Contents

FOXC2, Goosecoid, SIP1, Slug, Snail and Twist Induce EMT in Human Breast Cancer Cell Lines Through a Regulatory Network

Alicia Zhou, Kimberly Hartwell, Sendurai Mani, Jing Yang and Robert Weinberg1
Whitehead Institute for Biomedical Research, 1Massachusetts
Institute of Technology, Cambridge, MA 02142

During metastasis, carcinoma cells undergo a morphological change known as the epithelial-mesenchymal transition (EMT), which increases cell motility and promotes invasion. In many species, EMT is known to normally occur during embryonic development, aiding such key steps as gastrulation, neural crest emigration and organogenesis. E-cadherin, a cell-cell adhesion molecule, is a known EMT regulator whose down-regulation is required for EMT. Recent studies have indicated that another member of the cadherin family, N-cadherin, plays an inverse role to that of E-cadherin; its up-regulation promotes the induction of EMT. Six transcription factors – FOXC2, GSC, SIP1, Slug, Snail and Twist – have recently been identified as possible EMT regulators. Here, we study the expression, at the transcriptional level, of each of these transcription factors in order to characterize their roles in EMT. By RT-PCR analysis, we observed that the expression of FOXC2, SIP1, Slug, Snail and Twist was inversely correlated with that of E-cadherin and directly correlated with that of N-cadherin in the human breast cancer cell lines: MCF-7, MDA MB 231, MDA MB 435 and SUM 1315. To determine the regulatory role of these transcription factors, we created an in vitro over-expression system in the breast cell line, HMLE. Using this system, we have identified a regulatory network for these six transcription factors – FOXC2, GSC, SIP1, Slug, Snail and Twist – in human breast cancer cells.

SIRT4 regulates insulin secretion and production in pancreatic beta cells

Marcia C. Haigis1, **Danos C. Christodoulou**1, Kevin M. Haigis2, Gil Blander1, and Leonard Guarente1
1Department of Biology and 2 Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

In investigating the role of Sir2 proteins in aging, we characterized SIRT4, one of the mammalian homologues of Sir2. SIRT4 was found to encode a mitochondrial protein that possesses NAD-dependent ADP-ribosyltransferase activity. We studied the expression profile of SIRT4 and found that it was strongly expressed in pancreatic islets. The high levels of SIRT4 in beta-cells, prompted us to investigate whether SIRT4 is involved in insulin secretion. Reducing SIRT4 levels in murine beta-cells by RNA interference increased the levels of insulin contained and secreted by these cells. The effect of SIRT4 on insulin secretion appeared to be mediated by glutamate dehydrogenase; this inhibition required NAD as a cofactor. These results indicate that SIRT4 plays a role in diabetes and perhaps in aging.

Scanning for Alternative Conserved Exons

Name: **Eric Van Nostrand**
Faculty Advisor: Christopher Burge
Graduate/Postdoc mentors: Gene Yeo, Dirk Holste

Alternative pre-messenger RNA splicing affects a majority of human genes and plays important roles in development and disease. Alternative splicing (AS) events conserved since the divergence of human and mouse are likely of primary biological importance, but relatively few such events are known. Here we describe sequence features that distinguish exons subject to evolutionarily conserved AS, which we call "alternative-conserved exons" (ACEs), from other orthologous human/mouse exons, and integrate these features into an exon classification algorithm, ACEScan. Genome-wide analysis of annotated orthologous human-mouse exon pairs identified ~2,000 predicted ACEs, 9-fold more than found by transcript evidence alone. Alternative splicing was verified in both human and mouse non-disease tissues using an RT-PCR-sequencing protocol for 21 of 30 (70%) predicted ACEs tested, supporting the validity of a majority of ACEScan predictions. In contrast, AS was observed in mouse tissues for only 2 out of 15 (13%) tested exons that had EST or cDNA evidence of AS in human but were not predicted ACEs, and was never observed for eleven negative control exons in human or mouse tissues. Predicted ACEs were much more likely to preserve reading frame, and less likely to disrupt protein domains than other AS events, and were enriched in genes expressed in the brain and in genes involved in transcriptional regulation, RNA processing and development. Our results also imply that the vast majority of AS events represented in the human EST database are not conserved in mouse.

Control of Bacterial Chemotaxis through Regulated Gene Expression

Name: **Frederick Tan**
Faculty Advisor: Drew Endy

Engineering recombinant controls over bacterial chemotaxis is one possible way to program where bacteria move and what chemical signals produce a motile response. Control over motility demonstrates the ability to engineer novel phenotypes, which is an important step in the creation of synthetic microbial devices. Recombinant control over bacterial chemotaxis is achieved by regulating the expression of cheY, a gene coding for an essential chemotaxis signaling protein. Inducible expression of cheY in Escherichia coli mutants lacking cheY allows for the controlled restoration of chemotaxis signaling. Different gene expression levels are generated by varying the strength and placement of regulatory regions such as the promoter, the ribosome binding site, the protein degradation tag, and the transcriptional terminator. Ten different variations of a recombinant cheY construct have been designed and are currently being tested for an ability to produce natural CheY levels maintained in E. coli. One construct has been fully synthesized and tested. Although some problems involving leaky transcription were encountered, the development of more stringent transcriptional controls will limit its influence on cheY expression. Additionally, the synthesis and testing of the remaining nine constructs will be continued while considering other possible designs for regulated cheY expression. Isolating a recombinant construct capable of producing natural levels of CheY is expected in the near future.

Prevalence of the photosystem reaction center genes psbA and psbD in viruses infecting marine unicellular cyanobacteria.

Sullivan, M.B.*, D. Lindell*, **J.A. Lee**, L. Thompson, S.W. Chisholm.
Massachusetts Institute of Technology, Cambridge, MA 02139

Prokaryotic marine viruses (phages) are the most abundant biological entities on earth. They play at least two important roles: altering nutrient cycling through as predators and influencing evolution by facilitating horizontal gene transfer. Here we focus on phages (cyanophages) that infect the marine cyanobacteria, Prochlorococcus and Synechococcus, because these hosts are globally significant primary producers. The recent discovery of host-like photosynthesis genes in the genomes of eight out of nine sequenced marine cyanophages raises several questions as to how widespread such genes are in the cyanophage community, the evolutionary basis for their presence, and whether they are functional. This study examines the prevalence of two core photosystem reaction center genes, psbA and psbD, in cultured Synechococcus- and Prochlorococcus-infecting Podoviridae and Myoviridae. Thirty-nine cyanophage isolates were screened for both genes by PCR and sequenced for subsequent analyses. psbD was found in 31% (12 of 39) and psbA in 95% (37 of 39) of the cyanophages examined. Neither Synechococcus-infecting Podoviridae appear to have psbA, while psbD was not found in any of the fifteen Prochlorococcus- or Synechococcus-infecting Podoviridae and was uncommon among Procholorococcus-infecting Myoviridae (1 of 10). We suggest that these genes are functional and help maintain host photosynthesis during phage infection, as has been demonstrated for one host/phage pair. In addition, phylogenetic analyses were conducted PsbA and PsbD sequences, revealing a similar tree topology to those of many other genes previously examined (including 16S rDNA and the ITS region). One striking exception, the Prochlorococcus MIT9211 PsbD sequence, may indicate a horizontal gene transfer event. In addition to the potential evolutionary and physiological implications, the finding that photosynthesis genes are common elements of cyanophage genomes may offer a culture-independent means of identifying cyanobacteria-specific phages in the environment.

Transcriptional Silencing of a Transgene in C.elegans induced by RNAi

Alla Grishok1, **Jina L. Sinskey**1 and Phillip A. Sharp1,2
1Center for Cancer Research, 2Department of Biology and
McGovern Institute, Massachusetts Institute of Technology, 40
Ames Street, E17-529, Cambridge, MA 02139, USA.

RNA interference, or RNAi, refers to the ability of double-stranded RNA (dsRNA) to interfere with the expression of the corresponding endogenous gene. Most gene silencing due to RNAi in C.elegans occurs at the post-transcriptional level. We observed transcriptional silencing when worms containing the elt-2::gfp/LacZ transgene were fed RNA produced from the commonly used bacterial L4440 vector. The transgene and the vector share plasmid backbone sequences, and the homology between the two sequences leads to silencing upon feeding-induced RNAi initiation. A library of dsRNA-expressing bacteria with genes cloned into the L4440 vector was used in a candidate RNAi screen to identify genes essential for silencing of the GFP transgene. From the screen, we were able to determine which genes play a role in silencing of the transgenic strain. We found that the silencing depends on RNAi pathway genes, including dcr-1, rde-1, rde-4 and rrf-1, and also on alg-1/2 and on the HP1 homolog hpl-2. The latter is a chromatin silencing factor, and expression of the transgene is inhibited at the level of intron-containing precursor mRNA. We therefore define this type of silencing as RNAi-induced transcriptional silencing (RNAi-TGS).

Quantitative analysis of the abundance and rate of formation of peptide-MHC class I complexes on the cell surface

Lakshmi Nambiar and Professor Herman N. Eisen
Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

The abundance of Class I major histocompatibility complex (MHC) proteins on cells varies widely under physiological and pathological conditions. These proteins are especially abundant on dendritic cells which are essential for activating naïve T cells to mount an immune response. However, the expression of Class I MHC is often down-regulated on virus-infected or tumor cells in order to escape detection by CD8+ cytotoxic lymphocytes. These differences imply that expression of Class I MHC is subject to considerable regulation. Previous studies have determined relative steady state levels of Class MHC I expression through flow cytometry, but not in terms of absolute numbers of molecules per cell. We are developing a novel method using infrared fluorescence and doubly-labeled radioactive and fluorescent antibodies to determine the number of Class I MHC molecules per cell, the rate at which these molecules become accessible to peptide ligands, and the time it takes for cell surface peptide-MHC class I complexes to reach a steady state equilibrium. Through such quantitative analysis of MHC regulation, we may be better able to approach understanding and treatment of auto-immune disease, cancers, and infectious diseases.

Inhibition of β-sheet content in the Alzheimer's disease peptide Aβ1-42 by DAPH Analogs

Leslie Rozeboom, Vernon Ingram
Massachusetts Institute of Technology, Cambridge, MA 02142

The Alzheimer amyloid β peptide (Aβ1-42) aggregates into fibrils that form extracellular plaques in the brains of those suffering from Alzheimer's disease. These misfolded protein fibrils are rich in β-sheets and are suspected to be at least one of the causes of the damage caused by Alzheimer's disease. The dye thioflavin T (ThT) binds rapidly and specifically to β-sheet fibrils. ThT fluorescence was used as an assay for β-sheet content. Of more than 3,700 small molecules known to be biologically active, DAPH (4,5-dianilinophthalimide) was the most effective at reducing β-sheets in a high-throughput screen. In this study, we analyzed 11 analogs of DAPH and found one that decreases β-sheet content of aggregated Aβ better than DAPH. This and future DAPH analogs may be promising candidates for an Alzheimer's disease therapy.

Structure-Function Analysis of the Drosophila Het-A Gag Major Homology Region

UROP: **Malcolm Mattes**
Faculty Advisor: Mary-Lou Pardue
Mentor: Svetlana Rashkova

Drosophila melanogaster has an unusual mechanism of maintaining chromosome ends. Lacking telomerase, Drosophila instead has two non-long-terminal-repeat (non-LTR) retrotransposons Het-A and TART, which extend chromosome ends through successive transpositions, producing long head-to-tail arrays at the telomeres. Both retrotransposons encode Gag proteins, as do all retroviruses. Although there is much variation between species in the amino acid sequences of Gag, one region in particular, the Major Homology Region (MHR), is highly conserved, suggesting its importance in the proper functioning of the protein. Although the structure of the MHR in Het-A Gag is not well understood, studies of HIV1 Gag, whose MHR is most similar in sequence to that of Het-A Gag, has provided incite into our understanding of telomeric Gag proteins. These studies have implicated specific amino acid residues in the MHR as necessary for proper folding of the HIV1 Gag MHR, such that mutating any of these residues abolishes viral assembly, maturation, and infectivity. Molecular models suggest that the MHR of Het-A Gag could form a similar structure as that of HIV1 Gag, with conserved amino acids playing a similar role in protein folding. Our studies aim to determine the accuracy of this model, by mutating these conserved residues in the Het-A Gag MHR, and assaying for the mutant proteins, ability to interact with other proteins and target the telomeres as the wild type does in Drosophila cells. Preliminary results from analyzing seven such mutants show that mutating different amino acids has different effects in the targeting of the proteins, giving incite into the importance of these residues in the wild type proteins' activity.

Regulation of Muscle-Specific MicroRNAs

Mina Farkhondeh1, Prakash Rao2, Harvey Lodish1,2
1Department of Biology, MIT, Cambridge, MA 02139.
2Whitehead Institute for Biomedical Research, Cambridge, MA 02142.

MicroRNAs (miRNAs) are an abundant class of ~22 nucleotide RNAs that regulate gene expression. A number of tissue-specific miRNAs have been identified, suggesting that they may play a role in differentiation. Examples of tissue-specific miRNAs include miRNA1 (miR1) and miR133, which are highly expressed in skeletal and heart muscle. We have employed C2C12 myoblasts in order to study the regulation of miR1 and miR133 during differentiation. Levels of miR1 and miR133 increase during the differentiation of myoblasts into myotubes. Genomic analysis has revealed the existence of multiple loci for miR1 and miR133. Two copies of miR1 are present on chromosome 2 and 18, encoding miR1-1 and miR1-2 respectively. Likewise, three copies of miR133 are present on chromosomes 18, 2, and 1, encoding miR133a-1, miR133a-2, and miR133b respectively. In this study we used RT-PCR analysis to determine which of these loci give rise to the primary miRNAs (pri-miRNAs) for miR1 and miR133. We obtained products of the expected size for pri-miR1-1, pri-miR133a-2, and pri-miR133b; these were upregulated in differentiated cells, mirroring the expression pattern of the corresponding mature miRNAs. These results suggest that mature miR1 arises as a result of processing of pri-miR1-1. Similarly, mature miR133 could be processed out of pri-miR133a-2 and pri-miR133b. We are currently confirming the contributions of miR1-2 and miR133a-1. These results will guide future studies aimed at generating mice with loss of function alleles for miR1 and miR133.

Directed evolution of Lanthanide Binding Tag using yeast surface display system

Student Name: **Ryu Yoshida**
Mentors/Advisors: Dr. Bianca Sculimbrene, Langdon Martin, Professor Barbara Imperiali

Lanthanide binding tags (LBTs) are short sequences of natural amino acids, which bind lanthanide ions with high affinity. These short peptide sequences can be easily incorporated into recombinant proteins at the DNA level, and have several characteristics desirable for a biophysical probe. Lanthanide-bound LBTs provide long-lived millisecond luminescence, X-ray scattering power for phase determination of crystal structures, and paramagnetic properties to facilitate structural studies with NMR. Previously, a 17 amino acid-long LBT with low nano-molar affinity was discovered by screening chemically generated peptide libraries. To expand the search for LBTs with even better affinities for lanthanides, directed evolution of yeast surface protein is being implemented. This method allows for rapid screening of larger libraries of over 10^5 clones. In this method, libraries of LBTs are generated at the DNA level, expressed on yeast surface, and treated with V8 protease. As a result, LBTs with weaker lanthanide affinities are selectively digested because lanthanide-bound LBTs are protease-resistant. Yeast cells with undigested LBTs are collected using fluorescence activated cell sorting, and the peptide sequences are recovered by DNA sequencing.

Zebrafish hindbrain differentiation is regulated by nlz and vhnf1 through different pathways

Siddharth Bhola, Elizabeth Wiellette, and Hazel Sive
Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Cambridge, MA 02142.

The vertebrate hindbrain forms the cerebellum, medulla, and cranial nerves. It is responsible for maintaining many basic homeostatic functions, such as breathing and heartbeat. Development of the hindbrain begins early during embryogenesis when a human embryo is about 3-4 weeks old. Some of the first structures visible in the hindbrain are seven transient bulges called rhombomeres (r1 – r7), which form along the head to tail (anteroposterior) axis of the hindbrain and will give rise to different regions of the brain in the older animal. Using the zebrafish as model system, it is clear that long before rhombomeres appear, different regions of the hindbrain have begun to form. These regions are characterized by differential gene expression in domains that often indicate future rhombomeres. Many of these genes encode transcription factors that pattern the hindbrain. The focus of this study was on the interaction between Nlz and Vhnf1 function. The nlz gene family encodes novel zinc finger proteins expressed in the developing hindbrain. nlz functions to specify or maintain rhombomere 4 (r4) identity, and to limit r3 and r5 formation. vhnf1 is also expressed in the hindbrain and promotes r5 and r6 identity while repressing r4 fate. Since these two genes have opposite effects, a loss of function epistasis test was performed to determine whether they act through a common pathway. A combination of genetic mutants and antisense loss of function techniques was used. The double loss of function phenotype was found to be a combination of each single loss of function phenotype. These data suggest that nlz and vhnf1 function in separate pathways.

Contributions of the Hydrophobic Core of Human γD Crystallin to Protein Folding and Stability

Veronica Zepeda, Melissa Kosinski-Collins, Shannon Flaugh, Jonathan King
Department of Biology, Massachusetts Institute of
Technology, Cambridge, Massachusetts 02139

Human γD Crystallin (HγD-Crys), found in the human eye lens, is a 173 amino acid protein that has been attributed with the formation of cataracts. This primarily ß-sheet protein is made up of four Greek key motifs separated into two highly homologous domains. Four tryptophans buried in the hydrophobic core of HγD-Crys - positions 42, 68, 130, and 156 - provide probes of local conformation. The unfolding and refolding of HγD-Crys has been studied as a function of guanidinium hydrochloride (GuHCl) at 37°C and pH 7 (Kosinski-Collins and King, 2003). The wildtype HγD-Crys exhibited a reversible refolding above conentrations of 1.0M GuHCl. Below 1.0M GuHCl the formation of fibrillar aggregates was observed. Single mutants containing a tryptophan to alanine mutation have been constructed and expressed to determine the importance of these hydrophobic residues to the folding and stability of the protein. All trp to ala mutants retained a native like conformation and exhibited unfolding and refolding processes similar to wildtype protein. The formation of aggregates was also observed below concentrations of 1.0M GuHCl. Analysis of unfolding and refolding equilibrium data revealed that the overall stability of the proteins was not significantly decreased in comparison to wildtype HγD-Crys. This data suggests that these four hydrophobic residues do not play a major role in the folding or stability of the protein. Future studies with the trp to ala mutants will investigate whether these residues play a role in the absorption of UV and IR light that reaches the eye lens.