Research Digest

Synopses of Research Articles

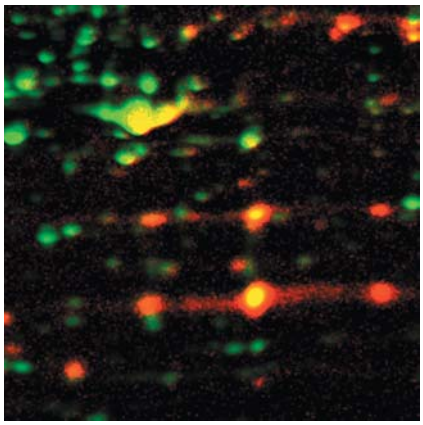
A New Way to Look at Oxidative Stress

DOI: 10.1371/journal.pbio.0020374

Chemical reactions lie at the heart of many biological processes, from photosynthesis and respiration to cell signaling and drug metabolism. Thanks to an atmosphere rich in oxygen, many organisms use oxygen to carry out these life processes. But oxygen metabolism produces highly toxic by-products called reactive oxygen species. When oxidation outpaces detoxifying reactions, oxidative stress occurs, and accumulating reactive oxygen species are free to wreak havoc on cellular machinery.

Cysteine, one of the 20 different amino acids that make up proteins, contains a thiol group, which can be modified upon oxidation. A thiol group can stabilize protein structures by forming covalent disulfide bonds and can mediate cysteine-regulated redox reactions. At the same time, however, the high reactivity of thiol groups makes them also particularly vulnerable to nonspecific reactions during conditions of oxidative stress.

Over the past few years, an increasing number of proteins have been discovered that use oxidative thiol chemistry to regulate their protein activity. In *PLoS Biology,* Lars Leichert and Ursula Jakob describe a novel method to monitor thiol modifications in proteins subjected to varying redox conditions in a living organism, the bacteria *Escherichia coli*. This technique is capable of providing a global snapshot of the redox state of protein cysteines during normal and oxidative stress conditions in the cell.



DOI: 10.1371/journal.pbio.0020374.g001

A differential thiol-trapping technique provides a snapshot of the in vivo thiol status of proteins upon variations in the redox homeostasis of cells

To detect proteins that have the ability to undergo stress-induced thiol modifications, Leichert and Jakob differentially labeled the thiol groups of thiol-modified and non-thiol-modified proteins. The proteins were then separated on two-dimensional gels based on their charge and molecular weight. If the technique worked, most thiol-modified proteins should be detected in the oxidizing environment of the *E. coli* periplasm (the region between the cell's membrane layers), and they were.

After proving the method's ability to detect proteins whose thiol groups were oxidized, the next logical step was to determine what proteins DsbA—the enzyme that catalyzes disulfide bond formation in the *E. coli* periplasm—was targeting. In *E. coli* mutant strains that lack DsbA, Leichert and Jakob identified a number of proteins with either substantially less or no thiol modification as compared to wild-type (non-mutant) strains, suggesting that these proteins are indeed DsbA substrates.

In contrast to the periplasm, the *E. coli* cytoplasm contains several reducing systems. When the researchers tested a mutant strain that lacked the gene for the reducing enzyme thioredoxin, they found that a large number of proteins accumulated in an oxidized state. Many of these proteins have cysteines and require a reduced thiol status for their activity. These results demonstrated that under normal growing conditions, many proteins contain cysteine residues that are vulnerable to even small amounts of reactive oxygen species and so require the constant attention of detoxifying enzymes.

In a final set of experiments, Leichert and Jakob discovered a number of proteins whose thiol groups get specifically modified in the presence of reactive oxygen species. These results start to explain some of the many metabolic changes that occur in oxidatively stressed cells.

Leichert and Jakob's technique should be applicable to many different cell types and organisms and can be used to investigate the in vivo thiol status of cellular proteins exposed to virtually any physiological or pathological condition that is accompanied by oxidative stress. The next step will be to investigate just how thiol modifications mediate the various functions of redox-regulated proteins.

Leichert LI, Jakob U (2004) Protein thiol modifications visualized in vivo. DOI: 10.1371/journal.pbio.0020333

Shut Down, Don't Stress Out

DOI: 10.1371/journal.pbio.0020373

Among the many stresses faced by a cell, one of the most serious is exposure to oxidizing agents. An invading organism, for example, must defend itself against the oxidative assault mounted by a host's immune system. Since oxidation can rapidly destroy many types of molecules, cells have developed multiple means of protecting against it. Rapid mobilization of these defenses requires diversion of resources and temporary suspension of many normal cellular functions, including protein synthesis. In a new study, Elise Hondorp and Rowena Matthews show that when the *Escherichia coli* bacterium confronts oxidative stress, an enzyme that stands at a central point in the amino acid supply line for protein synthesis is rapidly and reversibly inactivated.

Of the twenty amino acids that make up proteins, methionine plays a special role. It is the first amino acid added to every polypeptide chain, and without it, protein synthesis grinds quickly to a halt. Methionine is formed in *E. coli* through the action of the enzyme cobalamin-independent methionine synthase (MetE), which makes up between one and five percent of all protein in the cell. Thus, by turning off MetE in the face of oxidative stress, protein synthesis can be slowed or stopped, freeing cellular resources to be used elsewhere.

Hondorp and Matthews show that in *E. coli*, MetE is acutely vulnerable to oxidation under a variety of conditions. These results are in accord with a companion study by Leichert and Jakob, also in *PLoS Biology,* showing that MetE is one of the proteins most sensitive to oxidative damage. When the active site of MetE is stressed by an oxidant, Hondorp and Matthews show, it is temporarily blocked by the attachment of a glutathione subunit. Glutathione is a small molecule that includes a reactive sulfur atom. During "glutathionylation" of MetE, a sulfur on an amino acid of the enzyme is oxidized and links up with a sulfur on glutathione. This study shows that glutathionylation occurs only on a specific amino acid (cysteine 645), which recent structural work indicates sits at the entrance to the active site.

Attachment of the bulky glutathione subunit to this cysteine would be expected to block the active site, thus shutting down enzymatic activity. The results indicate that glutathionylation does indeed prevent activity of the enzyme, and furthermore, causes the enzyme to change its three-dimensional form. As the oxidative challenge abates, glutathionylation may be reversed, and the normal activity of the enzyme restored. Thus, glutathionylation of MetE may also serve to protect the active site from permanent oxidative damage. While glutathionylation is a common strategy in eukaryotes, MetE is so far one of the few proteins in bacteria known to be affected in this way.

Shutting down MetE and limiting methionine production may play another important role, namely, communicating the bacterium's metabolic state to other nearby *E. coli*. Methionine is a precursor for the signaling molecule AI-2, which is released extracellularly and appears to serve as a key indicator of colony health and density. This information enables neighboring cells to better respond to changing and potentially hostile environments. Thus, the glutathionylation and inactivation of MetE may provide a simple mechanism by which a bacterium and its neighbors attempt to deal with oxidative stress.

Hondorp ER, Matthews RG (2004) Oxidative stress inactivates cobalamin-independent methionine synthase (MetE) in *Escherichia coli*. DOI: 10.1371/journal.pbio.0020336

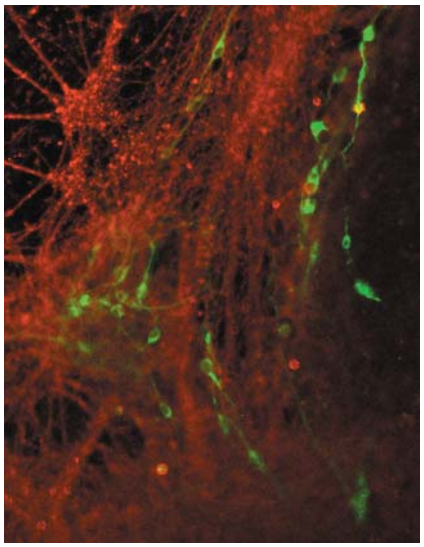
A New Cell Model for Parkinson's Disease

DOI: 10.1371/journal.pbio.0020385

Clinical descriptions of Parkinson's disease remain remarkably similar to those first described by James Parkinson nearly 200 years ago. Patients with "shaking palsy" experience a progressive loss of muscle control, increased muscle rigidity, inhibited movement, and tremors. These symptoms, it was later discovered, result from the loss of dopamine-producing neurons specifically in an area of the ventral midbrain called the substantia nigra. Midbrain dopamine neurons relay chemical signals that regulate motor control and less quantifiable attributes like mood and motivation, and therefore the loss of these cells is predicted to lead to the symptoms of Parkinson's.

Despite the well-characterized cellular basis of Parkinson's disease, the molecular mechanisms responsible for dopamine neurodegeneration remain unknown. There is evidence that both genetic and environmental components are involved. That a person with Parkinson's disease is three to four times more likely than an unaffected individual to have a close family member with "parkinsonian" symptoms suggests a genetic factor; furthermore, several genes have been associated with relatively rare, familial forms of the disease. For example, mutations of the protein alpha-synuclein (α-synuclein), which is found to aggregate in the brains of patients with Parkinson's, lead to a familiar parkinsonism syndrome. Mutations in a second gene called DJ-1 were recently found in two families with an inherited form of Parkinson's. Importantly, mutations in DJ-1 have previously been linked to the pesticide paraquat in unrelated research on cell stress and reactive oxygen species, and have been linked to dopamine neuron toxicity. Reactive oxygen species are molecular byproducts of oxygen metabolism that react with and damage cellular components like proteins and DNA, and there is evidence from postmortem studies that reactive oxygen species may play a role in Parkinson's disease.

Part of the challenge of untangling the relative contributions of all these components stems from the difficulty in finding a model that can adequately mimic the loss of dopamine cells. In two papers published in *PLoS Biology*, Asa Abeliovich and colleagues make the case that a model based on mouse embryonic stem cells offers a promising platform for dissecting the disease mechanism of Parkinson's. Working with these cells, the researchers report that DJ-1-deficient cells—and especially DJ-1-deficient dopamine neurons—display heightened sensitivity to oxidative stress. In a second paper, they link DJ-1 dysfunction to alpha-synuclein aggregation.



DOI: 10.1371/journal.pbio.0020385.g001

Oxidative stress has long been associated with neuronal cell death and neurodegenerative diseases like Parkinson's. Proving a causal relationship between oxidative stress and neurodegeneration, however, requires establishing a molecular mechanism.

In the first paper, to explore the hypothesis that DJ-1 contributes to the cellular response to oxidative stress, Abeliovich and colleagues created mouse embryonic stem cells lacking functional copies of DJ-1 and exposed them to hydrogen peroxide, a powerful oxidizer. Compared to normal cells, DJ-1 mutants showed signs of greater toxicity and higher levels of cell death. These defects were corrected when the researchers reintroduced the protein in the mutants, confirming DJ-1's responsibility for the defects. DJ-1 protects against oxidative damage, the results show, not by inhibiting the accumulation of the reactive oxygen species associated with hydrogen peroxide, but by mitigating the damage created by them.

Abeliovich and colleagues then explored DJ-1's function in dopamine neurons by inducing mutant and control embryonic stem cells to differentiate in cell cultures. Production of dopamine neurons was significantly reduced in the DJ-1-deficient cultures relative to the control cultures. And like DJ-1-deficient embryonic stem cells, DJ-1 dopamine mutants were vulnerable to oxidative stress. "DJ-1 deficiency," the authors conclude, "leads to reduced dopamine neuron survival and predisposes these cells to endogenous and exogenous insults." Inhibiting DJ-1 activity in neurons from the embryonic mouse midbrain produced the same results.

In the second paper, Abeliovich and colleagues go on to probe the molecular basis of DJ-1's activity. There have been several leads regarding how DJ-1 functions, based on homology to related genes, including a potential role as a molecular protein chaperone; protein chaperones assist in the folding and refolding of damaged proteins, and thus play a central role in the cellular response to oxidative stress. Abeliovich and colleagues found that DJ-1 acts as an unusual molecular chaperone that is specifically induced under oxidative conditions, and acts to prevent the aggregation of cellular proteins. Interestingly, the researchers go on to show that one substrate of DJ-1 activity is alpha-synuclein, thus providing a possible mechanism linking these two molecules implicated in Parkinson's disease. Altogether, these results support a link between toxin-induced oxidative damage and disease, and provide a tractable model for studying the molecular mechanisms of neurodegenerative disease.

Martinat C, Shendelman SB, Jonason A, Leete T, Beal MF, et al. (2004) Sensitivity to oxidative stress in DJ-1-deficient dopamine neurons: An ES-derived cell model of primary parkinsonism. DOI: 10.1371/journal.pbio.0020327

Shendelman S, Jonason A, Martinat C, Leete T, Abeliovich A (2004) DJ-1 is a redox-dependent molecular chaperone that inhibits α-synuclein aggregation formation. DOI: 10.1371/journal.pbio.0020362

Where to Start? Alternate Protein Translation Mechanism Creates Unanticipated Antigens

DOI: 10.1371/journal.pbio.0020397

In the spirit of good health, cells are constantly subjecting their protein contents to immunological surveillance by cytotoxic (killer) T cells. Tens of thousands of major histocompatibility complex (MHC) class I molecules cradle peptides (bits of proteins) on cell surfaces, and T cells detect any suspicious peptides with extreme sensitivity. If a cell is infected with a virus, peptides created from viral DNA will end up on the cell's surface as antigens, triggering immunological red flags.

Most—but not all—of the peptides presented by MHC class I molecules are created by conventional cellular mechanisms: with the help of a ribosome, three mRNA nucleotides (a codon) are decoded into a corresponding amino acid, which is strung as the next link on an elongating peptide. Most peptides begin with the amino acid methionine, coded by the mRNA nucleotide triplet A-U-G (AUG). But some peptides are "cryptic," arising from normally untranslated regions of mRNA or initiated with codons other than AUG.

Previous studies suggested that an unconventional translation mechanism creates some cryptic peptides. But how? And why? Only one type of translation initiator, transfer RNA (tRNA), specific for AUG and loaded with a methionine molecule, is known. Protein synthesis beginning at alternate codons has been attributed to imprecise pairing between the methionine translation initiator and mRNA. This, however, does not explain proteins that do *not* begin with methionine.

Only two mechanisms for building non-methionine-initiated peptides have been discovered. In a new study, Susan R. Schwab et al. characterize one of them, the CUG-initiated translation of a peptide starting with leucine instead of methionine.

The authors explored cellular translation by engineering cells to create peptides of interest and present them through matching MHC molecules on the cells' surfaces. Then, by harnessing the exquisite sensitivity of T cells to probe for antigens on MHC molecules, they could identify which peptides were created under different experimental conditions.

Their findings point to a unique translation mechanism. In the other known example of a non-methionine-initiated peptide, translation beginning at GCU or CAA is guided by a specific folded structure of mRNA nucleotides called the internal ribosome entry site. Schwab et al. have found that no similar structure is necessary for CUG-initiated translation. However, similar to the standard mechanism of AUG initiation, they found that ribosomes do scan for CUG. Additionally, the presence of a specific ribosome-binding sequence in mRNA (the "Kozak context") near a CUG site can enhance the efficiency of initiation there.

Schwab et al. have also suggested a possible purpose for this translation mechanism. Under stress, cells can down-regulate conventional translation, which curbs the production of viral proteins in the event of an infection but also suppresses the creation of antigens needed to flag down T cells for an immune response. Here, Schwab et al. report that peptides starting with leucine were produced in the absence of the protein eIF2, which normally aids in AUG-initiated peptide synthesis. Cells under stress slow conventional translation by restraining the function of eIF2. Therefore, CUG-initiated translation, which works without eIF2, might provide an out for stressed cells needing to create peptides. This alternative could be a great way to avoid pumping out viral proteins and still create antigens for T cell surveillance—unless, of course, viruses take advantage of the loophole for their own peptide production.

Schwab SR, Shugart JA, Horng T, Malarkannan S, Shastri N (2004) Unanticipated antigens: Translation initiation at CUG with leucine. DOI: 10.1371/journal.pbio.0020366

Developmental Milieu Influences a Gene's Role in Tumor Formation

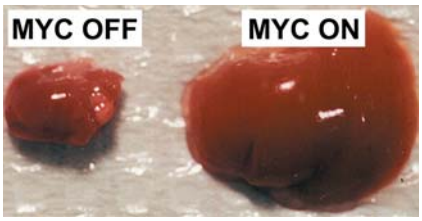
DOI: 10.1371/journal.pbio.0020375

Whether a person inherits a defective gene or acquires genetic damage by chance, two types of genes typically play a role in transforming a healthy cell into a cancer cell. Oncogenes and tumor suppressor genes are normally involved in cell growth, development, and cell differentiation. Both functions can be appropriated to ill effect by mutations. Single mutations in these genes rarely cause cancer on their own, but they predispose cells to additional insults that precipitate malignant transformation.

Susceptibility to cancer depends, among other things, on age. Though cancer in children is rare, the most common childhood cancers strike the hematopoietic system (leukemia), nervous system, and skeletal muscle system, while solid tumors of the lung, breast, prostate, and colon are more common in adults. This age differential suggests that an oncogene's ability to cause cancer in a particular cell type might depend on that cell's developmental stage. (A cell's gene expression profile differs with type and age; breast cells express different genes than liver cells, and immature cells express different genes than fully differentiated cells.) In a new study, Dean Felsher and colleagues show that age matters: activating oncogenes at different developmental time points in mouse liver cells produces different results.

Typically, once a cell is transformed, it stays in its "differentiative" state, that is, it stays in whatever developmental stage it was in when it became a tumor cell. But in a previous study, Felsher and colleagues found that turning off oncogenes in tumor cells allowed them to differentiate; these mature cells did not resume tumorigenesis after the oncogenes were reactivated. In this study, Felsher and colleagues show that the ability of the *MYC* oncogene to initiate liver cancer (hepatocellular carcinoma) in a transgenic mouse model varies with the age of the mouse.

To study the consequences of *MYC* overexpression in the liver cells of embryonic, neonatal, and adult mice, the authors used a biotech trick (called the Tet system) that controls gene expression



DOI: 10.1371/journal.pbio.0020375.g001

Developmental consequences of MYC overexpression

dose and timing with a drug. The system relies on the interplay of two elements: a gene (in this case, *MYC*) fused to a regulatory enhancer, and a transcription factor that binds to the enhancer and activates the gene. Administering a tetracycline-like drug (in this case, doxycycline) prevents the transcriptional activation of the gene.

Overexpressing the *MYC* oncogene in mice during embryonic development or at birth occasioned their demise fairly quickly (ten days and eight weeks after birth, respectively). In contrast, overexpression of *MYC* in adult mice resulted in tumorigenesis only after a long latency period. When the authors evaluated the cellular effects of *MYC* overexpression, they found that hepatocytes from neonatal transgenic mice showed evidence of increased proliferation (replicated DNA content) compared to normal hepatocytes, while transgenic adult hepatocytes showed increased cell and nuclear growth (some nuclei had as many as twelve genome copies instead of two) without dividing. Since these adult cells eventually developed into tumors, some clearly acquired the ability to divide, which the authors show is facilitated, among other events, by the loss of the p53 tumor suppressor.

Altogether these results suggest that whether oncogene activation can support tumor growth depends on the age of the host, which in turn suggests the role of genetically distinct pathways in young and adult mice. The consequences of *MYC* activation, Felsher and colleagues conclude, depend on the cell's developmental program, which determines whether a cell can grow and divide, or simply grow. In adult hepatocytes—which are normally quiescent—*MYC* requires additional genetic events to induce cell division and tumorigenesis; in immature hepatocytes—which are already committed to a program of cellular proliferation—*MYC* activation alone is sufficient. The next step will be to identify the epigenetic developmental factors, both internal and external, that lead to tumor formation, and how to prevent it.

Beer S, Zetterberg A, Ihrie RA, McTaggart RA, Yang Q, et al. (2004) Developmental context determines latency of MYC-induced tumorigenesis. DOI: 10.1371/journal.pbio.0020332

A Relay-Signal Model of Nematode Vulval Development

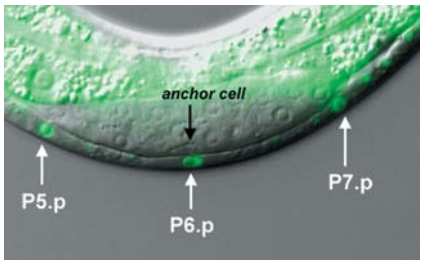
DOI: 10.1371/journal.pbio.0020376

A fundamental question in developmental biology is, how does a multicellular organism develop from a single cell? It's clear that one cell begets two, two beget four, and so on, but how do the newly created cells know which developmental fate to pick? Major insights into this question have come from identifying genes, molecules, and intercellular signaling pathways involved in a wide range of developmental processes. Operating in labyrinthine, often overlapping pathways, intercellular signals determine whether a cell divides, differentiates, migrates, and even lives or dies.

Scientists prefer to work out such problems in organisms with a manageable number of cells for obvious reasons, making the 959-cell soil nematode *Caenorhabditis elegans* a popular developmental model. *C. elegans* can exist as either a male or a hermaphrodite, and for some biologists, the hermaphrodite vulva—which consists of just 22 cells—is the perfect system for working out key aspects of intercellular signaling and cell fate.

In a new study, Alex Hajnal and colleagues challenge conventional thinking about vulval cell specification by identifying an enzyme that can amplify a signal's range and help turn three non-vulval precursors into vulval cells. Surprisingly, the enzyme, called ROM-1, accomplishes this feat by acting in the signal-receiving vulval precursor cells, rather than in the signal-sending cell that instructs the vulval cell fates.

The worm vulva forms a bridge between its gonad and the opening to the outer epidermal layer, called the cuticle. In the current model of vulval formation, a group of twelve epidermal cells, called Pn.p cells, lines the ventral surface of the worm.



DOI: 10.1371/journal.pbio.0020376.g001

Vulval precursor cells in *C. elegans*

Six of these cells, P3.p–P8.p, the vulval precursor cells (VPCs), have the potential to become vulval cells. During postembryonic development, the anchor cell in the larval gonad secretes an epidermal growth factor (called LIN-3) that activates the EGFR/RAS/MAPK signaling pathway and induces just three of the precursors to differentiate into vulval cells. The VPC closest to the anchor cell, P6.p receives most of the signal, and differentiates into eight vulval cells that form the tube linking the uterus to the gonad. Positioned on either side of P6.p, P5.p and P7.p receive a slightly attenuated signal, which, combined with a lateral signal from P6.p, gives rise to seven vulval cells that form vulval structures. The other three vulval precursors, P3.p, P4.p, and P8.p, it was thought, are too far away to receive the vulval induction signal and fuse into the surrounding epidermis.

The LIN-3 epidermal growth factors sit nestled within the cell membrane and must be "processed" to become active, prompting Hajnal and colleagues to look for candidate enzymes that could be doing the processing. They investigated the Rhomboid family of proteases, which are known activators of epidermal growth factor transmembrane proteins, and found one, ROM-1, with the amino acid profile required for catalytic protease activity. After showing that *rom* genes were not required for normal vulval development, the authors had a closer look at their role in vulval cell fate specification. Since loss of ROM-1 reduces the severity of a defect (in this case, multiple vulvas) caused by hyperactivation of the EGFR/RAS/MAPK pathway but has no effect on the precursors closest to the anchor cell, the authors conclude that ROM-1 enhances the EGFR/RAS/MAPK pathway, allowing it to reach the distant P3.p, P4.p, and P8.p precursors.

LIN-3 exists in two variant forms of different lengths, the longer one carrying a stretch of 15 extra amino acids in the region that is cleaved off to yield an active growth factor. Hajnal and colleagues show that ROM-1 only acts on the longer form to regulate the EGFR/RAS/MAPK pathway—and that the ROM-1/LIN-3 interaction occurs in the VPCs, independently of the anchor cell. They go on to propose a two-step model of vulval cell specification in which ROM-1 "extends the range" of the anchor signal, relaying it from the proximal to the more distant precursor cells by promoting the secretion of the long version of LIN-3. In normal development, LIN-3 secretion by the VPCs may serve initially to maintain the differentiation potential of all the precursors, while the anchor cell signal may seal their fates at a later phase.

Dutt A, Canevascini S, Froehli-Hoier E, Hajnal A (2004) EGF signal propagation during *C. elegans* vulval development mediated by ROM-1 rhomboid. DOI: 10.1371/journal.pbio.0020334

BMP Signaling Maintains Healthy Joint Cartilage

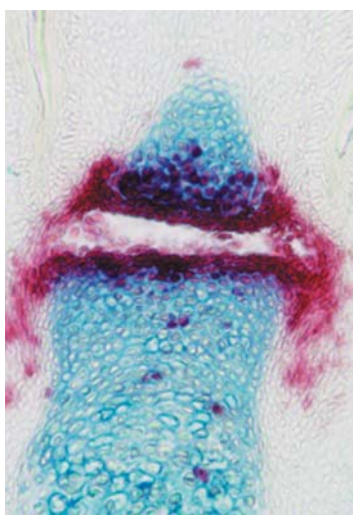
DOI: 10.1371/journal.pbio.0020395

The alarm clock rings and you jump straight out of bed, rattling downstairs to start the day. Or maybe you creak downstairs, each step a struggle because of stiffness and pain in your knees and other joints. If the second description fits the start of your day, then maybe, like 70 million Americans, you have arthritis, one of the most prevalent chronic health problems in the United States.

Arthritis is an umbrella term for more than 100 medical conditions. What all forms of arthritis have in common is that they affect our joints—places where two or more bones meet. In healthy joints, the ends of the bones are covered with cartilage, a tough but smooth tissue that, like the oil in a car engine, reduces friction between the moving parts. In the most common form of arthritis—osteoarthritis—breakdown of this cartilage, which is called articular cartilage, means the bones rub together, causing pain and loss of movement. Risk factors for osteoarthritis include age and family history.

If we could understand the molecular mechanisms that create and maintain articular cartilage, it might be possible to discover what goes wrong in our joints as we age and to find better treatments for arthritis. Embryologists have already discovered quite a bit about the earliest stages of joint formation. It is known, for example, that stripes of cells that form between developing bones subsequently develop into the permanent cartilage found in joints. Several members of a family of secreted proteins known as bone morphogenetic proteins (BMPs) are expressed in these stripes of cells, implicating BMP signaling (the transmission of messages produced by BMPs binding to cell-surface receptors) in early joint development.

David Kingsley's team has been investigating whether BMPs are also involved in the later development and maintenance of joint cartilage. To do this, the researchers designed a genetic system that inactivates BMP signaling late in mouse embryonic development. They inserted special DNA sequences called *loxP* sites on either side of *Bmpr1a*, a gene that encodes one of the BMP receptors. The *loxP* sites have no effect until an enzyme known as Cre is expressed, and then the DNA between the *loxP* sites is cut out and discarded. Because Kingsley's team knew that global inactivation of *Bmpr1a* early in development causes embryonic death, they linked the gene for Cre to DNA sequences that limit its expression to those regions of the embryo where joints eventually develop. The result: a mouse strain in which *Bmpr1a* receptor function is specifically lost only in tissues destined to become joints.



DOI: 10.1371/journal.pbio.0020395.g001

Targeting genes in joints

Most of the joints in this mouse strain formed normally. However, the mice rapidly developed severe arthritis after birth. By 7 days old, the expression of proteins normally found in cartilage was reduced, although at this stage the knee, for example, looked normal. By 7 weeks old (adulthood for mice), there were clear structural changes in the knee joints, and the articular cartilage was thinner and showed signs of wearing away. By 9 months old, the knees of the mutant mice largely lacked articular cartilage and the unprotected leg bones seemed to rub directly against each other.

All told, the joints in these mutant mice closely resembled those in people with osteoarthritis, suggesting that BMP signaling is necessary for the maintenance of healthy articular cartilage. This raises the possibility that mutations in BMP signaling components may underlie some of the genetic variation in human osteoarthritis risk and suggests that treatments designed to mimic or augment BMP signaling might help to maintain healthy joints. Finally, the genetic system described by Kingsley and coworkers should be useful for future investigations into joint formation and maintenance.

Rountree RB, Schoor M, Chen H, Marks ME, Harley V, et al. (2004) BMP receptor signaling is required for postnatal maintenance of articular cartilage. DOI: 10.1371/journal.pbio.0020355

Controlling the Timing of Gene Expression during Organ Development

DOI: 10.1371/journal.pbio.0020409

For more than 2,000 years, from the time of Aristotle onwards, it was thought that the complete body plan of human beings (and that of other animals) was present in the fertilized egg. During pregnancy, a preformed miniature human being, or homunculus, grew bigger and bigger; development was simply a process of growth.

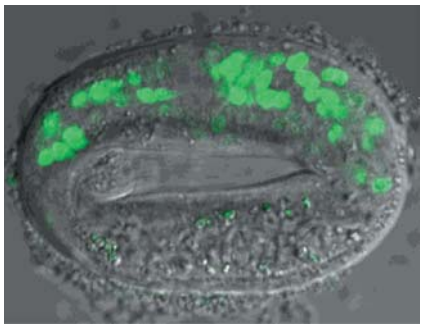
Then, in the mid-18th century, Carl Friedrich Wolff described how the chick gut, basically a tube, forms from an initially flat sheet of cells, overthrowing at a stroke the preformation theory of embryology. We now know that development is a complex series of coordinated processes that transforms the amorphous ball of cells produced from the fertilized egg by cell division into an intricate body containing numerous specialized tissues and organs. And we are beginning to understand how a wide array of transcription factors—proteins that bind to regulatory sequences within genes to control their expression—guide the sequential stages involved in development. It seems that these factors form regulatory networks that control the temporal and spatial waves of gene expression that underlie and are required for organized body building.

Susan Mango and her colleagues are studying the role of transcription factors in controlling organ development. The organ they are studying—the pharynx of the nematode worm—is relatively simple. This muscular tube, which passes bacteria (the food of this small soil-dwelling organism) from the mouth to the midgut, contains fewer than 100 cells of only seven different types.

To get an overall picture of the regulatory sequences within genes that are involved in the temporal control of pharyngeal development, the researchers identified 339 candidate pharyngeal genes by comparing gene expression profiles in mutant worm embryos that had excess pharyngeal cells with those in mutant embryos lacking pharyngeal cells. Then, by referring to a database that details gene expression patterns in nematode worms and embryos, the researchers classified 37 of their candidate genes as having early-onset expression and 34 as having late-onset expression.

Next, the scientists carefully examined the DNA sequence of each gene for candidate regulatory regions that might contribute to its temporal regulation. Of nine candidate motifs revealed by this search, six functioned as regulatory sites in in vivo assays. The researchers estimated that these six elements, together with sites that bind PHA-4—a member of a family of transcription factors that are important in digestive tract development in many animals—account for the timing of onset of expression of about half of the nematode's pharyngeal genes. Finally, the researchers used combinations of the newly discovered temporal regulation sites and PHA-4 sites in a genome-wide search that predicted pharyngeal genes and their time of onset of expression with greater than 85% accuracy.

From these results and those of previous studies, Mango and her colleagues propose a model to explain how the temporal control of pharyngeal gene expression needed for pharynx development is achieved. The earliest time for pharyngeal gene expression, they suggest, is determined by how well PHA-4 sticks to a particular gene's binding site. However, gene expression only occurs if other factors that bind to the regulatory sites are also present, and the exact combination of these factors determines which gene is active at any given time. The identity of these factors remains to be discovered. Nevertheless, at least for this simple organ, we now have a much better idea of how the complex process of organ formation is controlled at a molecular level, and it is likely that similar regulatory networks will underlie the formation of other organs as well.



DOI: 10.1371/journal.pbio.0020409.g001

Fluorescent reporter genes expressed in the developing *C. elegans* foregut

Gaudet J, Muttumu S, Horner M, Mango SE (2004) Whole-genome analysis of temporal gene expression during foregut development. DOI: 10.1371/journal.pbio.0020352.

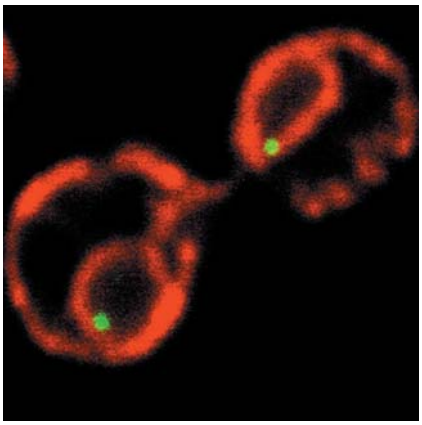
For Gene Activation, Location Matters

DOI: 10.1371/journal.pbio.0020381

Multicellular organisms contain a complete set of genes in nearly all of their cells, each cell harboring the potential to make nearly any protein in their genome. The same holds true for a single-celled bacterium or yeast. Yet a cell activates only a fraction of its genes at any given time, calling on a number of different mechanisms to activate the right genes at the right time. To metabolize sugar, for example, a cell needs to synthesize proteins involved in sugar metabolism, not protein repair, and vice versa. In a new study, Jason Brickner and Peter Walter report a mechanism for gene activation that depends on shuttling DNA to a particular location within the nucleus.

In organisms whose cells have nuclei (eukaryotes), genomes lie within the nucleus (called the nucleoplasm) but also interact with the inner nuclear membrane. Transcription factors activate gene expression by binding to a promoter sequence in the gene's DNA. The physical structure of DNA—which is packaged with proteins into chromatin—affects gene expression by controlling access to DNA. Where chromatin exists in the nucleus also influences gene expression. Heterochromatin—stretches of highly condensed chromatin—typically lines the nuclear periphery, and genes bundled into heterochromatin are typically silent. Active transcription generally occurs in the less condensed euchromatic regions. But since euchromatic regions are also silenced when they associate with heterochromatin along the membrane, it is thought that delivering chromatin to the nuclear periphery regulates transcriptional repression. Brickner and Walter, however, found evidence of the opposite effect—recruiting genes to the nuclear periphery can promote their activation—suggesting that nuclear membrane recruitment plays a much broader role than previously suspected in gene regulation.

To explore the consequences of chromatin location, the authors focused on a yeast gene called *INO1*, which encodes inositol 1-phosphate synthase, an enzyme involved in phospholipid (fat) biosynthesis. *INO1* is also a target gene of the "unfolded protein response," which



DOI: 10.1371/journal.pbio.0020381.g001

The INO1 gene (green) is recruited to the nuclear membrane (red) upon activation

is triggered when unfolded proteins accumulate in the endoplasmic reticulum, a subcellular organelle where secreted proteins are folded. The *INO1* gene contains a regulatory element (called UASINO) within its promoter region that responds to inositol availability. Genes under the control of this element are transcriptionally repressed by a repressor, Opi1, and activated by two transcription factors, Ino2 and Ino4. The presence of unfolded proteins sets off a chain of events to relieve Opi1 repression and allow activation of *INO1*.

Through a series of genetic and biochemical studies, Bricker and Walter show that Ino2 and Ino4 are always bound to the *INO1* promoter. Opi1 associates with the chromatin, restricting the *INO1* locus to the nucleoplasm and repressing transcription. Induction of the unfolded protein response bumps Opi1 off the chromatin and, with Opi1 out of the way, *INO1* travels to the membrane and transcription proceeds. Crucially, the authors show that artificial recruitment of *INO1* to the nuclear membrane can be enough to activate the gene. There are several mechanistic aspects of this model to figure out still, but Brickner and Walter argue that for *INO1*, gene recruitment to the nuclear membrane promotes its activation. In light of other recent work, this phenomenon may be emerging as a more general mechanism for regulating eukaryotic gene expression.

Brickner JH, Walter P (2004) Gene recruitment of the activated *INO1* locus to the nuclear membrane. DOI: 10.1371/journal.pbio.0020342

High Affinity: Making Up for Being Male

DOI: 10.1371/journal.pbio.0020387

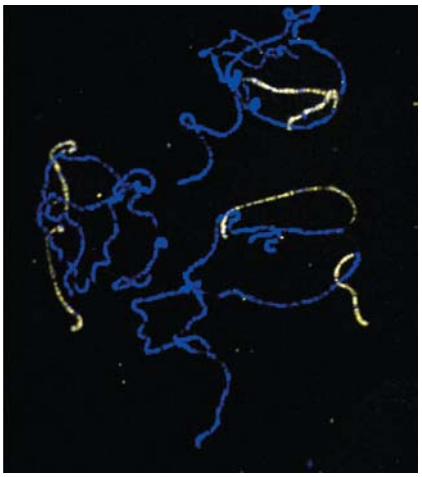
Because males and females possess different numbers of the two sex chromosomes (for instance, in mammals, XX in females versus XY in males), the potential "dose" of each gene differs. Without some compensating mechanism, female mammals would express twice the quantity of an X-linked gene as males. The same holds true in the fruitfly *Drosophila*, in which the female carries two X chromosomes, while the male carries only one.

In mammals, dosage compensation is achieved by silencing one of the X's in the female. *Drosophila* takes the opposite tack, doubling the output from the single male X chromosome. It does so through the creation of "compensasomes," protein–RNA complexes that bind to the X chromosome and boost gene transcription. One model of compensasome activity has posited a two-step mechanism, in which the complexes form only at 35–40 specific "entry sites" along the X, and then spread out to the surrounding regions. In this issue, Delphine Fagegaltier and Bruce Baker test this model and show that its predictions do not match experimental results.

The compensasome complex includes half a dozen proteins collectively known as MSLs (for "male-specific lethal"), along with two pieces of RNA, *roX1* and *roX2*. Fagegaltier and Baker reasoned that, according to the entry-site model, if a piece of the X not containing one of the entry sites was transposed to an autosome (non-sex chromosome), it should be unable to recruit MSLs and therefore be unable to form compensasomes. To test this prediction, they used autosomes into which various pieces of the X had been transposed. Contrary to prediction, they found that even the smallest pieces could recruit MSLs, whether or not they contained entry sites. Furthermore, the pattern of MSL binding was exactly the same as if the fragment of the X was still on its native chromosome, suggesting that each of the hundreds of sites at which compensasomes are found function autonomously to recruit them.

Another prediction of the entry-site model is that compensasomes should spread out from the entry site, along the

chromosome. And here again, the model does not hold up—Fagegaltier and Baker found that even when entry sites from the X chromosome are put close to an autosomal region, compensasomes never spread from the X onto these regions. These results suggest that spreading is not an innate function of the compensasome, and further strengthens the case for autonomous recruitment all along the X. In place of the two-step "entry site plus spreading" model, the authors propose a model based on differential affinity for compensasome components. They suggest that the 35–40 "entry sites" are simply high-affinity sites that recruit MSLs first, based on intrinsic differences that allow them to bind and hold MSLs more strongly than other sites. Once these sites are occupied, additional compensasome components can bind to lower-affinity sites. This mechanism can account for observed compensasome activity without the restriction to a limited number of entry sites and the requirement for spreading.



DOI: 10.1371/journal.pbio.0020387.g001

Compensasomes do not spread from the X chromosome onto autosomal material translocated onto the X

Fagegaltier and Baker note that while compensasome spreading does not normally occur during dosage compensation on the X chromosome, it has nonetheless been documented for some *roX* transgenes. They propose that the additional binding observed specifically around *roX* transgenes results from a mass action of compensasomes, as *roX* transgenes would act as assembly sites for compensasomes, just as ribosomal RNA genes do for ribosomes. Once formed, compensasomes may bind locally to other neighboring sites.

While the details of dosage compensation and the dosage compensation complexes now clearly differ between mammals and flies, there are broad similarities, including the widespread modification of chromatin structure and the use of RNA components in the compensation machinery. A deeper understanding of the process in flies may help shed light on the details of compensation in other organisms as well.

Fagegaltier D, Baker BS (2004) X chromosome sites autonomously recruit the dosage compensation complex in *Drosophila* males. DOI: 10.1371/journal.pbio.0020341

MicroRNA Is a Major Regulator

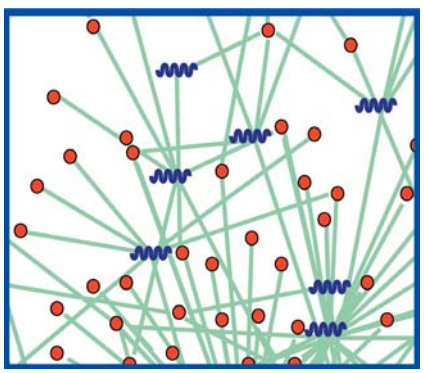
DOI: 10.1371/journal.pbio.0020396

Since their discovery a decade ago, microRNAs (miRNAs) have emerged as major regulators of gene expression in eukaryotes of all kinds. Only 20 to 40 nucleotides long, a miRNA binds to a specific target sequence within a much longer messenger RNA (mRNA), inhibiting its translation and thus controlling expression of the corresponding gene even after the DNA itself has been read. Within the human genome, there are about 250 genes that code for miRNAs. Each miRNA has the potential to bind to many different transcripts. Variations in miRNA sequence dictate the gene transcripts to which each will bind most strongly.

It has become clear that miRNAs play a critical role in controlling gene expression, for example, in larval developmental transitions and neuronal development in the worm *Caenorhabditis elegans*, growth control and apoptosis in the fruitfly *Drosophila melanogaster*, hematopoietic differentiation in mammals, and leaf development, flower development, and embryogenesis in the plant *Arabidopsis thaliana*. Despite their significance, the full range of genes miRNAs target is unknown, as is the best method for discovering them. In a new study, Debora Marks, Chris Sander, and colleagues describe an algorithm for determining the targets of miRNAs, and show they include more than 10% of all human genes.

The algorithm uses three factors to evaluate whether a potential target site is likely to actually be regulated by miRNA. First, the target site must have some degree of sequence complementarity to one or more of the known miRNAs. Second, the strength with which the predicted target and its miRNA bind together, which can be calculated from the sequence and other structural factors, must be higher than some threshold. Finally, evolutionary conservation—the presence of the target–miRNA pair in different organisms—is factored in, because the likelihood that the target and miRNA actually pair in vivo is greater if the pair is found in multiple types of organisms.

Using these principles, and the specific weighting they assigned to each factor, Marks and colleagues identified 2,273 genes in humans, rats, and mice that are likely targets for miRNA regulation. This is probably an underestimate of the total, since the researchers required each candidate gene to have at least two miRNA target sites. The authors identified another 2,128 genes with only one target site, but note that the false-positive rate here is likely to be high. Whatever the final number, the implication is that several thousand of our approximately 30,000 genes are under the control of miRNAs. Of special interest is that these putative targets include many genes known to be associated with the fragile X mental retardation protein, a crucial but still poorly understood player in mRNA regulation, whose absence leads to a type of mental retardation called fragile X syndrome.



DOI: 10.1371/journal.pbio.0020396.g001

microRNA gene networks

The researchers' findings also reinforce several emerging principles of miRNA-based regulation. First, it is widespread

among multicellular eukaryotes, and sequences are surprisingly conserved. Of the 78 known miRNAs in *Drosophila*, 28 have close relations in mammals. Second, an individual miRNA may regulate multiple genes—Marks and colleagues found that the average miRNA interacts with seven distinct mRNAs, with a range from 0 to 268. Third, the genes regulated by a single miRNA may be functionally related, such as components of the protein degradation system or specific signal transduction pathways. Fourth, single genes may be regulated by multiple miRNAs—the gene that encodes amyloid precursor protein, for example, has at least eight miRNA sites—suggesting that expression may be combinatorially controlled by numerous cellular influences.

These results provide resources for a host of experiments to elucidate the mechanism of miRNA action, which is not well understood. Several of the identified mammalian miRNA–target pairs have near-perfect matching sequences. In both plants (where miRNAs were first discovered) and animals, such matches are associated with degradation of the mRNA.

The authors fully recognize that their algorithm, called miRanda, is not the last word in miRNA target identification. In order to improve both the search for targets and the algorithm itself, they are making the algorithm and full sets of results in vertebrates available free to other researchers (www.microrna.org), who can modify its parameters as experimental results and new models dictate.

John B, Enright AJ, Aravin A, Tuschi T, Sander C, et al. (2004) Human microRNA targets. DOI: 10.1371/journal.pbio.0020363

Slime Mold Myosin Thick Filament Assembly Dissected

DOI: 10.1371/journal.pbio.0020392

The movements needed to read this synopsis—turning the page, tracking along the lines, sitting, breathing—all require myosin, a molecular motor in muscle that transforms chemical energy into small but deliberate motions. But beyond these macro-movements, myosin is also required for the micro-movements of individual cells and their organelles and for determining cellular architecture.

There are many different myosins, but they all have the same general structure. At one end is a globular head, which is responsible for motor activity. This head binds ATP—the cell's power supply—and actin, an important component of the cytoskeleton of cells. Next comes a helical neck or lever region. Finally, there is a long helical tail, which has different and somewhat poorly understood functions in the different myosins.

Myosin II, the classical form of myosin found in essentially all eukaryotic cells, is constructed from two heavy chains (which contain the three regions described above) and two pairs of light chains (which stabilize the neck region). The long helical tail of myosin II is formed by the two heavy chains wrapping around each other and is involved in getting myosin II to the right place in the cell, as well as in assembling it into filaments.

Individual myosin II molecules can make tiny molecular motions. ATP cleavage induces a shape change in the globular head, which is transmitted to the lever region of the molecule. Angular rotation of this region moves the myosin along the actin filament. But to achieve the larger movements that are necessary to, for example, split cells apart during cell division, individual myosin II molecules group together to form highly regular bipolar structures called bipolar thick filaments (BTFs). In these, the globular myosin heads are positioned on either side of the filament, and the tail regions are clustered in the middle. This geometry enables myosin II molecules in thick filaments to pull from either side, generating contractile forces.

James Spudich's team has been studying the assembly of these thick filaments in the slime mold *Dictyostelium discoideum*, an organism beloved by developmental and cellular biologists because of its simple development and ease of manipulation. In the present study, the researchers examined the physical properties of various fragments of the myosin tail to find out how the self-assembly and disassembly of the BTFs are regulated. They already knew that the addition of phosphate groups on three specific threonine amino acid residues in this region (through a chemical reaction called phosphorylation) is important for regulating BTF assembly; they knew this from studies showing that mutation of these residues to aspartic acid, which mimics phosphorylated threonine, inhibits BTF formation. Here, the researchers show that a specific tail fragment of the myosin heavy chain containing the three crucial threonine residues assembles into a structure with some, but not all, of the properties of BTFs. However, replacing these threonine residues with aspartic acid prevents any self-assembly of the fragment.

Further experiments in which different tail regions were nibbled away and the assembly properties of the remaining fragments were determined suggest that the myosin tail contains a series of elements that correlate with the distribution of charged amino acids along the tail, some of which favor assembly and some of which favor disassembly. But it's not just the tail that is important. For myosin II to form fully fledged BTFs of a defined size, it seems that the addition of some kind of globular head—in these experiments one composed of green fluorescent protein so that it could be examined—is necessary. The overall result is a molecule that is finely poised to self-assemble into BTFs in response to one or two charge changes produced by phosphorylation. Consequently, the myosin contractile system can respond rapidly to environmental changes.

Although *Dictyostelium* myosin II is somewhat different from vertebrate myosin II, the general principle by which myosin assembly and disassembly are regulated seems likely to hold for other myosins and so might throw light onto human disorders that involve myosin defects. But more fundamentally, similar principles may hold for spatial and temporal regulation of the many other macromolecular assemblies that are at the heart of cell and developmental biology.

Hotstetter D, Rice S, Dean S, Altman D, McMahon PM, et al. (2004) *Dictyostelium* myosin bipolar thick filament formation: Importance of charge and specific domains of the myosin rod. DOI: 10.1371/journal.pbio.0020356

Sleeping, Waking, … and Glucose Homeostasis

DOI: 10.1371/journal.pbio.0020415

We often think of ourselves as either a day person or a night person—one who rises with the sun, raring to go, or one who prefers to stay up through the night to get things done. Regardless, we each have our regular waking and sleeping cycles. It's been known for some time that variations in sleep and wakefulness are part of our circadian rhythm, or molecular clock. A portion of the brain called the hypothalamic suprachiasmatic nucleus (SCN) regulates this biorhythm. When this area of the hypothalamus is destroyed in animal models, the circadian rhythm is disrupted. Two transcription factors (proteins that regulate gene expression) called Bmal1 and Clock regulate aspects of circadian rhythm, possibly by regulating neurons in the SCN.

Other aspects of human physiology are also regulated in a circadian manner. Besides altering sleep and wakefulness patterns, ablation of the SCN alters the ability to regulate sugar levels. Sugar (glucose) levels must be maintained within fairly narrow limits for survival. This regulation is controlled in part by a balance between blood sugar level and insulin production (insulin lowers the blood sugar level). In people and in mouse models, both glucose level and insulin level are subject to circadian rhythms. It isn't clear, however, if this is a behavioral effect, whereby the disruption of the SCN might alter our feeling of being well fed—that is, being sated—as eating has a profound effect on blood sugar levels.

Garret FitzGerald and colleagues tested the effect of the molecular clock genes in glucose regulation (homeostasis) by examining mice in which Clock and Bmal1 were impaired. In normal mice they observed a peak in glucose levels early in the day. This diurnal regulation was lost in the mutant mice. Furthermore, whereas the normal mice could fairly easily return their glucose levels to normal when they were artificially treated with insulin, this ability was severely impaired in the mutant mice. What's more, a high-fat diet amplified this circadian variation in the normal animals, but the rhythm was abolished in the mutants on a high-fat diet. Thus, the authors demonstrated that circadian control of blood glucose levels is due directly to the presence of these transcriptional factors rather than due to some other behavioral effect that ablation of the hypothalamus might have caused. It's possible, therefore, that besides what we eat, our internal circadian clock could also be an important regulator of blood sugar levels.

DOI: 10.1371/journal.pbio.0020415.g001

Metabolic clock regulation of glucose homeostasis

What is still left to be explored is whether the change in glucose that results from disruption of the Clock and Bmal1 genes is due to the transcription factors' effect as circadian regulators or to an activity of these transcription factors that is unrelated to circadian rhythm generation. But the study does raise the possibility that when you eat may be as important to your health as what you are eating.

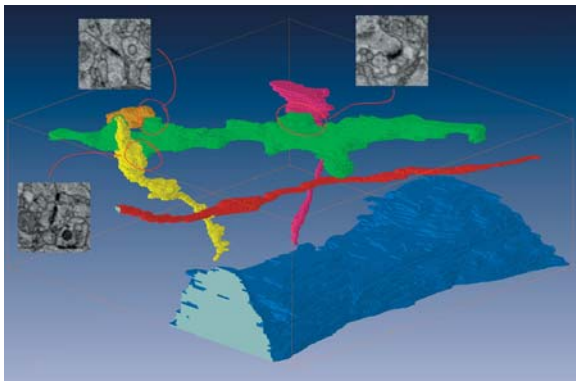
Rudic RD, McNamara P, Curtis AM, Boston RC, Panda S, et al. (2004) BMAL1 and CLOCK, two essential components of the circadian clock, are involved in glucose homeostasis. DOI: 10.1371/journal.pbio.0020377

Reconstructing Neural Circuits in 3D, Nanometer by Nanometer

DOI: 10.1371/journal.pbio.0020388

Understanding how the brain processes and stores information depends in large part on knowing which neurons are involved in a particular process and how they're organized into functional networks. Each of the 10 billion or so neurons in the brain has thousands of connections to other neurons, sending (via axons) or receiving (via dendrites) the signals that allow us to think. Each neuron can transmit signals to both local and distant neurons, and it is by mapping these networks that neuroscientists can discern correlations between neural connectivity and physiological responses and ultimately unveil the computational algorithms underlying brain function. Since the beginning of cellular neuroscience at the end of the 19th century, neuronal connections have been explored by tracing axons and dendrites under the light microscope. But even with the resolution of state-of-the-art light microscopy, this approach works only if a small subset of neurons is stained and thus leaves most of the network hidden.

Electron microscopy, on the other hand, can provide the spatial resolution necessary both to resolve processes in densely packed neural "wire bundles" and to identify synapses faithfully, but individual electron microscopic images are restricted to two dimensions. Transmission electron microscopy provides cross-sectional images through tissue, while scanning electron microscopy typically provides the appearance of 3D but in reality maps only the specimen surface and is thus blind to the connections within. It's possible to wrest 3D information from the transmission electron microscope by using tilt-series tomography, but sections can't be much thicker than 1 micron (a millionth of a meter). Data from thicker volumes can be obtained, but the process so far has been so painstaking and time-intensive—it involves, among other labor-intensive tasks, manually reconstructing serial sections—that few undertake it.



DOI: 10.1371/journal.pbio.0020388.g001

Neurite Reconstruction Manual reconstruction of selected processes in cortical tissue

It should, however, be possible to get similar data with "serial block-face imaging," which involves repeatedly cutting section after section from a plastic-embedded block of tissue and photographing what's left. Scanning electron microscopy is needed for this task, but sample preparation methods are like those used for transmission electron microscopy, albeit with a few additional steps to enhance contrast.

This is exactly what Winfried Denk and Heinz Horstmann have done to obtain "truly 3D datasets" using a method they call "serial block-face scanning electron microscopy" (SBFSEM), for which they constructed a "microtome" that goes inside the scanning electron microscope chamber. The resolution achieved is sufficient to reveal "even the thinnest of axons" and identify synapses. The SBFSEM method can generate stacks of thousands of ultra-thin sections, 50–70 nanometers (a nanometer is a billionth of a meter) thick, generating 3D datasets to reconstruct the topology and circuitry of neurons in brain tissue.

The authors' custom-designed microtome holds the tissue block in a way that ensures image alignment and maintains focus; all the while the specimen surface is positioned close enough to the objective lens to allow high-resolution imaging.

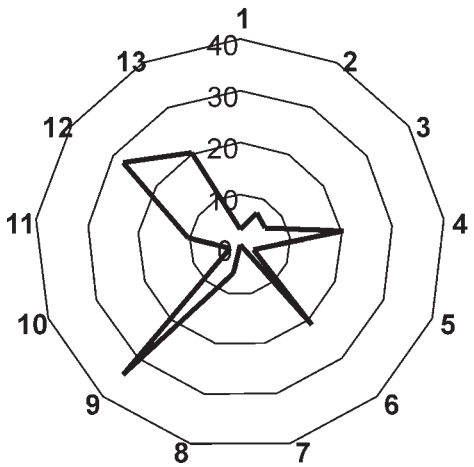
Denk and Horstmann expect that with this method they might ultimately be able to cut sections thinner than the 50 nanometers that their current setup manages. This then would allow them to cut sections even thinner than what is routinely possible in conventional transmission electron microscopy. While the authors doubt that the lateral resolution will ever reach that of transmission electron microscopy, they also argue that such high resolution may not actually be needed to trace neuronal connectivity. On the other hand, the method accelerates 3D electron microscopic data collection "by several orders of magnitude" by obviating the need for the labor-intensive adjustments to correct alignment and distortion required by other methods, an advance that is crucial for large-volume neuroanatomy and might, in addition, open up many hitherto inaccessible problems to ultra-structural investigations.

Denk W, Horstmann H (2004) Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. DOI: 10.1371/journal.pbio.0020329

Only Connect: The Functional Architecture of Brain Connectivity

DOI: 10.1371/journal.pbio.0020411

Imagine three cities, A, B, and C, splayed across the landscape to form a triangle, with each connected to the other two by two-lane roads. Such an arrangement of cities and roads constitutes a structural network. On any given day traffic may flow, say, only from A to B to C, or in both directions between A and B but from C only to A, or in both directions between all three, or any one of ten other arrangements. Within this structural network, then, there are 13 possible functional networks. If these cities are embedded within a larger network of routes and destinations, their particular triangular traffic pattern represents a "motif" of connectivity, akin to a recurring musical motif within a larger symphony.



DOI: 10.1371/journal.pbio.0020411.g001

Such connectivity networks are central to information processing in the brain, and understanding the recurring structural and functional motifs they contain is one way to begin to dissect how the symphony of brain function is composed. In this issue, Olaf Sporns and Rolf Kötter identify several common motifs in real brain networks, and show that brains tend to maximize the number of functional motifs while keeping the number of structural motifs relatively low.

The authors began with the frequency of motifs of different sizes (two, three, four, or five nodes) found in the visual cortex and whole cortex of the macaque monkey, the cat cortex, and the nervous system of the nematode *Caenorhabditis elegans*. For comparison, they generated matrices that contained an equivalent number of components (nodes and connections), but whose connections were either random or lattice-like, in which all nearest neighbors were connected. They found that, compared to the artificial networks, the biological ones were relatively low in structural diversity. For instance, macaque visual cortex contained instances of 3,697 different motifs with five nodes, versus 8,887 for equivalent random networks. Functionally, however, unlike the artificial systems, the biological systems were maximally diverse, with the maximum functional motif diversity (e.g., 13 for three vertices and 9,364 for five vertices) observed in all motif sizes they investigated.

The researchers also found some intriguing patterns within this maze of connectivity. For instance, not all motifs were found in equal numbers. A common functional motif for three vertices was for both A and C to communicate back and forth with B, but not with each other. This structure allows B to function as an integrator of signals from A and C, while keeping the activities of A and C distinct from one another. This kind of structure is widespread throughout the nervous system.

The authors then ran an evolutionary algorithm on their artificial networks. They showed that by selecting for maximal functional motif number, the structure of the artificial systems quickly came to resemble the structure of the real ones, with dense local connections and relatively fewer long-distance ones. Such a structure, termed " small world" connectivity, promotes cooperation between functional units, and efficient information exchange. Taken together, these results suggest that one factor that may drive the evolution of neural architecture is the maximization of functional connectivity within a network of relatively few neural actors.

Sporns O, Kötter R (2004) Motifs in brain networks. DOI: 10.1371/journal.pbio.0020369

Paying Attention to Memory

DOI: 10.1371/journal.pbio.0020407

If you could peer inside someone else's head, you'd see a scrunched-up gelatinous mass of tissue, weighing roughly a kilogram, homogeneous to the naked eye—in other words, a brain. The seeming uniformity of the overlying cerebral cortex, which has so outstripped other parts of the brain over the course of evolution that it makes up more than 80% of the brain, is belied by centuries of painstaking neuroscience. Some of the most compelling early evidence that parts of the cortex are specialized in their duties came from gun-shot wounds during the first world war. For instance, bullets lodged in the back of the brain disrupted sight in discrete portions of the visual scene, prompting insights into the localization and function of visual cortex.

The study of the front of the brain has a similar history of injury leading to insight. Phineas Gage, a railroad worker, had a 3.5-foot-long tamping iron blow straight through his frontal lobes and turned from a responsible, mild-mannered geek into an unruly exhibitionist overnight. Parts of the prefrontal cortex that he damaged have since been much studied for their involvement in motivation and emotional control.

More recent work has implicated other parts of the prefrontal cortex in working memory. Working memory is famously illustrated by your ability to temporarily remember a seven-digit telephone number, roughly the amount of information that you can store on-line in working memory for the duration of a task like phoning for a pizza.

Monkeys can be trained to remember information much like you remember a phone number, and then use the memory for gaining a reward (usually juice rather than pizza). They can learn to remember the specific location of a briefly flashed target on a screen and then, when cued, make an eye movement to look directly at that location. Previous research has shown that neurons in the prefrontal cortex maintain high rates of activity while monkeys remember the target location, and gradually the idea that the prefrontal cortex specializes in maintaining these transient memories has risen to dominance over other ideas about its functions.

In this issue of *PLoS Biology*, Mikhail Lebedev and his colleagues challenge this prevailing view with evidence that most prefrontal cortex neurons may not be so closely tied to working memory after all. As in previous research, they also trained monkeys to make an eye movement to a remembered target, but instead of only seeing one target, the monkeys saw two potential target locations during the course of the task. The monkeys had to pay attention to one of the potential targets, but this was not necessarily the one they would have responded to and was not the one they had to remember. To perform the task successfully, the animals had to engage their working memory, but most of the neurons the researchers recorded increased their activity selectively to the target that was the focus of attention.

Despite decades of research, the degree to which one region of the brain can be thought of as dedicated exclusively to a particular function is still much debated. These results do not refute the idea that the prefrontal cortex plays an important role in working memory. However, the authors suggest that this area may be more important in focusing the attention needed to remember that phone number, rather than actually holding that number in your mind.

Lebedev MA, Messinger A, Kralik JD, Wise SP (2004) Representation of attended versus remembered locations in prefrontal cortex. DOI: 10.1371/journal.pbio.0020365

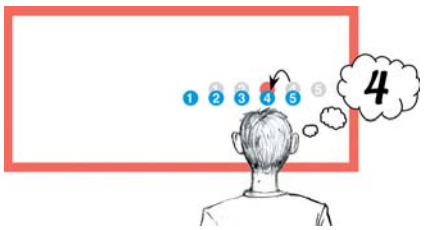
One Brain, One Vision

DOI: 10.1371/journal.pbio.0020414

Not all devices that measure the same property do it in the same way—a clock might use a spring system or it might be digitally synchronized to a transmitted signal. Although both have the same goal of reporting accurate time, each is subject to different errors. Sometimes even the same device uses different systems to measure the same property. A relatively simple device like a camera will use one sensor system to capture light intensity for an image and a second sensor to capture light intensity for making automatic adjustments of aperture and flash. It does not seem outlandish, therefore, that the brain might also have developed multiple sensory systems to achieve different goals. Indeed, an influential hypothesis has argued that people use two separate visual processing systems in much the same way as a camera—one for creating our perception of the world and another for guiding our actions within it.

One line of evidence supporting this dual hypothesis comes from an illusion known as Roelofs effect. Usually, people are pretty good at judging the location of even a small object. But if the small object is surrounded by a large frame and the frame itself is not centered in front of the person who is judging it, the viewer will perceive the object as shifted in a direction opposite that of the frame. This may not in itself be surprising, but the same person who perceives an offset of the object where none exists is nonetheless able to grasp it without difficulties.

In this issue of *PLoS Biology*, Paul Dassonville and his colleagues re-examine the seeming dissociation of visual analysis for perception and action, and call it into question. Through a careful quantitative analysis of the conditions under which the Roelofs effect occurs, they find that it traces not to an illusory perception of the object location but to an illusory perception of self. The large frame, presented under experimental conditions in which subjects sit in darkness without access to a normal rich sensory environment, actually causes people to incorrectly perceive their own centers as rotated towards the frame and therefore to conclude that the small object is offset with respect to themselves. This may seem like a subtle distinction, and yet, since it is the observer's frame of reference that is altered, that same distorted frame of reference will be used to guide movement. Thus, the error in movement planning should cancel the error in perception, and people should have no trouble reaching for the object despite their misperception, which is indeed what is observed.



DOI: 10.1371/journal.pbio.0020414.g001

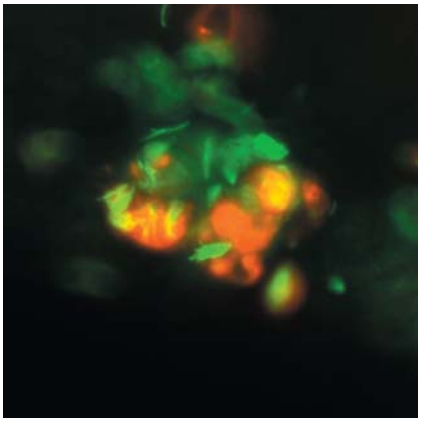
Others have questioned the hypothesis that two separable neural systems process the visual world for perception and action, but this study removes one of the strongest pieces of evidence in its favor with a precise alternative explanation. No two brains may see the world identically, but the authors suggest that it may be time to concede that a single brain, at least, has the same world view.

Dassonville P, Bala JK (2004) Perception, action, and Roelofs effect: A mere illusion of dissociation. DOI: 10.1371/journal.pbio.0020364

A Clear View of Mycobacterial Infection

DOI: 10.1371/journal.pbio.0020410

Fighting an infection might seem to be a battle between David and Goliath, given the relative sizes of bacterial infectious agents and the animals they infect. But on closer examination it is more often a chess match between two skilled opponents who have the uncanny ability to anticipate each other's moves. *Mycobacterium tuberculosis* causes tuberculosis (TB) in people, and related species that infect other animals are used as model systems for the study of TB. Much progress has been made in identifying the armaments (or virulence factors) of the bacteria. But the interplay, or chess match, between the bacterium and the animal it infects is much less clear. One of the host's first moves against the mycobacterium is the formation of a granuloma. Granulomas are tightly aggregated structures that consist of macrophages—one of the first lines of defense of the immune system—within which the infecting bacterium grows. Besides these and related cells that are present at the site of infection, additional macrophages and other immune cells are recruited in the formation of the granuloma. Although granulomas are required for the elimination of the infection, Lalita Ramakrishnan and colleagues have now shown that the bacteria have a game plan of their own.



DOI: 10.1371/journal.pbio.0020410.g001

Mycobacteria co-opt granulomas for their growth and spread

One problem in understanding the interaction between the mycobacterium and the host has been that it occurs deep in the lung of the infected animal, which makes it difficult to analyze how each of the animal or bacterial factors affect the strategic interplay between the host and pathogen. To overcome this limitation, Ramakrishnan and colleagues used zebrafish embryos, which are transparent and can be infected by a relative of the TB pathogen, *M. marinum*. This enables the researchers to watch cells as they are recruited into the granuloma.

Some of the virulence factors of mycobacteria are encoded in an area of the genome called the RD1 locus. In a mouse model, a strain of the bacteria missing RD1 causes far less pathology than a strain with the full complement of genes. The RD1 locus is also absent in the bacterial strain *M. bovis* that is used as an attenuated TB vaccine. But the precise role of RD1 in infection remains obscure.

By visualizing in zebrafish infections of a virulent strain of *M. marinum* and a strain with an RD1 deletion, Ramakrishnan and colleagues have observed that RD1 is actually required for granuloma formation but isn't needed for the bacteria to infect macrophages. What's more, macrophages that are infected with mycobacteria that contain RD1 produce a signal that further recruits macrophages to granulomas. This might seem an odd virulence strategy, as macrophages are required for mycobacterial elimination. But in this ongoing chess match, the virulent mycobacterium exploits the host's defense—granuloma formation—by providing additional macrophages for the bacteria to infect.

The end game of the chess match remains unclear. While granulomas are required for protection against mycobacteria, they are not completely effective. Thus, these bacteria have developed a strategy to recruit the normally defensive cells of the host to their advantage, but it remains to be shown what tips the balance between the macrophages' ability to clear the infection and their unwitting participation in the development of TB.

Volkman HE, Clay H, Beery D, Change JCW, Sherman DR, et al. (2004) Tuberculous granuloma formation is enhanced by a *Mycobacterium* virulence determinant. DOI: 10.1371/journal.pbio.0020367

Predicting Risk of Mosquito-Borne Disease in Variable Environments

DOI: 10.1371/journal.pbio.0020390

Malaria remains one of the greatest threats to global health, infecting more people than ever before. Confined mainly to the tropical areas of Africa, Asia, and Central America, malaria hits Africa the hardest; the poverty-stricken lands of sub-Saharan Africa account for 90% of malaria infections worldwide. Despite ongoing efforts to battle the disease—by controlling mosquito populations, reducing human contact, and developing drug prevention and treatment—the crisis continues to worsen.

The primary variables affecting risk of infection are the rate at which humans are bitten and the proportion of mosquitoes that are infectious. These two factors are often regarded as positively correlated, meaning that if the percentage of infectious mosquitoes increases, so will the human biting rate. But in a new study, David Smith, Jonathan Dushoff, and F. Ellis McKenzie challenge this assumption. Using a mathematical modeling approach to examine the relative contributions of the two factors across different landscapes and seasons, the authors show that the factors are not positively correlated. In fact, their calculations show that the rate humans are bitten and the proportion of infectious mosquitoes peak at different times and places.

Their modeling results suggest that the standard metric to estimate risk of infection—the number of times an infectious mosquito bites a person per day, called the entomological inoculation rate (EIR)—is flawed when variable conditions are taken into account. Using the average EIR to estimate average risk of infection in variable environments generates biased estimates because there is not a direct correlation between EIR and the proportion of humans who are infected.

The distribution of humans and suitable habitat for mosquito larvae varies across the landscape. And the density of mosquito populations varies seasonally, rising and falling with changes in rainfall, temperature, and humidity. Temporal and spatial variations in mosquito populations affect the rate humans get bitten, the number

of infectious mosquitoes, and the risk of infection. To understand how these space- and time-induced variations in mosquito populations shape the epidemiology of human infection, Smith and colleagues developed a set of mathematical models that calculate the relative impact of different parameters, in order to determine which factors most influence where and when risk of infection is highest.

First, they evaluated what factors affect the primary components of the EIR: the human biting rate and the proportion of infectious mosquitoes. As expected, the model predicts that fluctuations in mosquito density influence the EIR by changing the human biting rate. As more people are bitten, more people become infected; consequently, more mosquitoes feed on infected humans and so become infectious. Only adult mosquitoes transmit infection, so as mosquito populations age, the proportion of infectious mosquitoes increases. During the dry season, few mosquitoes are born, and so while the human biting rate and EIR decline, the proportion of infectious mosquitoes increases.

Because mosquito populations are densest near breeding sites—where younger mosquitoes outnumber adults—the human biting rate and the number of bites by infectious mosquitoes per person per day reflect shifts in mosquito density, not in the proportion of infectious mosquitoes. The model predicts that human biting rate is highest shortly after mosquito population density peaks, typically either near breeding sites or where human density is highest. The proportion of infectious mosquitoes, on the other hand, reflect the age of the mosquito population: it peaks where older mosquitoes are found—farther from breeding sites—and when populations are declining.

By mapping larval habitats against the local risk of mosquito-borne infections, Smith and colleagues conclude, epidemiological models can be developed to predict risk for local populations. Their results make the case that mathematical models can help public health officials calculate risk of infectious diseases in heterogeneous environments—that is, real world conditions—when vector ecology and the parameters of transmission are well characterized. Any plan to prevent and control the spread of mosquito-born infections would clearly benefit from paying attention to mosquito demography and behavior.

Smith DL, Dushoff J, McKenzie FE (2004) The risk of a mosquito-borne infection in a heterogeneous environment. DOI: 10.1371/journal.pbio.0020368

Endangered Frogs Coexist with Fungus Once Thought Fatal

DOI: 10.1371/journal.pbio.0020389

Amphibian declines have reached crisis proportions in various parts of the world. In many areas, habitat loss is the likely culprit. But when mass die-offs suddenly occurred in relatively undisturbed habitats, the cause was far less obvious. Fourteen species suffered either extinctions or major declines in the pristine rainforests of Queensland, Australia, between 1979 and 1993. It was suggested in 1996 that some unknown disease had spread through the populations, but no pathogen was discovered until 1998, when the fungus *Batrachochytrium dendrobatidis* was identified from sick and dead frogs. Since then, several lines of evidence suggest that *B. dendrobatidis* may be involved in frog declines: the fungus has been found on frogs in afflicted areas; lab studies show that it's highly pathogenic to some frog species; and pathological evidence links it to host mortality. But with little information about the prevalence of this fungal infection in wild frogs, or how the disease impacts frogs in the wild, the causal role of this chytrid fungus remains unclear.

DOI: 10.1371/journal.pbio.0020389.g001

Taudactylus eungellensis **(Photo: Richard Retallick)**

To evaluate the effects of *B. dendrobatidis* on frogs in their natural habitat, Richard Retallick et al. focused on six species living in the high-elevation rainforest streams of Eungella National Park in Queensland, Australia, where frog losses were "particularly catastrophic." Two species vanished between 1985 and 1986: the Eungella Gastric-Brooding Frog (*Rheobatrachus vitellinus*), which is now thought extinct, and the Eungella Torrent Frog (*Taudactylus eungellensis*), which later reappeared in a few small populations. Other local frog species escaped this period relatively unscathed.

Retallick captured frogs from six sites from 1994 to 1998, clipped one or two toe tips from each frog to age and identify them, and then released the frogs back into the wild. At the time, *B. dendrobatidis* had yet to be identified, but Retallick retained the toe tips, and the authors tested the toes for disease in 2002–2003. Fungal infections were found in two species—*T. eungellensis* and *Litoria wilcoxii/jungguy* (the latter consists of two species that are indistinguishable without genetic analysis); the other four species were infection-free. *L. wilcoxii/jungguy* did not decline to any great extent during the 1985–1986 die-off.

The proportion of infected *T. eungellensis* frogs was greatest at three particular sites, which showed peak infections during cooler months. Prevalence of infection was highest during winter and spring, but did not vary from year to year, suggesting that the infection is now endemic. Fungal infections were found in 27.7% of *L. wilcoxii/jungguy* frogs, with no evidence that prevalence differed among sites, seasons, or individuals (males, females, or subadults). The probability of recapture was significantly lower for frogs that were already infected when first captured. While this might suggest a correlation between infection and death, it's impossible to distinguish death from simple failure to recapture the animal. On further analysis, McCallum and colleagues found no evidence that survival differed between infected and uninfected frogs, suggesting that this potentially devastating amphibian disease now coexists with the frogs, with little effect on their populations.

These results, the authors conclude, "show unequivocally" that remaining populations of *T. eungellensis*, a rainforest frog listed as endangered, "now persist with stable infections of *B. dendrobatidis*." While these findings do not exonerate the fungus as the agent of mass declines, they do rule out the possibility that the fungus caused the decline, then vanished from the area, allowing frog populations to recover. The authors allow that it's possible that *B. dendrobatidis* did not cause the

initial *T. eungellensis* declines. Or alternately, the fungus could have emerged as a novel pathogen in the ecosystem, causing massive casualties before some form of evolutionary response took hold. Surviving frog populations may have evolved resistance to the pathogen, for example, or less virulent strains of the fungus may have evolved. If it turns out that frog populations can develop resistance to the chytrid fungus, the researchers point out, then a conservation program of captive breeding and selecting for resistance could potentially thwart the extinction of these, and other, critically endangered frogs. A critical next step, then, is to determine whether frogs and fungus do coevolve.

Retallick RWR, McCallum H, Speare R (2004) Endemic infection of the amphibian chytrid fungus in a frog community post-decline. DOI: 10.1371/journal.pbio.0020351

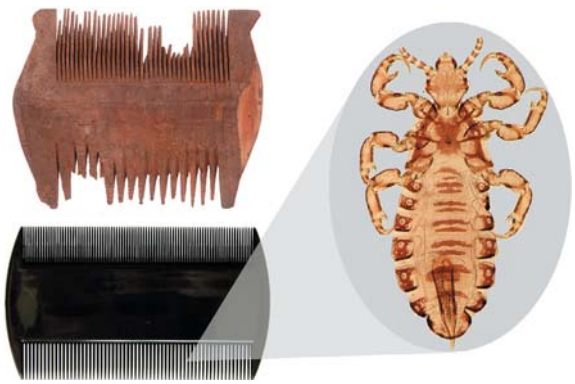
Did We or Didn't We? Louse Genetic Analysis Says Yes

DOI: 10.1371/journal.pbio.0020378

If you're an evolutionary biologist, the tired old saw, "You can tell a lot about a person by the company they keep," represents a fresh new approach to a longstanding problem. Especially if the company in question is a parasite—say, for example, lice—and the problem is tracing the path of human evolution.

One of the hottest debates in the study of human origins centers around whether modern *Homo sapiens* interbred with archaic humans. While genetic data have provided insight into recent human evolution, fossils remain the only available evidence for many archaic human species—and the fossil record is notoriously spotty, leaving the data open to multiple interpretations. Two prominent models and a subset of variants have emerged, differing mainly on the question of gene flow: one asserts that modern humans emerged from an archaic ancestor in Africa about 130,000 years ago and then replaced archaic forms in Africa, Asia, and Europe with no gene flow between them; the other proposes gene flow between modern human populations as well as interbreeding between modern and archaic forms in different parts of the world. Both models find support in the available data, but neither can claim a perfect fit with all the data, leaving the possibility of interbreeding an open question.

Faced with a relative paucity of human fossil and genetic data, scientists have been forced to rely on other data sources. Mounting evidence suggests that parasites with an established coevolutionary history with their hosts can serve as a proxy for host evolutionary history, an especially handy tool in the event of insufficient host data. Following this approach in a new study, David Reed and colleagues circumvent the lack of human data by analyzing the next best thing: head lice.



DOI: 10.1371/journal.pbio.0020378.g001

Ancient nit combs (above) resemble modern ones (below). (Egyptian wooden comb courtesy of Te Papa, Wellington, New Zealand, negative number F.003884/5. Modern louse comb and head louse images by Vincent S. Smith.)

As host-specific, obligate parasites—that is, occurring on a single species and not able to survive off that host—lice require direct physical contact between hosts for transmission. As human parasites, lice harbor in their genetic sequence hints of the slings and arrows of evolutionary fortune (and touches of grace, for that matter) that strike their host. Recent studies of the evolutionary history of other human parasites (tapeworms, malarial parasites, and human papillomaviruses), for example, fall in line with fossil and genetic data that place our origins in Africa.

Two louse species parasitize humans, head/body *(Pediculus humanus)* and pubic *(Pthirus pubis)*. Head and body lice obviously occupy different habitats, but are not genetically distinct. Interestingly, *P. humanus* contains two ancient lineages, offering the opportunity to shed light on this murky period in human evolution. To do this, Reed and colleagues had to reconstruct the evolutionary history of *P. humanus,* which they did using morphological and genetic analyses of this and other species of lice. They confirmed that *P. humanus* comprises two lineages—one contains both head and body forms and has worldwide distribution; the other contains only the head louse and is restricted to the New World—but discovered that *P. humanus* originated long before its *H. sapiens* host.

Humans went through a population bottleneck around 100,000 years ago, followed by an expansion; one would expect to see the same thing in lice. Population genetics studies revealed, however, that only the worldwide lineage went through a bottleneck and subsequent expansion. The New World lineage not only maintained a relatively stable population size but followed an evolutionary path distinct from the worldwide lineage for the past 1.18 million years. It is unlikely, the authors argue, that two ancient louse lineages could embark on such different evolutionary histories on the back (or head) of a single host. More likely, the New World louse evolved on an archaic form of humans and then cast its lot with a modern version.

While the split between *H. sapiens* and *H. neanderthalensis* was too recent (about 700,000 years ago) to support a concurrent split between the worldwide and New World lice lineages, the split between *H. sapiens* and *H. erectus* (about 1.8 million years ago) could. Reed and colleagues propose a scenario in which *H. sapiens* and *H. erectus* carried distinct types of lice owing to a million years or so of isolation. As the first waves of modern humans left Africa about 100,000 years ago and modern humans replaced archaic forms, the two forms engaged in enough contact—whether in the form of fighting, swapping clothes, or interbreeding—for archaic lice to make the switch to modern human hosts. Tackling the question of interbreeding, the authors suggest, might best be pursued by studying *P. pubis,* which requires sexual contact for transmission.

Reed DL, Smith VS, Hammond SL, Rogers AR, Clayton DH (2004) Genetic analysis of lice supports direct contact between modern and archaic humans. DOI: 10.1371/journal.pbio.0020340

A Richer Map for Searching Scientific Literature

DOI: 10.1371/journal.pbio.0020343

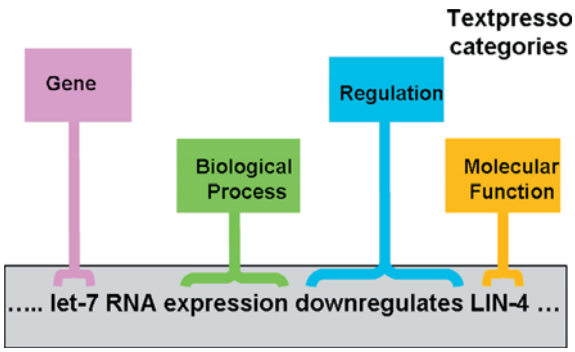
Anyone who has used a search engine quickly becomes familiar with both their power and limitations. A key-word search for "bush kerry butterfly county" turns up almost twenty-eight thousand documents, but a scant few of them are of any interest to the botanist studying the butterfly bush in County Kerry. The problem is more circumscribed but no less significant within specialized databases, such as the fourteen-plus million medical journal articles catalogued by PubMed.

It is not so much that the literature is too vast, but that the search strategies are too weak. A simple list of key words used to tag and retrieve a document cannot begin to capture the richness of the information within, especially when that wealth is expressed in syntactically complex sentences like "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."

But there is another way. Rather than simply extracting a limited list of key words from an article's abstract, the entire text of the document can be categorized into classes: some words represent entities (e.g., gene, cellular component, molecular function) and others relationships (e.g., physical association, purpose, comparison, regulation). The entire set of entities and relationships can be linked to create a map of the information within the document, which, like a physical map, captures some of the complexity of the territory it describes.

Humans excel at this type of concept mapping, but their labors are slow and expensive. In this issue, Paul Sternberg and colleagues at the California Institute of Technology (Caltech) describe a computer-based system that performs the same task, and show that it is almost as good as humans at mapping out the scientific literature concerning the laboratory nematode, *Caenorhabditis elegans*.

Sternberg's system, called Textpresso, includes 33 categories of terms, both of entities and relationships, and a full list of all possible examples for each entity (for genes, for example, this would be specific gene names) and relationship (for physical association, this would include bind, adhere, link, etc.). This collection, called an ontology, is then applied to sentences within the text of a document to map out the relationships within—for instance, the mention of two genes within a sentence along with any form of the word stem "regulat-" indicates that one gene probably regulates the other, and the sentence is marked accordingly. With scores of tags applied, the full markup of a sentence is typically much longer than the sentence itself. Currently, Textpresso has marked up almost 4,000 full-text articles on *C. elegans*, representing 60% of the entire literature.



DOI: 10.1371/journal.pbio.0020343.g001

Identifying terms in scientific prose

The final result is a document that can be queried in subtle ways impossible with mere key words. For instance, to find entities (whether they be transcription factors, small molecules, or anything else described to date) that interact with the aging-related gene *daf-16*, one enters the terms "daf-16" and "association." Textpresso returns 125 publications, with citations and links to the articles. (Textpresso is available at www.textpresso.org.) The results can be further refined by adding other entities or relationships, as well as by specifying author, journal, or year of publication.

Textpresso's ability to find relevant documents, and ignore irrelevant ones, is still not as great as an expert human curator of the same literature. But the system can be constantly tweaked to get better and better. This process requires human intervention, and the Caltech team does not think this is likely to be automated anytime soon. On the other hand, the structure of Textpresso, and to some extent the ontological lists from *C. elegans*, can be used for literature analysis of other model organisms. Finally, the fully annotated literature within a field is not only a repository of scientific facts, but also a data mine of human communication, which can be queried for patterns having little to do with model organisms and much to do with how scientists communicate with each other.

Müller HM, Kenny EE, Sternberg PW (2004) Textpresso: An ontology-based information retrieval and extraction system for biological literature. DOI: 10.1371/journal.pbio.0020309