Science and Medicine

Ze'ev Hochberg and Ken Ong

Skewed inheritance

The architecture of parent-of-origin effects in mice

Mott R, Yuan W, Kaisaki P, Gan X, Cleak J, Edwards A, Baud A, Flint J
Wellcome Trust Centre for Human Genetics, Oxford, UK
Cell 2014;156:332–342

Background: The number of imprinted genes in the mammalian genome is predicted to be small, yet we show here, in a survey of 97 traits measured in outbred mice, that most phenotypes display parent-of-origin effects that are partially confounded with family structure.

Methods: To address this contradiction, the authors used reciprocal F1 crosses to investigate the effects of knocking out two non-imprinted candidate genes, *Man1a2* and *H2-ab1*, that reside at non-imprinted loci but that show parent-of-origin effects.

Results: Expression of multiple genes becomes dysregulated in a sex-, tissue-, and parent-of-origin-dependent manner. Evidence is shown that non-imprinted genes can generate parent-of-origin effects by interaction with imprinted loci.

Conclusions: The importance of the number of imprinted genes is secondary to their interactions. This gene network effect may account for some of the missing heritability seen when comparing sibling-based to population-based studies of the phenotypic effects of genetic variants.

Parent-of-origin effects are those in which the phenotypic manifestations of a gene are determined by the identity of its origin, the mother or the father. Typically, such effects were thought to result from genomic imprinting, DNA methylation marks that silence genes when transmitted specifically from mother or father. Less than 100 of the estimated 20,000 human genes are imprinted and, despite some classical examples of monogenic imprinting disorders (e.g. the Prader-Willi, Angelman, Beckwith-Wiedemann and Silver-Russell syndromes), there are yet few examples of parent-of-origin allelic effects on complex human traits and diseases [1].

Here, the authors show, by computational re-analysis of published data in mice, that a surprisingly large number of traits show parent-of-origin effects. Remarkably, 91 out of 97 (93%) of the traits studied showed such inheritance patterns. They then identified multiple individual genetic variants that show parent-of-origin specific associations with these traits. Surprisingly, many of these variants were located in non-imprinted genes. They went on to show a remarkable network of interactions between imprinted and non-imprinted genes that contributes to parent-of-origin effects.

There has been increasing awareness that imprinted genes are co-regulated in wide cross-chromosomal networks through epigenetic mechanisms beyond DNA methylation, for example chromatin methylation and long-range regulatory RNAs [2]. This paper shows that such networks likely also involve many non-imprinted genes. Consideration of parent-of-origin effects, even at non-imprinted genes, could explain part of the 'missing heritability' in many complex human traits and diseases.

Global epigenomic reconfiguration during mammalian brain development

Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghghi FG, Sejnowski TJ, Behrens MM, Ecker JR
Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA
Science 2013;341:1237905

Background: DNA methylation is implicated in mammalian brain development and plasticity underlying learning and memory.

Methods: The authors performed genome-wide composition, patterning, cell specificity, and dynamics of DNA methylation at single-base resolution in human and mouse frontal cortex throughout their lifespan.

Results: Widespread methylome reconfiguration occurs during fetal to young adult development, coincident with synaptogenesis. Highly conserved non-CG methylation (mCH) accumulates with development in neurons, but not glia, and is the dominant form of methylation in human neuronal DNA. A specific mCH signature identifies genes that escape X-chromosome inactivation. Finally, 5-hydroxymethylcytosine (hmC) marks fetal brain cell genomes at putative regulatory regions that are CG-demethylated and activated in the adult brain and that CG demethylation at these hmC-poised loci depends on Tet2 activity.

Conclusions: Whole-genome single-base resolution methylcytosine and hydroxymethylcytosine maps reveal profound changes during frontal cortex development in humans and mice. Overall, brain cell DNA methylation has unique features that are precisely conserved, yet dynamic and cell-type specific.

Changes in brain DNA methylation have been implicated in learning, memory, and age-related cognitive decline. For example, disruption of the more typically-studied gene-repressive methylation of CpG sites (5-methylcytosine) leads to cognitive defects in the mouse. Here, the authors studied at single base-pair resolution, the changes with age and developmental stage in various epigenetic marks. These included 5-hydroxymethylcytosine, which is oxidized from 5-methylcytosine, and is generally correlated with increased (rather than decreased) gene expression, and also methylation of non-CpG sites ('mCH', where H = A, C, or T). Unlike other differentiated cell types, they found that neurons accumulate substantial mCH during the early years of life, coinciding with the period of synaptogenesis and brain maturation.

The brain continues to develop physically, in humans, well into the third decade of life. An early post-natal burst of synaptogenesis is followed by pruning of excess synapses during adolescence, allowing experience-dependent plasticity and learning in children and young adults. Disruption of this process leads to behavioral and psychiatric disorders. As the authors conclude, these results extend our knowledge of the unique role of DNA methylation in brain development and function, and offer a new framework for testing of pathological disorders of neurodevelopment.

A new appetite-controlling circuit

Genetic identification of a neural circuit that suppresses appetite

Carter ME, Soden ME, Zweifel LS, Palmiter RD
Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA
Nature 2013;503:111–114

Background: Appetite suppression occurs after a meal and in conditions when it is unfavorable to eat, such as during illness or exposure to toxins. A brain region proposed to play a role in appetite suppression is the parabrachial nucleus, a heterogeneous population of neurons surrounding the superior cerebellar peduncle in the brainstem. The parabrachial nucleus is thought to mediate the suppression of appetite induced by the anorectic hormones amylin and cholecystokinin, as well as by lithium chloride

and lipopolysaccharide, compounds that mimic the effects of toxic foods and bacterial infections, respectively. However, the identities of neurons in the parabrachial nucleus that regulate feeding are unknown, as are the functionally relevant downstream projections.

Methods and Results: The authors used genetically encoded anatomical, optogenetic and pharmacogenetic tools to identify calcitonin gene-related peptide-expressing neurons in the outer external lateral subdivision of the parabrachial nucleus that project to the laterocapsular division of the central nucleus of the amygdala as forming a functionally important circuit for suppressing appetite. Activation of these neurons projecting to the central nucleus of the amygdala suppresses appetite. In contrast, inhibition of these neurons increases food intake in circumstances when mice do not normally eat and prevents starvation in adult mice whose agouti-related peptide neurons are ablated.

Conclusions: These findings identify a neural circuit from the parabrachial nucleus to the central nucleus of the amygdala that mediates appetite suppression in conditions when it is unfavorable to eat. This neural circuit may provide targets for therapeutic intervention to overcome or promote appetite.

You are probably familiar with papers on the roles of leptin and MC4R in hunger-promoting GABAergic AgRP neurons located in the arcuate nucleus of the hypothalamus. That hypothalamic circuit maintains homeostatic control of energy status, by responding to peripheral and central signals of energy depletion or energy repletion to promote hunger or satiety, respectively. Here, the authors explored the basis for another circuit that regulates eating behavior, namely what turns you off your food when you are ill, as in infectious or inflammatory conditions, or when exposed to toxins. They found a neural circuit connecting the hindbrain to the amygdala that when stimulated leads to illness-associated anorexia. This circuit lies downstream of those other hypothalamic and brainstem nuclei that coordinate appetite and satiety, and its activation is likely responsible for the unpleasant feeling or discomfort experienced in adverse conditions when it is unfavorable to eat. Therapeutic regulation of this circuit might help to promote re-establishment of appetite in debilitating chronic illness, while its activation might represent new treatment targets for obesity.

A Neanderthal T2DM gene

Sequence variants in *SLC16A11* are a common risk factor for type 2 diabetes in Mexico

SIGMA Type 2 Diabetes Consortium, Williams AL, Jacobs SB, Moreno-Macias H, Huerta-Chagoya A, Churchhouse C, Marquez-Luna C, Garcia-Ortiz H, Gomez-Vazquez MJ, Burr T NP, Aguilar-Salinas CA, Gonzalez-Villalpando C, Florez JC, Orozco L, Haiman CA, Tusie-Luna T, Altshuler D
Nature 2014;506:97–101

Background: Performing genetic studies in multiple human populations can identify disease risk alleles that are common in one population but rare in others, with the potential to illuminate pathophysiology, health disparities, and the population genetic origins of disease alleles.

Methods: The authors analyzed 9.2 million single nucleotide polymorphisms (SNPs) in each of 8,214 Mexicans and other Latin Americans: 3,848 with type 2 diabetes (T2DM) and 4,366 non-diabetic controls.

Results: In addition to replicating several previous findings, a novel locus was found associated with T2DM at genome-wide significance spanning the solute carriers genes *SLC16A11* and *SLC16A13* ($P = 3.9 \times 10^{-13}$; odds ratio (OR) = 1.29). The association was stronger in younger, leaner people with T2DM, and replicated in independent samples ($P = 1.1 \times 10^{-4}$; OR = 1.20). The risk haplotype carries four amino acid substitutions, all in *SLC16A11*; it is present at ~50% frequency in Native American samples and ~10% in East Asian, but is rare in European and African samples. *SLC16A11* mRNA is expressed in liver, and its protein localizes to the endoplasmic reticulum. Analysis of an archaic genome sequence indicated that the risk haplotype introgressed into modern humans via admixture with Neanderthals. Expression of *SLC16A11* in heterologous cells increased intracellular triacylglycerol levels.

Conclusions: Study of Mexican and Latin American individuals identified *SLC16A11* as a novel candidate gene for T2D with a possible role in triacylglycerol metabolism.

What is special about one more T2DM locus? Indeed, there are 76 other T2DM loci found in much larger genome-wide association studies (GWAS), the latest study involving 26,488 cases and 83,964 controls across a wide range of ancestries [3]. This GWAS identified a new risk factor for T2DM in Mexicans and Mexican Americans: common variants in the gene *SLC16A11*.

The novelty of this study is that it demonstrates the value of investigating genetic determinants of disease in specific ancestral populations. Some populations harbor 'private' variants that are not found in other samples. These may contribute to population differences in disease prevalence, and may open new avenues in understanding of disease pathogenesis. The prevalence of T2DM in Mexican and other Latin American populations is double that seen in US non-Hispanic Whites. The authors estimate that this single haplotype, which is rare or absent in populations from Africa and Europe, could on its own account for 20% (9.2–29%) of the difference in prevalence of T2DM between Mexican Americans and European Americans. Furthermore, this gene identifies a new pathway of intracellular triacylglycerol metabolism in the liver that contributes to T2DM development.

Finally, multiple lines of evidence indicate that the T2DM-predisposing haplotype, which is not present in Africans, entered modern humans relatively recently through admixture between humans and Neanderthals. Remarkably, those individuals who are homozygous for the 5 SNP haplotype have a near-identical 73-kb long DNA sequence to that of a Neanderthal from Denisova Cave, a cave in the Altai Mountains of Siberia, in a region thought to have been inhabited concurrently by Neanderthals and modern humans. This is a remarkable example of genetic introgression, by admixture between modern humans and Neanderthal, affecting physiology and disease susceptibility.

***FTO* or not *FTO* – that is the question**

Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*

Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gomez-Marin C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviondran V, Tam D, Shen M, Son JE, Vakili NA, Sung HK, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui CC, Gomez-Skarmeta JL, Nobrega MA
Department of Human Genetics, University of Chicago, Chicago, IL, USA
Nature 2014;507:371–375

Background: Genome-wide association studies (GWAS) identified variants within introns of *FTO* that are robustly-associated with risk for obesity and type 2 diabetes (T2DM). Subsequent studies in mice showed that *FTO* expression levels influence weight gain and body composition. However, no direct link between the obesity-associated variants and *FTO* expression or function has been shown.

Methods and Results: The authors show that the obesity-associated non-coding sequences within *FTO* are functionally connected with the homeobox gene *IRX3*. The obesity-associated *FTO* region directly interacts with the promoters of *IRX3* as well as *FTO* in the human, mouse and zebrafish genomes. Furthermore, long-range enhancers within this region recapitulate aspects of *IRX3* expression, suggesting that the obesity-associated interval regulates *IRX3*. In human brains, obesity-associated variants are associated with expression of *IRX3*, but not *FTO*. *Irx3*-deficient mice show a 25–30% reduction in body weight, primarily through the loss of fat mass and increase in basal metabolic rate with browning of white adipose tissue. Hypothalamic expression of a dominant-negative form of *Irx3* reproduces the metabolic phenotypes of *Irx3*-deficient mice.

Conclusions: These findings suggest that *IRX3* is a functional long-range target of obesity-associated variants within *FTO* and represents a novel determinant of body mass and composition.

While many other genetic loci have since been linked to obesity risk, common variants in introns 1 and 2 of the *FTO* gene still represent the strongest genetic association signals for polygenic obesity. However, how it regulates obesity risk has proven to be a difficult challenge [4]. Alarming, the current paper suggests we have been looking in the wrong place! These authors investigated long-range DNA interactions using chromatin conformation capture, which measures how frequently different regions of the genome interact. They found that, despite being located nearly half a million

base-pairs apart, intron 1 of the *FTO* gene physically interacts with the promoter of the *IRX3* gene. They go on to show that experimental disruption of whole-body and specifically hypothalamic *IRX3* expression alters weight gain and energy homeostasis.

Long-range interaction between DNA sequences has long been recognized. For example, an intron in the gene *HERC2*, which itself regulates neurodevelopment, contains an enhancer for *OCA2*, an oculocutaneous albinism gene, and this explains how variants in *HERC2* are associated with eye, hair and skin color. Increasing understanding of such long-range DNA interactions will hopefully help to understand the mechanisms of many disease-associated genetic loci, and in particular some remarkably strong signals that occur in gene deserts – regions of the genome that are completely devoid of coding sequences.

A new hormone links metabolism and cancer

Endotrophin triggers adipose tissue fibrosis and metabolic dysfunction

Sun K, Park J, Gupta OT, Holland WL, Auerbach P, Zhang N, Goncalves Marangoni R, Nicoloso SM, Czech MP, Varga J, Ploug T, An Z, Scherer PE

Touchstone Diabetes Center, Department of Internal Medicine, University of Texas Southwestern Medical Center, TX, USA

Nat Commun 2014;5:3485

Background: The authors recently identified endotrophin as an adipokine with potent tumor-promoting effects. However, the direct effects of local accumulation of endotrophin in adipose tissue have not yet been studied.

Methods: The authors used a doxycycline-inducible adipocyte-specific endotrophin overexpression model to demonstrate that endotrophin plays a pivotal role in shaping a metabolically unfavorable microenvironment in adipose tissue during consumption of a high-fat diet (HFD).

Results: Endotrophin is a potent co-stimulator of pathologically relevant pathways within the ‘unhealthy’ adipose tissue milieu, triggering fibrosis and inflammation and ultimately leading to insulin resistance. Blocking endotrophin action using a neutralizing antibody ameliorates the metabolically adverse effects and reverses the metabolic dysfunction induced by HFD.

Conclusions: These findings demonstrate that endotrophin has a key role in adipose tissue, resulting in systemic elevation of pro-inflammatory cytokines and insulin resistance. Endotrophin is a potential mechanism in metabolism and cancer.

Excess caloric balance causes rapid expansion of adipose tissue, which leads to local hypoxia and a fibrotic response. Adipose tissue fibrosis then restrains further adipocyte expansion and leads to an increased adipocyte cell death and enhanced inflammatory responses. The authors had previously identified ‘endotrophin’ as the cleaved product of a pro-fibrotic protein, and showed that it has distal effects on mammary tumor expansion and metastatic growth [5]. Here, they examined its local effects in adipose tissue and found that it has a similarly adverse effect on metabolic health. Endotrophin is upregulated in genetic and diet models of mouse obesity. In turn, endotrophin overexpression upregulates pro-fibrotic and pro-inflammatory genes in fat pads of the transgenic mice, eventually leading to insulin resistance. Furthermore, a neutralizing antibody against endotrophin ameliorates the metabolically adverse effects induced by a high-fat diet.

In summary, endotrophin represents a new mechanism that potentially explains links between high levels of adiposity to both poor metabolic health and increased cancer risk. It can function either locally in adipose tissue to trigger fibrosis, inflammation and unfavorable metabolic effects as reported here, or more distally in tumor tissues to promote tumor progression. Blocking endotrophin might serve as a future therapeutic target.

A secreted PTEN phosphatase that enters cells to alter signaling and survival

Hopkins BD, Fine B, Steinbach N, Dendy M, Rapp Z, Shaw J, Pappas K, Yu JS, Hodakoski C, Mense S, Klein J, Pegno S, Sulis ML, Goldstein H, Amendolara B, Lei L, Maurer M, Bruce J, Canoll P, Hibshoosh H, Parsons R
Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA
Science 2013;341:399–402

Background: Phosphatase and tensin homolog on chromosome ten (PTEN) inhibits cell growth and proliferation by antagonizing the phosphoinositide 3-kinase (PI3K) pathway.

Methods and Results: By analysis of *PTEN* mRNA, the authors identified a novel 576-amino-acid variant of *PTEN*, which they named PTEN-Long as it contains an additional 173 N-terminal amino acids compared to the normal PTEN open reading frame. This variant arises from an alternative translation start site 519 base-pairs upstream of the ATG initiation sequence. PTEN-Long is a membrane-permeable lipid phosphatase that is secreted from cells and can enter other cells. When administered exogenously, PTEN-Long antagonizes PI3K signaling and induces tumor cell death in vitro and in vivo.

Conclusions: PTEN-Long may have therapeutic uses as a functional tumor-suppressor protein in tumor cells.

Our previous *Yearbook* chapters described rare mutations in the PI3-kinase/AKT/PTEN pathway as causing constitutive overactivity of this signalling pathway leading to overgrowth, severe hypoglycemia, and linking tumor susceptibility to insulin sensitivity [6]. As reviewed above in chapter 2, rapamycin and other inhibitors of downstream mTOR signalling are proving to be effective in many of these conditions [7]. The, quite unexpected, findings reported here point to other ways to modulate this key pathway.

As expected of an intracellular signalling molecule, PTEN is normally found in the cytosol and sometimes the nucleus. Here, the authors found an alternative initiation codon ~500 bp upstream of the typical initiation codon, which translates to a longer variant of PTEN protein. Crucially, the additional 173 amino acids contain a secretion signal sequence (which allows PTEN-Long to exit cells) and also a string of arginine residues, which resembles the cell-penetrating domain of the transactivator of transcription (TAT) protein of the human immunodeficiency virus and allows PTEN-Long to enter other cells. They found that PTEN-Long is secreted from cells and can decrease PI3K signalling distally. Indeed, intraperitoneal injection of PTEN-Long into mice carrying xenograft tumors from several human cancer cell lines reduced PI3K-Akt and led to tumor regression. The relevance of PTEN-Long on insulin metabolism was not yet examined, but would be expected to reduce insulin sensitivity. In addition to suggesting new therapeutic targets in cancer, synthesis of derivative fusion proteins may allow cell delivery of proteins in many other experimental and clinical settings.

Fix your eyes on autism

Attention to eyes is present but in decline in 2- to 6-month-old infants later diagnosed with autism

Jones W, Klin A
Marcus Autism Center, Children's Healthcare of Atlanta, Department of Pediatrics, and Center for Translational Social Neuroscience, Emory University, Atlanta, GA, USA
Nature 2013;504:427–431

Background: Deficits in eye contact are a well-recognized hallmark of autism spectrum disorders (ASDs). However, the age at onset of these deficits has not been known.

Methods: The authors performed a prospective longitudinal study of 110 infants: 59 at high risk for ASD (full siblings of a child with ASD) and 55 at low risk (without affected first-, second- or third-degree relatives). While infants were shown a video of a woman, an eye-tracker was used to measure 'prefer-

ential attention to the eyes', as opposed to other body parts or objects, at 10 time points: ages 2, 3, 4, 5, 6, 9, 12, 15, 18 and 24 months. Diagnostic status of ASD was ascertained at 36 months.

Results: In the 12 infants who were later diagnosed with ASDs, growth curves of social visual engagement followed a different developmental course compared to other infants; they showed a mean decline in eye fixation from 2 to 6 months of age, a pattern not observed in infants who did not develop ASD. These observations mark the earliest known indicators of social disability in infancy. They also falsify a prior hypothesis: in the first months of life, this basic mechanism of social adaptive action, eye fixation, is not immediately defective in infants later diagnosed with ASD, but starts at normative levels prior to decline.

Conclusions: The observation of this decline in eye fixation, rather than outright absence, offers a promising opportunity for early intervention that could build on the apparent preservation of mechanisms subserving reflexive initial orientation towards the eyes.

You don't need to be a pediatrician to know that infants look more at another person's eyes than at their mouth and other body or object regions. This eye fixation is thought to be an innate skill that guides normal infant socialization; it is already apparent at age 2 months and remains constant during the first 2 years of life. In contrast, mouth fixation typically increases during the first year, peaking at approximately 18 months, while fixation on other body and object regions declines sharply throughout the first year.

Impaired eye fixation has long been recognized as a key feature of autism spectrum disorders (ASDs), is tested for in clinical diagnostic instruments, and deficits have been noted as early as 2 years of age. This leads to subsequent impairments in social orienting; affected individuals are also poor at extracting complex information from faces, such as emotions and intention. Here, the authors used state-of-the-art eye-tracking methods in a very detailed and long-running study to document the derailment in visual engagement in children who were later diagnosed with ASDs. The authors honestly admit that the findings disproved their (and most other investigators') hypothesis that defective eye fixation would be diminished from birth, thus contradicting the prevailing idea that ASD is caused by a 'congenital absence of social adaptive orientation'. As well as allowing possible much earlier diagnosis of ASDs, the findings indicate a remarkable potential opportunity for treatment during a narrow developmental period, spanning the transition from experience-expectant to experience-dependent mechanisms, and before atypical processing embeds itself as a life-long condition.

Chemical synthesis – two papers

Total synthesis of a functional designer eukaryotic chromosome

Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM, Cai Y, Zeller K, Agmon N, Han JS, Hadjithomas M, Tullman J, Caravelli K, Cirelli K, Guo Z, London V, Yeluru A, Murugan S, Kandavelou K, Agier N, Fischer G, Yang K, Martin JA, Bilgel M, Bohutski P, Boulier KM, Capaldo BJ, Chang J, Charoen K, Choi WJ, Deng P, DiCarlo JE, Doong J, Dunn J, Feinberg JJ, Fernandez C, Floria CE, Gladowski D, Hadidi P, Ishizuka I, Jabbari J, Lau CY, Lee PA, Li S, Lin D, Linder ME, Ling J, Liu J, London M, Ma H, Mao J, McDade JE, McMillan A, Moore AM, Oh WC, Ouyang Y, Patel R, Paul M, Paulsen LC, Qiu J, Rhee A, Rubashkin MG, Soh IY, Sotuyo NE, Srinivas V, Suarez A, Wong A, Wong R, Xie WR, Xu Y, Yu AT, Koszul R, Bader JS, Boeke JD, Chandrasegaran S

Department of Environmental Health Sciences, Johns Hopkins University School of Public Health, Baltimore, MD, USA
Science 2014;344:55–58

Context: Rapid advances in DNA synthesis techniques have made it possible to engineer viruses, biochemical pathways and assemble bacterial genomes.

Methods and Results: Here, the authors report the synthesis of a functional 272,871-base-pair designer eukaryotic chromosome, synIII, which is based on the 316,617-base-pair native *Saccharomyces cerevisiae* chromosome III. Changes to synIII include TAG/TAA stop-codon replacements, deletion of subtelomeric regions, introns, transfer RNAs, transposons, and silent mating loci as well as insertion of loxP sites to enable genome scrambling. SynIII is functional in *S. cerevisiae*. Scrambling of the chromosome in a heterozygous diploid reveals a large increase in a-mater derivatives resulting from loss of the MAT α allele on synIII.

Conclusion: The complete design and synthesis of synIII establishes *S. cerevisiae* as the basis for designer eukaryotic genome biology.

The design and generation of a functional, synthetic chromosome in a yeast is a milestone in engineering vital cells. To generate the synthetic chromosome, the designer DNA sequence was broken down into overlapping stretches of 70 nucleotides, which were chemically synthesized in parallel. The original chromosome sequence was then systematically replaced with the synthetic DNA in vivo, by introducing up to 12 overlapping 3-kb synthetic fragments at a time into the yeast cell, in 11 successive rounds of integration. They now have a yeast strain that contains an engineered chromosome III, and that grows just as well as the original strain. It contains more than 50,000 sequence alterations, and is 14% shorter than the natural sequence. This work addresses several issues that will be important in eukaryotic genome design – the removal of redundant DNA between genes that perform similar functions, altering the genetic code.

Erythropoietin derived by chemical synthesis

Wang P, Dong S, Shieh JH, Peguero E, Hendrickson R, Moore MA, Danishefsky SJ
Laboratory for Bioorganic Chemistry, Sloan-Kettering Institute for Cancer Research, New York, NY, USA
Science 2013;342:1357–1360

Context: Erythropoietin is a signaling glycoprotein that controls the fundamental process of erythropoiesis, orchestrating the production and maintenance of red blood cells. As administered clinically, erythropoietin has a polypeptide backbone with complex heterogeneity in its carbohydrate domains.

Results: The authors describe the total synthesis of homogeneous erythropoietin with consensus carbohydrate domains incorporated at all of the native glycosylation sites. The oligosaccharide sectors were built by total synthesis and attached stereospecifically to peptidyl fragments of the wild-type primary sequence, themselves obtained by solid-phase peptide synthesis. The glycopeptidyl constructs were joined by chemical ligation, followed by metal-free dethiylation, and subsequently folded. This homogeneous erythropoietin glycosylated at the three wild-type aspartates with N-linked high-mannose sialic acid-containing oligosaccharides and O-linked glycoporphin exhibits Procrit-level in vivo activity in mice.

Most biologic therapeutics, such as insulin and growth hormone, are large, complex molecules or heterogeneous mixtures of molecules; as such they are manufactured through recombinant DNA technology in living system (microorganism, plant cell, or animal cell). Biologic therapeutics are now being used to treat a wide range of diseases, including cancer and autoimmune disorders, while their utility in endocrine conditions has been recognized now for three decades. Here, the authors accomplished the total chemical synthesis of the biologic erythropoietin (EPO) in a single, pure form. It was made possible by development of native chemical ligation (NCL), a selective reaction that enables two unprotected polypeptide chains to be joined. The reaction requires a cysteine residue which EPO (and other proteins) lack, but the authors overcame this obstacle by showing that alanine (far more prevalent than cysteine) can do the job.

Bigger brains in the city

Anthropogenic environments exert variable selection on cranial capacity in mammals

Snell-Rood EC, Wick N
Department of Ecology, Evolution and Behavior, University of Minnesota, Twin Cities, MN, USA
Proc Biol Sci 2013;280:20131384

Context: It is thought that behaviorally flexible species will be able to cope with novel and rapidly changing environments associated with human activity. However, it is unclear whether such environments are selecting for increases in behavioral plasticity, and whether some species show more pronounced evolutionary changes in plasticity.

Methods: To test whether anthropogenic environments are selecting for increased behavioral plasticity within species, the authors measured variation in relative cranial capacity over time and space in 10 species of mammals. They predicted that urban populations would show greater cranial capacity than rural populations and that cranial capacity would increase over time in urban populations. Based on relevant theory, they also predicted that species capable of rapid population growth would show more pronounced evolutionary responses.

Results: Urban populations of two small mammal species had significantly greater cranial capacity than rural populations. In addition, species with higher fecundity showed more pronounced differentiation between urban and rural populations. Contrary to expectations, they found no increases in cranial capacity over time in urban populations – indeed, two species tended to decrease cranial capacity over time in urban populations. Furthermore, rural populations of all insectivorous species measured showed significant increases in relative cranial capacity over time.

Conclusions: The results provide partial support for the hypothesis that urban environments select for increased behavioral plasticity, although this selection may be most pronounced early during the urban colonization process. Furthermore, these data also suggest that behavioral plasticity may be simultaneously favored in rural environments, which are also changing because of human activity.

Trends in body size, such as the secular trends in growth and puberty, originate from phenotypic plasticity and reflect directional selection in response to environmental change. Behaviors are prone to as much plasticity in response to the emotional but also the physical environment. This was modeled here and tested in 10 species of mammals, with the prediction that urban populations would show greater cranial capacity than rural populations and that cranial capacity would increase over time in urban populations. They also predicted that species capable of rapid population growth would show more pronounced evolutionary responses. The findings were most pronounced in urban populations of two small mammal species that had significantly greater cranial capacity than rural populations. In addition, species with higher fecundity showed more pronounced differentiation between urban and rural populations. Contrary to expectations, they found no increases in cranial capacity over time in urban populations; in fact, two species tended to have a decrease in cranial capacity over time in urban populations. Furthermore, rural populations of all insectivorous species measured showed significant increases in relative cranial capacity over time. Examples of rapid natural selection processes are accumulating. The directional changes in cranial size that they detected might represent a microevolutionary adaptive tracking as now observed in multiple studies.

Bones control insulin

Bone-specific insulin resistance disrupts whole-body glucose homeostasis via decreased osteocalcin activation

Wei J, Ferron M, Clarke CJ, Hannun YA, Jiang H, Blauer WS, Karsenty G
J Clin Invest 2014;124:1–13

Context: Insulin signaling in osteoblasts has been shown recently to contribute to whole-body glucose homeostasis in animals fed a normal diet; however, it is unknown whether bone contributes to the insulin resistance that develops in animals challenged by a high-fat diet (HFD).

Methods and Results: Here, the authors evaluated the consequences of osteoblast-specific overexpression of or loss of insulin receptor in HFD-fed mice. They determined that the severity of glucose intolerance and insulin resistance that mice develop when fed a HFD is in part a consequence of osteoblast-dependent insulin resistance. Insulin resistance in osteoblasts led to a decrease in circulating levels of the active form of osteocalcin, thereby decreasing insulin sensitivity in skeletal muscle. Insulin resistance developed in osteoblasts as the result of increased levels of free saturated fatty acids, which promote insulin receptor ubiquitination and subsequent degradation.

Conclusions: These results underscore the involvement of bone, among other tissues, in the disruption of whole-body glucose homeostasis resulting from a HFD and the involvement of insulin and osteocalcin

cross-talk in glucose intolerance. Furthermore, the data indicate that insulin resistance develops in bone as the result of lipotoxicity-associated loss of insulin receptors.

A growing body of evidence implicates bone and its cells as an important metabolic organ linked to other metabolically active tissues by insulin. The *Yearbook* has previously cited that insulin stimulates bone cells to produce and activate osteocalcin, which in response increases the efficiency of glucose metabolism through its actions on β cells and other peripheral tissues. This paper now directly explores the contribution of insulin signaling in osteoblasts to the disturbances in whole-body glucose metabolism associated with a high-fat diet. In mice fed a high-fat diet, increased uptake of saturated fatty acids by the osteoblast accelerates the ubiquitination and degradation of the insulin receptor, reducing serum osteocalcin, which in turn exacerbate insulin resistance in muscle and white adipose tissue. Most likely, insulin signaling in osteoblasts also affects global glucose homeostasis through osteocalcin-independent mechanisms.

Small differences have big consequences

Biophysical mechanisms for large-effect mutations in the evolution of steroid hormone receptors

Harms MJ, Eick GN, Goswami D, Colucci JK, Griffin PR, Ortlund EA, Thornton JW
Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA
Proc Natl Acad Sci USA 2013;110:11475–11480

Context: The genetic and biophysical mechanisms by which new protein functions evolve is a central question in evolutionary biology, biochemistry, and biophysics. Of particular interest is whether major shifts in protein function are caused by a few mutations of large effect and, if they are, what mechanisms mediate these changes.

Methods: The authors combine ancestral protein reconstruction with genetic manipulation and explicit studies of protein structure and dynamics to dissect an ancient and discrete shift in ligand specificity in the steroid receptors, a family of biologically essential hormone-controlled transcription factors. They previously found that the ancestor of the entire steroid receptor family was highly specific for estrogens, but its immediate phylogenetic descendant was sensitive only to androgens, progestogens, and corticosteroids.

Results: They show that this shift in steroid receptor function was driven primarily by two historical amino acid changes, which caused a ~70,000-fold change in the ancestral protein's specificity. These replacements subtly changed the chemistry of two amino acids, but they dramatically reduced estrogen sensitivity by introducing an excess of interaction partners into the receptor/estrogen complex, inducing a frustrated ensemble of suboptimal hydrogen bond networks unique to estrogens.

Conclusion: This work shows how the protein's architecture and dynamics shaped its evolution, amplifying a few biochemically subtle mutations into major shifts in the energetics and function of the protein.

Steroid receptors are of two major clades. One contains the estrogen receptors (ERs) and the other – the non-aromatized steroid receptors (naSRs) – includes receptors for androgens, progestogens, mineralocorticoids, and glucocorticoids. Here, the authors synthesized cDNAs for the ligand-binding domains, expressed them, and characterized their sensitivity to a broad panel of hormones, showing hormone-specific binding. Using a library of steroid hormones, they established that ER's estrogen specificity is determined primarily by the requirement for an aromatized A-ring, whereas naSRs specifically exclude aromatized steroids. Combining ancestral protein reconstruction with genetic manipulation, they show that a shift in function was compelled by two historical amino acid changes, which caused a marked change in the ancestral protein's specificity, reducing estrogen sensitivity by introducing an excess of interaction partners into the receptor/estrogen complex. Among others, these findings show that subtle changes in the biochemistry of a few amino acids can be magnified into major evolutionary shifts in function.

Reprogramming of the circadian clock by nutritional challenge

Eckel-Mahan KL, Patel VR, de Mateo S, Orozco-Solis R, Ceglia NJ, Sahar S, Dilag-Penilla SA, Dyar KA, Baldi P, Sassone-Corsi P

Center for Epigenetics and Metabolism, Department of Biological Chemistry, University of California, Irvine, CA, USA
Cell 2013;155:1464–1478

Context: Circadian rhythms and cellular metabolism are intimately linked.

Results: A high-fat diet (HFD) generates a profound reorganization of specific metabolic pathways, leading to widespread remodeling of the liver clock. Strikingly, in addition to disrupting the normal circadian cycle, HFD causes an unexpectedly large-scale genesis of de novo oscillating transcripts, resulting in reorganization of the coordinated oscillations between coherent transcripts and metabolites. The mechanisms underlying this reprogramming involve both the impairment of CLOCK:BMAL1 chromatin recruitment and a pronounced cyclic activation of surrogate pathways through the transcriptional regulator PPAR- γ . Finally, the authors demonstrate that it is specifically the nutritional challenge, and not the development of obesity, that causes the reprogramming of the clock and that the effects of the diet on the clock are reversible.

A large number of physiological events exhibit circadian rhythmicity. At the cellular level, these rhythms are controlled by transcriptional feedback loops that produce oscillations in gene expression, a process associated with circadian changes in chromatin architecture, mRNA processing, protein activity, and protein turnover. It is controlled by specialized factors, including CLOCK and BMAL1, which are the center of this report. Such coordination is necessary for tissue-specific oscillations that control circadian physiology. A high-fat diet (HFD) affects the circadian clock in several ways, including a phase advance of metabolite and transcript oscillations that are maintained on the diet, as well as an abolition of otherwise oscillating transcripts and metabolites. In addition to these disruptive effects, they find a surprising, elaborate induction of newly oscillating transcripts and metabolites. Furthermore, remodeling of the circadian clock by nutrients is rapid and reversible. HFD has pleiotropic effects that lead to a reprogramming of the metabolic and transcriptional liver pathways. It reprograms the circadian transcriptome and metabolome and impairs BMAL1 recruitment to target chromatin sites. This transcriptional reprogramming involves PPAR- γ -driven oscillation in gene expression.

Unexpected relations

A mitochondrial genome sequence of a hominin from Sima de los Huesos

Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, Arsuaga JL, Martinez I, Gracia A, de Castro JM, Carbonell E, Paabo S

Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
Nature 2014;505:403–406

Context: Excavations of a complex of caves in the Sierra de Atapuerca in northern Spain have unearthed hominin fossils that range in age from the early Pleistocene to the Holocene. One of these sites, the 'Sima de los Huesos' ('pit of bones'), has yielded the world's largest assemblage of Middle Pleistocene hominin fossils, consisting of at least 28 individuals dated to over 300,000 years ago.

Results: The skeletal remains share a number of morphological features with fossils classified as *Homo heidelbergensis* and also display distinct Neanderthal-derived traits. Here they determine an almost complete mitochondrial genome sequence of a hominin from Sima de los Huesos and show that it is closely related to the lineage leading to mitochondrial genomes of Denisovans, an eastern Eurasian sister group to Neanderthals.

Conclusions: The results pave the way for DNA research on hominins from the Middle Pleistocene.

In *Yearbook 2012*, we cited work on the Denisovans genome (linked only to a cave in southern Siberia) who left marks in the Australian Aboriginals' DNA. Now, a study of mitochondrial DNA from a 400,000-year-old femur from Spain has revealed a surprising link between the Denisovans, who lived thousands of kilometers away and hundreds of thousands of years later, and Europe's hominin inhabitants of the time, greater than any of these two had with *H. sapiens*. This is the oldest hominin sequence yet published. Previous work suggested that Neanderthals and Denisovans had a common ancestor that lived up to 700,000 years ago. The Spanish hominin could represent a founder population that once lived all over Eurasia and gave rise to the two groups.

A truly wild horse

Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse

Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PL, Fumagalli M, Vilstrup JT, Raghavan M, Korneliussen T, Malaspina AS, Vogt J, Szklarczyk D, Kelstrup CD, Vinther J, Dolocan A, Stenderup J, Velazquez AM, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Roed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KA, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MT, Kjaer K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J, Willerslev E
Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
Nature 2013;499:74–78

Context: The rich fossil record of equids has made them a model for studying evolutionary processes. This article presents a 1.12-times coverage draft genome from a horse bone recovered from permafrost dated to approximately 560–780 thousand years before present (kyr BP). The data represent the oldest full genome sequence determined so far by almost an order of magnitude. For comparison, the authors also sequenced the genome of a Late Pleistocene horse (43 kyr BP), the modern genomes of five domestic horse breeds (*Equus ferus caballus*), a Przewalski's horse (*E. f. przewalskii*) and a donkey (*E. asinus*).

Results: The *Equus* lineage that gave rise to all contemporary horses, zebras and donkeys originated 4.0–4.5 million years before present (Myr BP), twice the conventionally accepted time to the most recent common ancestor of the genus *Equus*. Also, horse population size fluctuated multiple times over the past 2 Myr, particularly during periods of severe climatic changes. It is estimated that the Przewalski's and domestic horse populations diverged 38–72 kyr BP, and there is no evidence of recent admixture between the domestic horse breeds and the Przewalski's horse investigated.

Conclusions: These findings support the contention that Przewalski's horses represent the last surviving wild horse population. The similar levels of genetic variation among Przewalski's and domestic populations indicate that the former are genetically viable and are worthy of conservation efforts. Evidence was found for continuous selection on the immune system and olfaction throughout horse evolution. Finally, 29 genomic regions were identified among horse breeds that deviate from neutrality and show low levels of genetic variation compared to the Przewalski's horse. Such regions could correspond to loci selected during early domestication.

This horse genome is the oldest DNA sequence ever assembled, being about 700,000 years old. It was compared to a 43,000-year-old horse, a modern donkey and five contemporary domestic horse breeds, reconstructing their ancestors to about 4–4.5 million years ago. Other than the amazing technological accomplishment, this paper solves the question whether the Przewalski's horse, which is a rare and endangered subspecies of wild horse native to Mongolia, is really wild and not domesticated like all other horse breeds. The researchers concluded that Przewalski's horse is a separate, truly wild species that split around 38,000–72,000 years ago from the lineage that led to domestic horses.

Modelling the effects of subjective and objective decision-making in scientific peer review

Park IU, Peacey MW, Munafo MR

Department of Economics, University of Bristol, Bristol, UK

Nature 2014;506:93–96

Context: The objective of science is to advance knowledge, primarily in two interlinked ways: circulating ideas, and defending or criticizing the ideas of others. Peer review acts as the gatekeeper to these mechanisms. Given the increasing concern surrounding the reproducibility of much published research, it is critical to understand whether peer review is intrinsically susceptible to failure, or whether other extrinsic factors are responsible for distorting scientists' decisions.

Results: The authors show that, even when scientists are motivated to promote the truth, their behavior may be influenced and even dominated by information gleaned from their peers' behavior, rather than by their personal dispositions. This phenomenon, known as herding, subjects the scientific community to an inherent risk of converging on an incorrect answer and raises the possibility that, under certain conditions, science may not be self-correcting. They further demonstrate that exercising some subjectivity in reviewer decisions, which serves to curb the herding process, can be beneficial for the scientific community in processing available information to estimate truth more accurately.

Conclusions: By examining the impact of different models of reviewer decisions on the dynamic process of publication, and thereby on eventual aggregation of knowledge, the authors provide a new perspective on the ongoing discussion of how the peer-review process may be improved.

There is increasing concern that many published research findings are in fact false. It is common for early findings to be refuted by subsequent evidence, often leading to the formation of groups who interpret the same evidence in notably different ways. The authors cite the false theory of stress as a cause of gastric ulcers. Once established, these can become surprisingly difficult to refute, 'they may become more "vampirical" than "empirical" – unable to be killed by mere evidence'. Science may therefore not be as self-correcting as is commonly believed, and the selective reporting of results can produce literatures that 'consist in substantial part of false conclusions'.

This paper aimed to understand how different criteria of reviewing decisions influence the publication outcome, and how the resulting publication histories and the information inherent in the relevant peer-review criterion influence the community's eventual understanding of the topic. To this end they used two models: (1) the subjective criterion of how strongly the reviewer agrees with the conclusion of the research and other more objective criteria such as research design and methodology, and (2) the decision reflects objective criteria only. The results show that scientists' behavior is dominated by information gleaned from their peers' behavior ('herding', for example, by what has been published previously), rather than by their own personal views. Worryingly, there is evidence that increased popularity of a particular research theme reduces the reliability of published results, and that findings published in prestigious journals are less reliable and more likely to be retracted [8].

The authors encourage: 'It is important that reviewers put weight on their private signals, in order to be able to escape from herding'. Other approaches, such as the use of post-publication peer review, may achieve the same end. Finally, we wish to point out that *Yearbook* editors are encouraged to exercise complete subjectivity when choosing their papers!

References

1. Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, et al: Parental origin of sequence variants associated with complex diseases. *Nature* 2009;462:868–874.
2. Stelzer Y, Sagi I, Yanuka O, Eiges R, Benvenisty N: The noncoding RNA IPW regulates the imprinted DLK1-DIO3 locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nat Genet* 2014;46:551–557.
3. DIAGRAM Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes Consortium, South Asian Type 2 Diabetes Consortium, Mexican American Type 2 Diabetes Consortium, T2D-GENES Consortium: Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* 2014;46:234–244.
4. McMurray F, Church CD, Larder R, Nicholson G, Wells S, Teboul L, et al: Adult-onset global loss of the *fto* gene alters body composition and metabolism in the mouse. *PLoS Genet* 2013;9:e1003166.

5. Park J, Scherer PE: Adipocyte-derived endotrophin promotes malignant tumor progression. *J Clin Invest* 2012;122:4243–4256.
6. Pal A, Barber TM, Van de Bunt M, Rudge SA, Zhang Q, Lachlan KL, et al: PTEN mutations as a cause of constitutive insulin sensitivity and obesity. *N Engl J Med* 2012;367:1002–1011.
7. Senniappan S, Alexandrescu S, Tatevian N, Shah P, Arya V, Flanagan S, et al: Sirolimus therapy in infants with severe hyperinsulinemic hypoglycemia. *N Engl J Med* 2014;370:1131–1137.
8. Brembs B, Button K, Munafo M: Deep impact: unintended consequences of journal rank. *Front Hum Neurosci* 2013;7:291.