiple human RNA libraries. Total RNA was isolated from human lung and placenta tissues and subsequently fractionated using flashPAGE. The small RNA fraction was converted to double stranded DNA templates suitable for sequencing using the Legend Small RNA Expression Kit. Five independent small RNA libraries were clonally amplified by ePCR and run on different SOLiD™ instruments from Foster City and Beverly (four quadrants on each slide). Each library generated between 80 and 175 million of 35 bp (colors) length reads. There are more than 2,000 known unique small RNA sequences, with lengths ranging from 18 to 31 bp. The short and variable sizes of small RNA targeted in these experiments make the reads generated by SOLiD™ system to contain a P2 adaptor sequence 'tail', and therefore unsuitable for current matching pipeline. A new RNA matching pipeline has been developed and implemented to overcome these main issues, allowing us to map 55–60% of the reads generated from one library. Mapped data was used to answer question related to detectability of known and new isomirs (miRNA sequences with variable 5′ and 3′ ends) which in some cases the most abundant version differs from Sanger reference sequence. Additionally, library saturation, reproducibility and variability estimates are generated using these libraries. Fold changes generated from SOLiD sequencing data were also compared to those of 350 Taqman® miRNA assays (TLDA cards). Significant correlation levels were observed confirming the applicability of this approach for small RNAs expression profiling. Moreover, more than 3,000 potentially novel miRNAs or non-coding RNAs were discovered. These potential novel small RNAs are currently being further validated.

547: Micro RNA, simple sequence repeats and neuron specific genes

[1]**S. Trivedi**, [2]G. Ramakrishna, [1]S. Jethra, [1]H. Soni

[1]Computational Biol. Lab., Department of Zoology, JN Vyas University, Jodhpur (Raj.), 342011, India, [2]Lab. of Cancer Biology, Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad (AP), India

Micro RNA (miRNA), a class of non-coding RNA, are known to play a role in gene regulation that may influence development, differentiation or tissue/organ specific gene expression. Though the exact mechanism of miRNA action in neuronal differentiation and neuron specific gene expression is still under investigation, it has been found that this class of non-coding RNA not only modulates neuronal differentiation but also plays a role in neurodegeneration/protection. Some of the neuron specific miRNA target genes are associated with neurodegenerative disorders due to expansion of repetitive DNA called simple sequence repeats (SSRs). In order to investigate whether neuron specific miRNA have repeat sequences, in silico analysis of human miRNA was carried out. Further, an association of miRNA with repeats was sought with known neuron specific target genes to seek whether repeats were present in target genes as well.

548: Prediction and validation of precursor miRNAs in human genomic sequences

[1]**Candida Vaz**, [2]Priyanka Nath, [1]Sonika Tyagi, [2]Sachin Maheshwari, [1]Alok Bhattacharya

[1]Centre for Computational Biology and Bioinformatics, School of Information Technology, Jawaharlal Nehru University, New Delhi, India, [2]Department of Computer Sciences, Indian Institute of Technology, New Delhi, India

MicroRNAs (miRNAs) are small regulatory non coding RNAs, that regulate the expression at the post-transcriptional level by arresting the translation of mRNAs. These miRNAs are encoded in the intergenic region of the genome and to some extent in the intronic regions too. The generation of miRNAs from their genes involves complex enzymatic processing which is manifested in the creation of hairpin like pre-miRNAs of average 70 nt in length. Mature miRNAs of average 22 nt in length are then generated by further processing. Several computational methods have been devised so far to identify the miRNAs, as experimental methods are slow and miRNAs are difficult to isolate by cloning, due to low expression, stability and tissue specificity. Not all of these methods are suitable for identification of novel miRNAs or can be applied to scan large genomes. In this presentation we describe two different approaches for miRNA detection and show results of genome wide scans. The first miRNA prediction method is based on 'Stochastic Context Free Grammar' and is trained on human miRNAs, and the second method models the pre-miRNA detection as a path finding problem in a graph obtained from the genome using 'Suffix Trees' and includes a set of filters obtained by analyzing all the known human miRNAs. These filters are used to remove the false positives. The predicted miRNAs are validated by finding non coding transcripts, Drosha cutting sites and potential targets.

549: Noncoding RNAs in nervous system function

Claes Wahlestedt

The Scripps Research Institute, 5353 Parkside Dr, Jupiter, FL 33458, USA

In the more than 7 years since the first publications on the entire human genome, the number and general architecture of protein coding genes, comprising only some 1.2% of the genome, has remained remarkably stable. By contrast, surprising changes have come from noncoding regions, which have created an entirely new paradigm for all aspects of genome-related research. Several high-throughput efforts, by ourselves and others, have provided strong evidence that, in contrast to earlier understanding, a great majority of the mammalian genome is transcribed into noncoding RNAs. Why do organisms and perhaps the mammalian brain in particular, consume so much of its energy to transcribe all these RNA transcripts? And is this newly recognized universe of noncoding RNAs of functional (regulatory) importance? A primary objective of higher life forms is to gather larger amounts of more relevant information, and to apply that information toward the improvement of its future information-gathering ability. Noncoding RNAs with regulatory function may be critical in this regard. This presentation will highlight some of our recent examples of long and small noncoding RNAs with regulatory activities on specific genes and pathways (e.g., references 1–4).

1. Katayama S, Tomary T, Kasukawa T, Waki K, Suzuki M, Nishida H, Choo YP, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang K, Mattick J, Hume D, Lipovich L, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk A, Mottagui-Tabar S, Lenhard B, Wahlestedt C (2005) Antisense transcription in the mammalian transcriptome. Science 309:1564–1566
2. Khalil AM, Faghihi MA, Modarresi F, Brothers SP, Wahlestedt C (2008) A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. PLoS ONE 3(1):e1486

3. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St. Laurent III G, Kenny PJ, Wahlestedt C (2008) A noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase expression. Nature Med (in press)

4. St-Laurent III G, Wahlestedt C: Noncoding RNAs (2007) Couplers of analog and digital information in nervous system function? Trends Neurosci 30:612–621