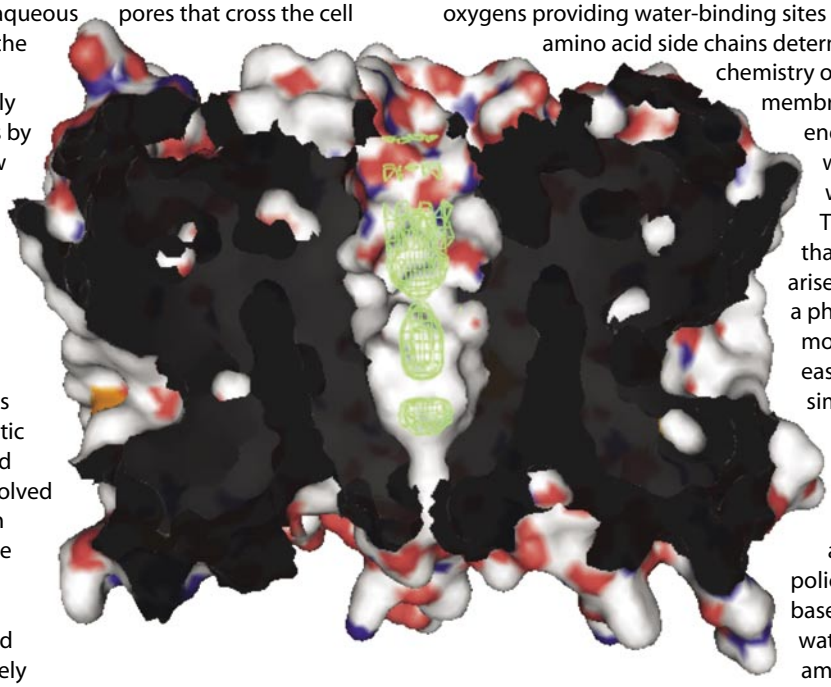
Research Digest

Synopses of Research Articles

Structure of Aquaporin Reveals Mechanism for Transport Selectivity

Biochemists aren't much accustomed to seeing their work in the popular press, save for annual coverage of the Nobel prize in chemistry. This year, Roderick MacKinnon was recognized for working out the atomic structure of an ion channel and Peter Agre for discovering that a major protein found in red blood cells functions primarily as a water channel. Agre went on to establish the family of related channels, which he named "aquaporins." Channel proteins have aqueous pores that cross the cell membrane and regulate the flow of molecules in and out of cells. Water passively pours though aquaporins by osmosis, moving from low to high concentrations of solutes. Solving the structure of these channels provided a platform for exploring the underlying molecular mechanisms that allow the proteins to function as filters and maintain osmotic equilibrium. Robert Stroud and colleagues recently solved the atomic structure of an aquaporin (GlpF) and have now solved the structure of another water channel from *Escherichia coli*, called aquaporin Z, that selectively conducts only water at high rates.



Structure of aquaporin Z

Aquaporins form a large, diverse family of proteins and have been found in bacteria, plants, and animals. There are 11 family members in the human proteome. Less than a decade ago, scientists discovered the aquaporin Z gene (*aqpZ*) in *E. coli*, pointing to the protein's role in regulating water transport in this prokaryote. The high-resolution X-ray structures of recombinant aquaglyceroporin glycerol facilitator (GlpF)—a channel protein that transports both glycerol and water in *E. coli*—determined by the Stroud group in 2000 and of bovine aquaporin 1 (AQP1) from red blood cells, determined a year later, revealed how these aquaporins regulate their transport and selectivity. The aquaporin Z channel protein in *E. coli* can accommodate a flow of water at rates six times higher than GlpF, making it the prime subject for studying the selectivity of a high-conducting water channel. And because the two main classes of aquaporins occur in *E. coli*—which means they're exposed to the same cellular environment—and were both expressed recombinantly, the opportunities for comparative structural and functional analyses, combined with site-directed mutagenesis, promise to provide valuable insights into the molecular underpinnings of the selectivity of functionally different aquaporins.

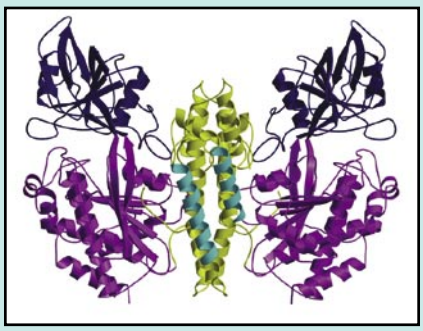
After fabricating and growing a recombinant form of AqpZ in *E. coli*, David Savage in the Stroud group recovered the proteins from the bacterial colonies, then purified and concentrated them. The proteins were crystallized—capturing five water molecules inside—and then analyzed by state-of-the-art high-resolution X-ray diffraction techniques. The architecture of aquaporin Z, the researchers report, is typical of aquaporins, with a spiral of eight oxygens providing water-binding sites inside the channel and amino acid side chains determining the size and chemistry of the channel. The outer membrane and cytoplasmic ends of the channel are wider than the interior, which is long and narrow. This structure confirms that aquaporin selectivity arises in part from erecting a physical barrier: small molecules, like water, can easily pass, but larger ones simply can't fit. And the strategic positioning of amino acid residues with hydrophilic or hydrophobic properties along the channel helps police the influx of molecules based on their affinity for water. While it seems two amino acid chains located in the middle of the channel also provide a water-friendly surface, Stroud et al. say they play a more intriguing role. Noting that the water molecules occupy the channel in single file, the scientists explain that such an orientation would normally facilitate the random flow of protons (or hydrogen ions), which would be lethal to the cell. This central amino acid pair, they say, restricts the behavior of water molecules in the center of the channel in such a way that prevents "proton jumping" yet keeps the water flowing.

With two structural models of aquaporins down to the atomic level in the same species, scientists can now begin to investigate the molecular mechanisms that facilitate their selectivity. The importance of understanding these widely distributed channel proteins was underscored by the Nobel awards this year. Water transport is fundamental to life, and aquaporins are found throughout the body. Knowledge of their structure will help reveal the molecular mechanics of their specialized feats and promise to offer insights into a wide range of human disorders.

Savage DF, Egea PF, Robles Colmenares Y, O'Connell JD III, Stroud RM (2003) Architecture and selectivity in aquaporins 2.5 Å x-ray structure of aquaporin Z. DOI: 10.1371/journal/pbio.0000072

Structural Mechanism Shows How Transferrin Receptor Binds Multiple Ligands and Sheds Light on a Hereditary Iron Disease

Iron is an essential nutrient for sustaining life-forms as diverse as plankton and humans. But too much iron, or too little, can spell trouble. Mammalian cells maintain the proper balance partly with the help of a specialized cell surface protein called the transferrin receptor (TfR). TfRs bind to the iron-carrying transferrin protein (Fe-Tf) and escort their cargo to the cell's interior. (To learn more about iron metabolism, see the primer by Tracey A. Rouault in this issue [DOI: 10.3171/journal/pbio.0000079].) This receptor also binds the hereditary hemochromatosis protein (HFE), which is mutated in individuals who have the common iron-overload disorder hereditary hemochromatosis. While the molecular pathway that mediates cellular intake of iron through the TfR is known, it was not clear just how TfR assists in iron release to the cell and how it binds HFE and transferrin. By introducing multiple mutations in human TfRs, Pamela Bjorkman and colleagues identified functional binding sites for transferrin in both its iron-loaded and iron-free (apo-Tf) forms and for HFE. From these data, the researchers mapped out a scenario of the dynamic interactions between receptor and ligands (the bound molecule) and worked out a structure-based model for the mechanism of TfR-assisted iron release from Fe-Tf.



Ribbon diagram of transferrin receptor homodimer

Bjorkman's lab, which had previously solved the structures of both HFE and HFE bound to the TfR, used their structural information to investigate how the proteins interact, which amino acid residues are required for binding, whether the two ligands bind differently to the receptor, and how HFE binding affects transferrin binding. They found that Fe-Tf and HFE occupy the same or an overlapping site on the receptor, but since transferrin is much larger than the HFE protein, it appeared that transferrin could also interact with other parts of TfR. And it remained to be seen whether TfR discriminated between the iron-loaded and iron-free states of transferrin.

In this study, Bjorkman and colleagues expanded their library of TfR mutants to clarify the transferrin binding signature on TfR and to see how the TfR mutations affect the way apo-Tf and Fe-Tf interact with the receptor. They characterized the binding affinities of 30 TfR mutants to HFE and Fe-Tf and to apo-Tf, and they report that mutations in 11 of the TfR residues interfere with either one or both forms of transferrin. Four of these residues are essential for transferrin binding and are conserved in all known TfR DNA sequences. Since residues that didn't have much impact are not conserved, the scientists say the results are likely to describe transferrin–TfR interactions for other species as well.

As expected, the most critical residues required for transferrin binding fall within the receptor's helical domain and have significant physical overlap with residues required for HFE binding; though some residues that are required for apo-Tf binding do not affect Fe-Tf binding. Bjorkman et al. also identify additional residues in another domain on TfR (called the protease-like domain) that support Fe-Tf but not apo-Tf binding, confirming that the receptor binding footprints of the two metal-binding states of transferrin are indeed different. With a structural model showing *where* Fe-Tf and apo-Tf bind to the receptor, they could evaluate *how* they bind and thus explain how the receptor mediates iron release.

By suggesting a mechanism through which TfR binding regulates iron release, this structural model of the transferrin–TfR complex will bolster efforts to elucidate the molecular details of this process. Confirmation that transferrin and HFE do indeed compete for docking privileges reveals a possible role for HFE in maintaining iron homeostasis and will provide valuable insights into the dysregulation that leads to the warehousing of iron and resulting tissue and organ damage associated with hemochromatosis.

Giannetti AM, Snow PM, Zak O, Björkman PJ (2003) Mechanism for multiple ligand recognition by the human transferrin receptor. DOI: 10.1371/journal.pbio.0000051

Chromosome Locus and Candidate Gene for Osteoporosis Identified

While osteoporosis is commonly thought of as a disease of old women, it's really more a disease of old age. Marked by a deterioration of bone density and strength, osteoporosis (meaning "porous bones") often goes undetected until a fateful slip results in a serious fracture of the hip, spine, or wrist. While it's true that postmenopausal women are much more susceptible than men—largely owing to their smaller frames—risk for men increases exponentially with age. Risk factors traditionally reflect lifestyle choices, including lack of exercise and poor diet, though genetics appears to be a major determinant of low bone mineral density (BMD), a characteristic feature of osteoporosis. Genetic factors also influence an individual's rate of bone loss, bone size, and likelihood of falling. In a large genetic study of osteoporotic Icelanders and their extended families, Unnur Styrkársdóttir and colleagues at deCODE Genetics in Reykjavik identified a candidate gene associated with a predisposition for osteoporosis.

Linking specific genes with complex diseases like osteoporosis is a tricky business. There are likely to be several genetic causes, and to find them researchers need large populations, abundant genetic markers, and extensive patient data. In addition to powerful genetic resources, researchers at deCODE can take advantage of a nationwide genealogical database of native Icelanders stretching back to the country's origins 1,100 years ago. By screening hundreds of affected individuals and their families, the scientists searched for candidate genes underlying osteoporosis and its harbinger, low BMD.

While low BMD is the best predictor of osteoporotic fracture, peak bone mass (mineral mass begins to diminish after young adulthood) and the rate of postmenopausal bone loss also appear to influence risk. Genetics contributes to all of these factors. Styrkársdóttir and colleagues screened the genealogical data on 207 extended families for the most clinically relevant phenotypic factors—low BMD and osteoporotic fractures—to focus their search. Screening only affected families with these attributes in a genome-wide scan, they say, seemed a reasonable method for finding susceptibility genes.

They conducted a series of scans and found a significant linkage to the short arm of Chromosome 20. This region contains six known genes, including four genes involved in bone formation and osteoblast (bone-forming cell) differentiation.

To winnow the list of most likely candidate genes, Styrkársdóttir et al. screened the genomes of 705 individuals with osteoporosis in a case-control study, using closely spaced genetic markers within the region of interest. This analysis pointed to *BMP2* as the most likely candidate—an enticing finding, because *BMP2* (bone morphogenetic protein 2) is known to be involved in bone development. Sequencing the *BMP2* gene in 188 patients and 94 controls to look for functional variants that might account for a predisposition to osteoporosis flagged several possible sequence alterations (which were subsequently shown to be associated with osteoporosis in the larger cohort). Furthermore, an independent replication study with two groups of postmenopausal Danish women—one group with low BMD and one with osteoporotic fractures—found comparable results, with a higher incidence of the *BMP2* variants in the affected women compared to the controls. These variants alone can't explain all of the results, however, because a linkage analysis run with patients who do not carry these variants still shows a likely association in the *BMP2* region, implying that there may be other variants in *BMP2* or adjacent genes that might be influencing osteoporosis.

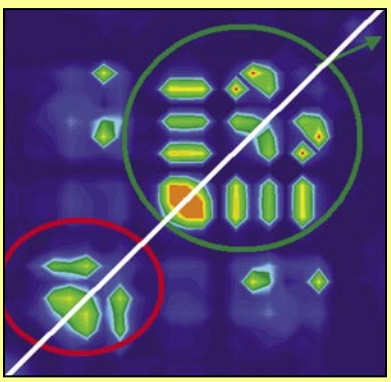
These results support *BMP2* as a likely gene for osteoporosis. Future studies will determine whether it increases risk through its control over peak bone mass, as the researchers suggest, or through some other mechanism. Styrkársdóttir et al. stress that other as-yet-unidentified variants within or near *BMP2* may also account for the linkage they found. The researchers hope their work will inspire others to replicate their results, to confirm and elucidate the role of this gene in osteoporosis. Understanding the mechanisms and signaling pathways of *BMP2*'s effects could not only identify drug targets for osteoporosis therapies, but also catch those at high risk before they take that fateful fall.

Styrkársdóttir U, Cazier J-B, Kong A, Rolfsson O, Larsen H (2003). Linkage of osteoporosis to chromosome 20p12 and association to *BMP2*. DOI: 10.1371/journal/pbio.0000069

GAD2 Identified as Candidate Gene for Obesity

Like many complex physiological conditions, obesity results from the interaction of multiple genes and environmental factors. Obesity often runs in families, but since families tend to share lifestyle habits as well as genes, teasing out genetic factors has proved difficult. The search for candidate genes for susceptibility has been further complicated by the likelihood that each gene contributes a small effect and that these genes may differ in different populations. In a large case-control study of nuclear families in France, Philippe Froguel and colleagues used genome-wide scans of a chromosomal region linked with susceptibility to identify a new candidate gene for obesity.

The genetic variation that accounts for the near boundless diversity in human form and behavior arises from an average 0.1% difference in genome sequence among individuals. Much of this variation is due to single nucleotide polymorphisms, or SNPs, which occur on average at one locus per 1,000 DNA basepairs, out of a total 3 billion. Many of these variants occur outside of genes, but those found in or around genes affect everything from individual appearance to susceptibility (or resistance) to disease. Specific gene variants, called alleles, are linked to increased risk for a number of diseases, including "monogenic" obesity, a rare condition in which single genes can cause obesity. But most cases of obesity are "polygenic," influenced by many genes.



Visualization of linkage disequilibrium

Froguel and colleagues had previously found a strong linkage between a region of chromosome 10 and obesity, with SNPs in the *GAD2* gene turning up most frequently. *GAD2* codes for an enzyme (glutamic acid decarboxylase, or GAD65) that catalyzes the production of a neurotransmitter (gamma-aminobutyric acid, or GABA) that interacts with a neuropeptide (NPY) in the hypothalamus to help stimulate appetite. In an average-weight individual, the metabolic effects of orexigenic neuropeptides (which stimulate the appetite, decrease energy use, and support fat storage) and those of anorexigenic neuropeptides (which suppress appetite, increase energy use, and help burn fat) are kept in balance. It has long been thought that alleles affecting metabolism may increase susceptibility by inhibiting anorexignic neuropeptides.

Screening genomic data on 575 obese subjects and 646 controls for *GAD2* alleles, Froguel et al. identified one group of alleles as "protective" against obesity and another as increasing risk. These links were supported by genetic screens of the nuclear families participating in the study, which showed a higher frequency of the protective group of alleles in the controls. Obese individuals with two copies of one at-risk allele also had more difficulty controlling food intake compared to control subjects, based on a standardized assessment of eating behavior. When the researchers investigated the physiological effect of the "highest risk" *GAD2* variant in a mouse cell line, they found it caused a 6-fold increase in transcriptional activity compared to the "wild-type" cells without the variant.

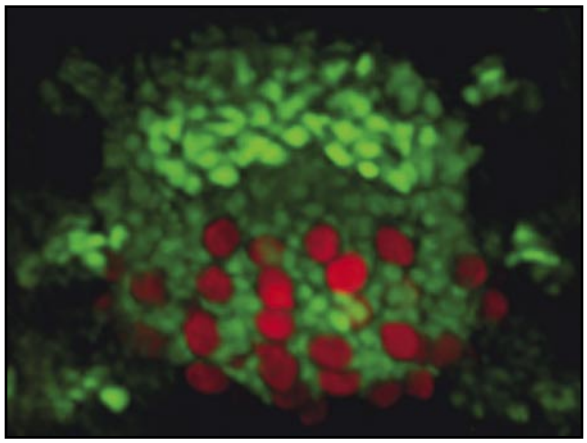
The researchers hypothesize that an overexpression of the *GAD2* gene may increase the amount of the neurotransmitter GABA in the hyopothalamus, thereby increasing GABA's orexigenic effects and leading to overeating. Studies in turkeys injected with muscimol, a GABA activator, and mice overexpressing a protein that regulates GABA localization support this notion; the more musimol the turkeys received, the more they ate, while the mice gained more weight and showed higher fat deposits.

Genetic factors can't explain the rapid rise in obesity rates, but they can provide clues to preventive and therapeutic approaches to ease the health burden associated with the condition. The results to date suggest that *GAD2* is a promising candidate gene. Future work on the genetics and biology of *GAD2* will resolve whether and how the gene influences body weight.

Boutin P, Dina C, Vasseur F, Dubois S, Corset L, et al. (2003) *GAD2* on chromosome 10p12 is a candidate gene for human obesity. DOI: 10.1371/journal.pbio.0000068

Novel Receptor Guides Germ Cell Transepithelial Migration in *Drosophila*

From the moment a fertilized egg starts dividing, its cellular progeny fall into a highly coordinated regimen of motion, growth, and change. As one cell quickly becomes hundreds, some cells passively ride the waves of dividing cells, while others must navigate uncharted territory to reach their destination. Such is the fate of germ cells, the guardians of genetic inheritance. In the fruitfly *Drosophila*, primordial germ cells are among the first cells to develop, appearing in the posterior embryo. Yet the cells that form the somatic tissue of the gonad—where germ cells mature into sperm and eggs—arise in the middle layer of cells, called the mesoderm. Primordial germ cells must make their way through the densely packed cells that form the posterior midgut epithelium and then negotiate a series of topographical landmarks before arriving at the gonadal mesoderm cells.



***Drosophila* germ cells migrating through the gut epithelium**

The genes and molecules guiding germ cells' improbable migration through the posterior midgut are not well known. Scientists have only recently discovered that a chemokine receptor called CXCR4—a member of the G protein-coupled receptors (GPCRs), the largest family of cell-surface receptors—controls the migratory pattern of a variety of cell types, including leukocytes, neurons, and cancer cells. Subsequent studies in zebrafish and mice showed that this receptor and its ligand (bound molecule) guide germ cell migration in vertebrates. Now Ruth Lehmann and colleagues have identified a novel type of GPCR gene called *tre1* that is required for the earliest stage of active germ cell migration—passage through the posterior midgut epithelium.

Until now, no mutations had been identified that affected transepithelial migration without disrupting the midgut itself. To identify likely candidates, Lehmann et al. introduced a series of mutations in *Drosophila*, ranging from overexpression to no expression, and stained the germ cells to observe the mutations' effects on germ cell migration. A laborious series of experiments led the authors to a gene that had a very interesting germ-cell phenotype (the physical effect of the mutation) when misexpressed in these cells. After closer analysis, they discovered that the gene was not normally expressed in germ cells and that loss of gene function had no effect on these cells. But the authors realized they were on the right track since the protein sequence of the gene was a GPCR and another closely related gene produced a striking germ-cell phenotype: in some cases, a gonad wound up with only one out of the typical 12–15 germ cells, while the other cells remained trapped in the gut.

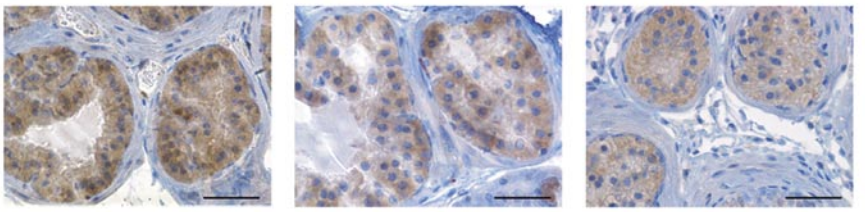
This gene, called *tre1*, had previously been considered a taste receptor gene until the true taste receptor was discovered. Expression pattern studies of *tre1* showed that it is expressed in germ cells and that germ cells lacking the Tre1 GPCR display a significant defect in migration. Maternal genes (gene products from the mother that wind up in the embryo) regulate the earliest stages of development, and the researchers conclude that normal germ-cell migration through the posterior midgut—the first active migratory step—depends on maternal *tre1* RNA and protein deposited in the germ cells. And *tre1* may be needed only for this early step, since the few *tre1*-mutant cells that do pass through the epithelium still make it to the gonads. Given the fact that *tre-1* already had a name, Lehmann et al. decided to keep its abbreviation but, following genetics tradition, change its meaning to describe the phenotype of its mutation: trapped in endoderm. While it is unclear what ligand activates the Tre1 receptor, the scientists identified a likely downstream target of the receptor, called Rho1. When this well-known GPCR signaling component is disrupted, germ cells cannot pass through the posterior midgut epithelium.

Tre1, which belongs to a new subclass of GPCRs, has three *Drosophila* homologs that are expressed in various types of migrating cells. Because GPCRs are evolutionarily ancient proteins, Tre1 and its homologs may have conserved functions in directional cell migration. Lehmann et al. propose that Tre1 functions similarly to related chemokine receptors active during transepithelial migration during an immune response—a process that is not well understood—and that this new group of GPCRs could well control this process, as well as a variety of other migratory functions. If it turns out that these same molecules are also active in cancer—in which invasive migratory behavior leads to tumor cell metastasis—scientists will have a promising lead on finding ways to block their action in targeted therapies. Lehmann et al. hope these results will lay the foundation for such efforts by shedding light on the molecular mechanisms that drive cell migration through tissue.

Kunwar PS, Starz-Gaiano M, Bainton RJ, Heberlein U, Lehmann R (2003). Tre1, a G protein-coupled receptor, directs transepithelial migration of *Drosophila* germ cells. DOI: 10.1371/journal/pbio.0000080

Dose of *PTEN* Gene Drives Progression of Prostate Cancer

Each year some 29,000 American men die of prostate cancer. With about 221,000 more diagnoses expected in the United States in 2003, the disease trails only skin cancer as the most common male cancer. As a complex disease, multiple genetic and environmental factors contribute to risk and, like most cancers, its onset and progression depends on the combination of a series of genetic disruptions rather than on a single event. But as Pier Paolo Pandolfi and colleagues report, protein "dose"—that is, the level of remaining activity—also influences cancer progression.



Hyperplasia in *Pten* hypomorphic mice

Focusing on the tumor suppressor gene *PTEN*, the researchers used the mouse as a model system to study tumor progression in prostate cancer. *PTEN* is among the most commonly mutated tumor suppressor genes in human cancer. And like many other tumor suppressors, *PTEN* targets proteins in signaling pathways that regulate cell growth and apoptosis in healthy tissue and contributes to cancer when dysfunctional. Humans, as diploid organisms, generally have two versions of most genes, including *PTEN*. In the event that one copy is damaged or lost, gene function is usually maintained by the other copy. In the classic definition of a tumor suppressor, *both* copies must be lost for a tumor to occur. Yet in many cases of advanced cancer, including prostate cancer, only one copy is lost at the time a patient shows symptoms. It is then not unreasonable to hypothesize that the degree of remaining *PTEN* activity controls the course of the disease: loss of one copy could influence tumor initiation, while further slight reductions might be sufficient to facilitate the invasion and metastatic behavior of late-stage cancers.

Pandolfi and colleagues chose two strategies to investigate this hypothesis. In the first approach, they genetically engineered one series of mice with minimal levels of murine Pten protein (complete loss results in embryo death). This novel "hypomorphic" strain of mice exhibits only 25%–35% active *Pten*, which appears to be the minimum level of Pten needed to survive embryonic development. This hypomorphic model adds to existing strains of fully functional and 50% active *Pten* mice. In order to model the full loss of Pten protein, the researchers generated another series of mice in which *Pten* genes were selectively disabled in the prostate only. The researchers found that subtle reductions in *Pten* dose did indeed produce progressive changes in the biology of the tumor, while mice having no functional *Pten* genes showed the most invasive and aggressive cancers. These results, the researchers say, show that *Pten* plays a "crucial dose-dependent role in prostate cancer tumor suppression" and that progressive reduction of gene function induces progressive changes in the quantity and quality of molecular and pathological effects on the pathway to full-blown cancer.

By coupling the molecular genetics and dose of Pten protein with the physiological progression of cancer in the prostate, these new mouse models may not only shed light on cancer progression in humans, but also help bolster diagnostic, prognostic, and therapeutic techniques. While evaluation of tumor status has traditionally been determined by pathological analysis of tissue samples, these new models allow scientists to pair anatomical stages with underlying molecular events—such as the expression level of a single gene or protein—to allow more accurate assessments. These molecular profiles can also help researchers design targeted, more efficient prostate cancer treatments. For example, if prostate tissue is hypersensitive to reduction of PTEN in humans—which the results suggest may be the case, since male mice with only 30% of normal Pten levels show massive and selective enlargement of the prostate, and even invasive tumors—then ongoing monitoring of PTEN levels could help tailor therapies based on promoting *PTEN* expression. For patients with complete loss of *PTEN* function, where this would not be an option, inhibiting the proteins made overactive through *PTEN* loss could prove effective. And these approaches could well hold true for other cancers involving *PTEN* mutations, including endometrial, brain, and breast cancer.

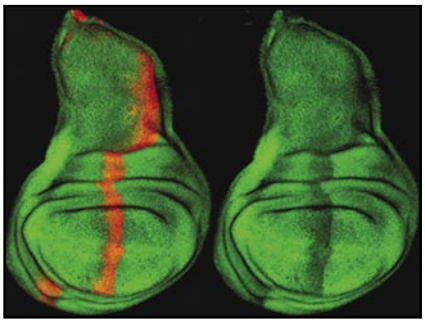
Trotman LC, Niki M, Dotan ZA, Koutcher JA, Di Cristofano A, et al. (2003) Pten dose dictates cancer progression in the prostate. DOI: 10.1371/journal.pbio.0000059

New Screen Identifies Elusive MicroRNA Targets

Biology has come a long way since the days of DNA makes RNA makes protein. This "central dogma of molecular biology" does not tell the whole story, but then it was not meant to. Francis Crick proposed the term in 1958 to describe the unidirectional transfer of information from a genetic alphabet to an amino acid alphabet and saw it as a useful model for guiding investigations of protein synthesis, at a time when little was known about molecular biology. Twelve years later, the discovery of reverse transcription—the transfer of genetic information from RNA to DNA, which is how retroviruses invade their hosts—was an unexpected change in the paradigm. But today the study of RNA is a burgeoning field, growing along with new discoveries of the molecule's diversity of form and function.

Among the most recently recognized RNAs are the "small RNAs." They do not code for proteins, but regulate how and whether coding RNA, called messenger RNA (mRNA), functions. Unknown before 1993, the 21- to 23-nucleotide transcripts known as microRNAs (miRNAs) are thought to regulate gene expression by binding to mRNAs and inducing their cleavage or blocking their translation into proteins. Hundreds of miRNAs have been identified in plants, worms, flies, and humans, but function has been attributed to only four animal miRNAs. Now Alexander Stark, Julius Brennecke, Robert Russell, and Stephen Cohen have developed a computational model to predict miRNA targets in *Drosophila*, the first step in understanding the functions of miRNAs.

Progress in identifying the targets and associated functions of animal miRNAs has been slow, partly because traditional genetics methods—which investigate gene function by inducing mutations and observing what happens when a gene is disabled—are less efficient since the genes that code for miRNAs are small. The small size of the miRNA sequence also hampers computer-based predictions of targets. Further complicating matters, miRNA/mRNA pairings in animals tend to be "imperfect"—the miRNA sequence and the mRNA sequence are not entirely complementary—making them harder to detect.



Presence of microRNA (red) inhibits expression of the target (green)

Stark, Brennecke, et al. have overcome these problems by combining a computational genome screen customized to identify potential miRNA targets with an RNA-folding program that evaluates the structural and thermodynamic plausibility of the predicted pairs and distinguishes real from random matches. The functional sites for the few validated miRNA/mRNA pairs are all located in the 3' untranslated region (UTR) of mRNA and are conserved in the same region of homologous genes from related species. (Many regulatory elements are thought to reside in noncoding regions of the genome. Two such elements, the 3' UTR and 5' UTR of mRNAs, are known to harbor sequences essential for gene expression and regulation.) To create a database of such conserved UTRs, Stark, Brennecke et al. compared the UTRs of functionally equivalent genes in the fruitfly species Drosophila melanogaster and Drosophila pseudoobscura. Then they searched these conserved genetic elements for potential target sequences by requiring partial complementarity to the known miRNAs. Going a step further, the researchers also searched the genome of the mosquito Anopheles gambiae—which diverged from Drosophila some 250 million years ago—to see whether the predicted targets were conserved there as well. They found that requiring conservation in Anopheles was more useful in confirming predicted Drosophila targets rather than identifying them.

Efforts to find more animal miRNA targets should get easier, the researchers suggest, as more is learned about the fly and mosquito genomes and as more is understood about the structural and biochemical nature of miRNA/mRNA pairing—for example, which regions and structures are essential for pairings. Until then, the targets predicted in this new screen will help scientists validate more miRNA targets and ultimately reveal just how large a role these small RNAs play in regulating the genome.

Stark A, Brennecke J, Russell RB, Cohen SM (2003) Identification of *Drosophila* microRNA targets. DOI: 10.1371/journal.pbio.0000060

Gene Expression Profile Created for Mouse Stem Cells and Developing Embryo

While the controversy surrounding the ethics of stem cell research shows no signs of abating, scientists continue to demonstrate the promise of stem cell–derived therapies for a wide range of degenerative diseases. The hope is that stem cells, which retain a unique "pluripotent" ability to morph into any of the 200 cell types of the human body, could be used to repair or replace damaged or diseased tissue. Many animal studies have supported this potential for both embryonic and adult stem cells, with some findings indicating that hematopoietic (blood-forming) stem cells could be cultured and used to help cancer patients who need bone-marrow transplants, and others suggesting that adult brain stem cells could repair damaged nerve tissue and help paralysis patients recover movement.

Mus musculus

Despite these advances, little is known about the molecular events that trigger differentiation and determine a cell's developmental potential. Such information will help scientists better manipulate this potential in stem-cell therapies. Having compiled a comprehensive database of genes expressed in mouse early embryos and stem cells, Minoru Ko and colleagues present a model that takes the first step toward characterizing the molecular profile of stem cells.

Arguing that a broad understanding of these molecular determinants requires a broad selection of cell types, the scientists combined new gene expression data on early embryos and stem cells with existing gene expression data to compare transcription patterns across a wide range of cell types and developmental stages. The expanded mouse transcriptome (record of transcribed genes) included data on unfertilized eggs; "totipotent" fertilized eggs, which have the potential to become any cell; pluripotent embryonic cells; various embryonic and adult stem cells; and fully differentiated cells.

Ko et al. characterized gene activity in this diverse cross-section of cells and looked for molecular differences, including the level and type of gene expression, as a measure of developmental potential. Applying standard statistical tools to spot major trends and clusters in gene expression, the researchers found 1,000 new gene candidates, which they grouped according to particular embryonic stage and stem cell type. From these signature gene sets, they identified a cluster of 88 genes whose average expression level decreased as cells became more specialized. This finding indicates not only that totipotent and pluripotent cells have distinct genetic profiles, but that these 88 genes may serve as molecular markers of developmental potential. Further support of this predictive power comes from the finding that adult stem cells—cells derived from adult organs that retain a measure of pluripotency—were clustered with early embryonic tissues of similar potential.

These results are consistent with previous findings that cells gradually lose developmental potential and that adult stem cells retain plasticity, but more importantly they link signature genes with different stem cell types and stages—thus providing a preliminary set of molecular markers for characterizing the function and potential of different stem cells. Identifying the genes that shape the unique properties of stem cells will shed light on the molecular pathways that guide development and suggest ways to best exploit the full therapeutic potential of these embattled cells.

Sharov AA, Piao Y, Matoba R, Dudekula DB, Qian Y, et al. (2003) Transcriptome analysis of mouse stem cells and early embryos. DOI: 10.1371/journal/pbio.0000074

Distinct Mechanisms Control Transposon Inheritance through Overlapping Pathways

When Barbara McClintock discovered transposable genetic elements in maize in the early 1950s at Cold Spring Harbor Laboratory (CSHL) in New York, she dubbed them "controlling elements," since their manner of jumping from one chromosomal site to another controlled their own function as well as that of nearby genes. Just as genetic material is transferred from "parent" to "daughter" cell, these changes in "transposon" activity, and corresponding changes in gene function, are inherited by cellular progeny. The significance of McClintock's jumping genes was not widely recognized at first, but transposable elements are now considered a classical model for epigenetic inheritance—the study of heritable changes in gene expression and regulation that arise independently of changes in DNA sequence.

Transposons are normally "silent"—that is, inactive and stationary—but various mechanisms can rouse them and thus influence their regulation of gene expression. They can be inherited in this active state. Because hyperactive transposons would cause genetic chaos—cancer and other diseases arise from misplaced genes—epigenetic controls are crucial. A process called DNA methylation is thought to keep transposons from running amok, but other mechanisms that affect epigenetic inheritance may also play a role. While the molecular basis of these regulatory pathways is not clear, it is apparent that they interact. Investigating the nature and consequences of these interactions, Rob Martienssen and his colleagues at CSHL found that different transposons respond to different types of epigenetic regulation and identified two distinct mechanisms of transposon silencing that likely interact in a common pathway.

The challenge of teasing out the individual contributions of these highly complex, overlapping regulatory pathways is complicated by the fact that few model organisms suit the task. Fission yeast, the organism of choice for many fundamental cell biology investigations, seems to lack DNA methylation, while mice can't live without it, ruling them out as viable subjects for experiments that disturb this form of DNA modification. The *Arabidopsis thaliana* plant, on the other hand, not only can survive such disturbances but can also produce offspring ripe for studying epigenetic inheritance patterns. Plus, this plant has already helped researchers identify many genes involved in epigenetic regulation.

Jumping genes discovered in maize through their effect on kernel pigmentation

To study the molecular basis of epigenetic inheritance and the interaction of the various regulatory networks—DNA methylation, histone modification, and RNA interference—Martienssen et al. used several *Arabidopsis* strains with mutations known to affect these processes as well as the epigenetic inheritance of active transposons. DNA methylation chemically modifies DNA; histone proteins are modified by other molecules and alter chromatin structure (the complex of DNA and proteins that packages DNA in cells); and RNA interference (RNAi) blocks DNA transcription. Focusing on a representative group of transposons, Martienssen et al. crossed mutants in DNA methylation, histone modification, and chromatin remodeling with nonmutant plants and characterized the impacts on each transposon.

Using transposon expression (transposons must be expressed to jump around) in the offspring to determine whether transposable elements were silent, transiently activated, or heritably activated, the researchers found that when each of the transposons were subjected to the same mutations, they did not all respond in the same way. While some mutations affected all of the elements, other mutations affected only a subset. The quality of the responses also varied; some mutations caused changes that were transient, that is, lost in the next generation, while others were inherited.

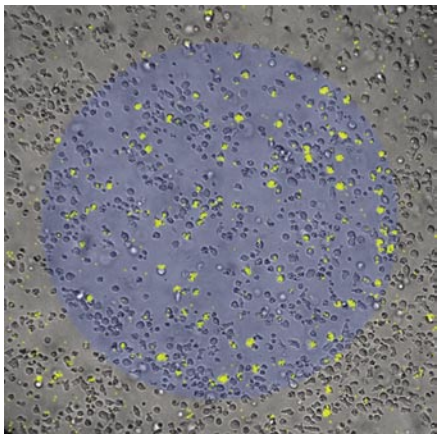
They also found that RNAi, which silences transposable elements in fruitflies (*Drosophila*) and worms (*Caenorhabditis elegans*), influences epigenetic inheritance. Small interfering RNA (siRNA), however, somehow interacts with other pathways, such as DNA methylation and chromatin remodeling, to do so.

Martienssen et al. say these results indicate that transposons differ in their pattern of regulation and tend to respond to different types of epigenetic regulation, suggesting there are distinct mechanisms of transposon silencing but that the mechanisms or pathways probably interact. Given the potentially damaging effects of transposition and its seemingly ubiquitous presence in living things, it's reasonable to wonder whether transposable elements evolved as a flexible response to special circumstances, like environmental stress, a possibility proposed by McClintock. The epigenetic mechanisms that control these elements are vital in regulating the structure and organization of the genome and in establishing the right balance between genetic variation and fidelity, the molecular foundation of evolution. It seems only fitting that biologists at Cold Spring Harbor would continue to plumb the implications of McClintock's work—and that their model system of choice would be a plant.

Lippman Z, May B, Yordan C, Singer T, Martienssen R (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. DOI: 10.1371/journal/pbio.0000067

New Microarray Detects Antigen-Specific T Cells and Immune Responses

When a pathogen slips into your body, its chances of escaping detection are slim, thanks to the highly specialized surveillance team of the immune system. If a virus, for example, succeeds in duping a cell's molecular machinery into manufacturing new copies of itself, the cell responds by breaking a few of the pathogen's proteins into fragments and displaying some of these peptides (or antigens) on its surface. There, held in the grips of cell surface proteins called MHC (major histocompatibility complex) proteins, antigens can be recognized as foreign to circulating T lymphocytes. The body produces billions of these white blood cells every day, each outfitted with specialized T cell receptors and MHC proteins outfitted to recognize a unique antigen. This exquisitely specific recognition of antigens—displayed on the cell surface by an MHC protein—is the critical step leading to the proliferation, activation, and differentiation of T cell clones specially equipped to destroy that pathogen.



Calcium flux in activated lymphocytes

The detection of "antigen-specific" T cell populations can provide insight into the physiology of the immune system and how it responds appropriately to disease or inappropriately to host proteins in autoimmune disorders. Scientists have linked different T cell responses to specific antigens associated with microbes, autoimmune diseases, allergens, and cancer cells. Understanding how the immune system responds to such pathogens—and how it functions in the absence of disease—depends on being able to detect and evaluate these responses. But since different populations of T cells can interact with diverse antigens simultaneously, identifying and characterizing these populations has presented a huge challenge. Yoav Soen, Daniel Chen, Daniel Kraft, Mark Davis, and Pat Brown have developed a high-volume screen that not only sorts and identifies multiple antigen-specific T cell populations from a diverse sample, but also determines which populations are active, even when the response is weak.

Faced with the challenge of identifying huge numbers of antigen-specific cells with a wide range of peptide-MHC complexes, Soen et al. turned to the high-throughput technology of microarrays. But instead of using bits of DNA as probes to latch on to active genes in cell samples, the researchers used arrays of peptide-MHC complexes to capture antigen-specific T cells. They printed tiny spots of different peptide-MHC complexes on glass slides and then layered populations of T cells onto the slides, where the T cells could interact with each of the printed peptide-MHC complexes. The rare cells that recognize each specific peptide-MHC complex are captured at the corresponding spot on the microarray, where they can be counted and assayed. With this new microarray, immunology researchers—traditionally restricted to identifying and quantifying only a few lymphocyte populations at a time—can now characterize hundreds or thousands at a time.

To test the reliability of the microarray, the researchers labeled different populations of T lymphocytes based on their antigen-binding specificities and found that the array accurately detected each population. The ability to sort out and assay rare cells that recognize specific antigens will be useful for a wide range of applications, including vaccine development. The researchers also demonstrated the array's sensitivity by successfully detecting a weak, specific immune response in cells extracted from vaccinated mice. Such an application would be a valuable tool for monitoring the global population of T cells in a living organism—including human patients—in response to vaccination, infection, autoimmunity, and other diseases.

Soen Y, Chen DS, Kraft DL, Davis MM, Brown PO (2003) Detection and characterization of cellular immune responses using peptide-MHC microarrays. DOI: 10.1371/journal/pbio.0000065

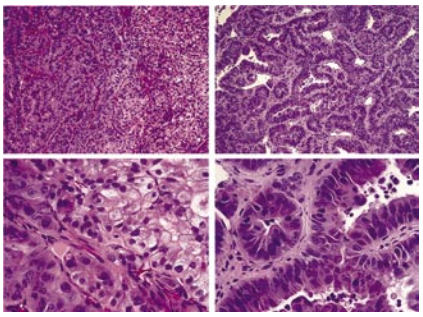
Inhibition of HIF2α Protein Suppresses pVHL-Driven Tumor Growth

Scientific progress in biology typically relies on incremental advances that explain a small piece of an impossibly complex puzzle. But sometimes, understanding that one piece is enough. Scientists have long known that a rare cancer syndrome, called von Hippel–Lindau (VHL), had a hereditary component. First seen clustering in families over a century ago, the disease is characterized by vascular tumors, including retinal and central nervous system "hemangioblastomas" and clear-cell carcinomas of the kidney and adrenal glands. Just ten years ago, scientists identified the tumor suppressor gene, called *VHL*, that causes the disease and later discovered how loss of *VHL* leads to cancer.

It had appeared this pathway was uncommonly simple. *VHL* regulates HIF (hypoxia-inducible factor), a transcription factor that controls cellular responses to below-normal oxygen levels in the blood or tissue, a condition known as hypoxia. In the presence of oxygen, the *VHL* gene product pVHL forms a complex with other proteins and targets HIF for degradation. In response to hypoxia, HIF activity increases and induces the expression of a number of proteins that help cells adapt to low levels of oxygen. If *VHL* function is lost, it cannot target HIF, leading to a stable accumulation of HIF and a corresponding overproduction of the hypoxia-induced proteins. When aberrantly produced, these proteins contribute to tumor formation and growth through angiogenesis (the mass of blood vessels characteristic of hemangioblastomas) and unregulated cell growth. But subsequent studies suggested that *VHL* might have other targets, raising the possibility that HIF might not be the sole, or even primary, mediator of tumor formation in cells lacking pVHL. Now William Kaelin and colleagues report that inhibiting HIF is not only necessary to support pVHL's function as a tumor suppressor, but that it is also enough.

The researchers set out to determine whether the inhibition of HIF was specifically responsible for pVHL's reported tumor suppressive effects. When pVHL is reintroduced into human

renal carcinoma cells lacking pVHL and overexpressing HIF2α, HIF2α levels go down, hypoxia-inducible gene expression goes down, and tumorigenesis is suppressed. Kaelin et al. found that introducing HIF2α variants lacking specific pVHL-binding sites into such cells was sufficient to negate pVHL's tumor suppressor activity, implying that inhibition of HIF is necessary for pVHL to suppress tumor growth. To see whether simply removing HIF2α would mimic the tumor suppressor activity of pVHL, the researchers introduced "short hairpin" RNAs (shRNAs) to "silence" HIF2α protein production. The shRNA-treated cells (which lacked functional *VHL* genes) showed decreased levels of HIF2α activity and were grossly impaired with respect to facilitating tumor formation in mice. These results suggest that, while pVHL may have multiple functions, inhibition of HIF can account for its ability to suppress tumor growth, at least in the context of renal carcinoma.



Tumor suppression by HIF2α shRNA

Tumor suppressor genes are defined by what happens in their absence—while their normal function is not always apparent, it is clear that their loss of function results in cancer susceptibility—and by the fact that they usually require loss of both versions, or alleles, for damage to occur. While patients with VHL inherit one mutated copy and later acquire a mutation in the second copy, patients with similar but nonhereditary vascular cancers have also lost both *VHL* alleles. Knowing HIF's role as the fulcrum between health and disease, scientists can now concentrate on just that one piece of the puzzle—HIF and its target genes—to develop customized preventive and therapeutic treatments not only for VHL patients but for those living with nonhereditary tumors, such as kidney cancers, that frequently result from pVHL inactivation.

Kondo K, Kim W, Lechpammer M, Kaelin WG Jr (2003) Inhibition of HIF2α is sufficient to suppress pVHL-defective tumor growth. DOI: 10.1371/journal/pbio.0000083