

Figure 1 Effect of a single SNP (rs12821256) on hair color. The distal upstream hair follicle enhancer of *KITLG* contains a predicted LEF-binding site overlapping the blond-associated SNP (underlined nucleotide). LEF strongly binds the ancestral allele, whereas weak binding of the derived allele reduces gene expression by ~20%, resulting in lighter hair color.

the precise mutations involved in phenotypic variation that is not medically relevant^{5,6}, questioning whether we learn anything other than ‘molecular details’. Despite the fact that there is an almost overwhelming amount of data documenting changes in gene expression—from comparative studies using approaches from *in situ* hybridization to high-throughput RNA sequencing (RNA-seq)—in a wide range of species, we know little about the general mechanism by which these changes occur, especially in cases of adaptation. And there are many possibilities: evolution could result in the creation of new

regulatory elements or the ablation of existing ones or might tinker with their effectiveness (as in the change in affinity for transcription factor binding shown here), to name a few. Although a general understanding of how gene expression evolves, especially in eukaryotes, will require the accumulation of many empirical examples, this study represents a trusted data point.

Beyond blond: the next steps

The SNP identified by Guenther *et al.* is, of course, not the sole determinant of hair color. Change in *KITLG* expression in human hair

follicles is one of many factors contributing to differences in hair pigment. But, interestingly, whereas many pigmentation-related genes affect both hair and eye color (explaining the strong but imperfect correlation between the two), this change in gene expression is associated with hair pigmentation alone. The next step in this research is to understand exactly how this small decrease in *KITLG* gene expression leads to the relatively large change in hair color phenotype. In mice, changes in *Kitl* expression can affect the number, size, dendricity, differentiation and pigment synthesis of melanocytes, but as of now it remains unclear how the decrease in *KITLG* expression causes blond hair in Europeans. Nonetheless, it is precisely this type of detailed study that provides the necessary first step toward a complete understanding of how changes in genotype translate to changes in phenotype and, at the same time, infuses some much needed optimism that, at least in some cases, we will be able to track down causal variants first suggested by GWAS analyses.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

1. Edwards, S. *et al.* *Am. J. Hum. Genet.* **93**, 779–797 (2013).
2. Walsh, S. *et al.* *Forensic Sci. Int. Genet.* **7**, 98–115 (2013).
3. Guenther, C.A., Tasic, B., Luo, L., Bedell, M.A. & Kingsley, D.M. *Nat. Genet.* **46**, 748–752 (2014).
4. Sulem, P. *et al.* *Nat. Genet.* **39**, 1443–1452 (2007).
5. Rockman, M.V. *Evolution* **66**, 1–17 (2012).
6. Travisano, M. & Shaw, R.G. *Evolution* **67**, 305–314 (2013).

Panning for molecular gold in whipworm genomes

Elodie Ghedin

Two new studies report genomic data on three species of whipworm, soil-transmitted parasitic worms responsible for trichuriasis. These genomes provide insights into host-pathogen interactions and potential for new drug targets for helminth therapy.

Trichuriasis is one of the three major soil-transmitted gastrointestinal helminth infections that cause chronic and debilitating diseases in an estimated 1 billion people worldwide. Trichuriasis, together with ascariasis and hookworm, accounts for a major burden of global disease, with ensuing economic

Elodie Ghedin is in the Department of Biology, Center for Genomics and Systems Biology and at the Global Institute of Public Health, New York University, New York, New York, USA. e-mail: elodie.ghedin@nyu.edu

consequences¹. On page 693 of this issue, Matthew Berriman and colleagues² report the genome sequences of two *Trichuris* species: *Trichuris trichiura*, a human pathogen, and *Trichuris muris*, its mouse counterpart. On page 701, Aaron Jex and colleagues³ report the genome sequence of the pig whipworm *Trichuris suis*. According to the most recent estimates, *T. trichiura* alone infects more than 700 million people, mostly children⁴. *T. muris*, used as a model of the human disease, has been exploited to better understand the role of the host immune system in control

of the infection⁵. Importantly, the ova of this *T. suis* are currently being investigated as a new therapeutic agent against autoimmune diseases⁶.

Map to new targets

Foth *et al.*² and Jex *et al.*³, along with the recent report of the genome sequence of *Trichinella spiralis*⁷, provide substantial genomic information for clade I parasitic nematodes. Foth *et al.*² report the whole-genome sequencing and assembly of the mouse whipworm genome for *T. muris*, using multiple worms and a

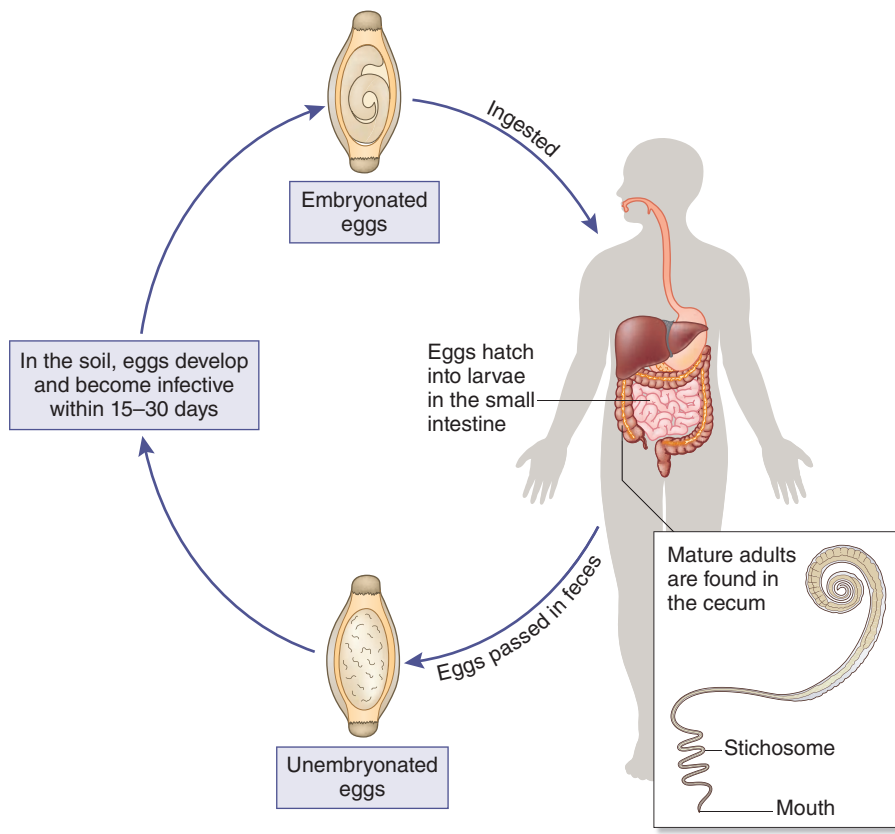


Figure 1 The life cycle of the whipworm *T. trichiura* infecting a human host. Embryonated eggs are ingested by the human host as a result of soil-contaminated hands or food. In the small intestine, the eggs hatch into larvae that mature into adults, which eventually establish themselves in the mucosa of the cecum and ascending colon. Female worms can shed thousands of eggs per day.

combination of complementary sequencing approaches. This high-quality *T. muris* reference map was then used to assemble the draft genome of a single *T. trichiura* male isolated from an infected patient in Ecuador. Jex *et al.*³ sequenced to ~140× coverage the whole genomes from single adult male and female *T. suis* worms, estimating the genome sizes to be 83.6 Mb and 87.2 Mb, respectively. The haploid genome for the *T. muris* female, in contrast, at 106 Mb is 20% larger than its female counterpart in *T. suis*. Gene numbers across the three species vary widely from 9,650 in the 75-Mb *T. trichiura* draft genome to 11,000 in the *T. muris* genome and to more than 14,500 in the *T. suis* genome. Although intrachromosomal rearrangements were apparent in the comparison of *T. muris*, *T. trichiura* and *T. spiralis*², gene content on homologous chromosomes appears largely conserved. Both studies found high conservation of genes across the *Trichuris* species, with species-specific differences largely due to genes of unknown function. Surprisingly, no evidence was found for genes specific to the Y chromosome, which was shown to contain mostly repetitive sequence.

Foth *et al.*² report whole-transcriptome analysis for *T. trichiura* and *T. muris*, and they used this dual-species transcriptome analysis to identify genes showing sex- or life stage-specific expression. In addition, they characterized the transcriptional landscape of the bacillary band and stichosome, specialized cellular structures found in the anterior region of whipworms. To explore the host-parasite interaction in more detail, they also looked at changes in host gene expression during the helper T 1 (T_H1)-like immune response of whipworm-infected mice. Jex *et al.*³ report the whole-transcriptome analysis for *T. suis*, along with stage-, sex- and tissue-specific gene expression patterns.

Deworming—or anthelmintic treatment—is carried out primarily through mass administration of benzimidazole drugs, but there is a high rate of reinfection⁸. Foth *et al.*² identify close to 400 genes and putative drug targets for anthelmintic treatment on the basis of their essentiality, expression patterns and inferred druggability. For a number of these genes, drugs already exist to treat other diseases. Jex *et al.*³ identify specific short noncoding RNAs that have an

important regulatory role in stage-specific differentiation and the sexual development of *T. suis*, providing another avenue for gene disruption. Although *T. trichiura*-specific antigens have not yet been identified for vaccine development as with other helminths⁸, the availability of these new genome sequences also provides an incentive to develop trichuriasis vaccines.

Host-pathogen interactions

By combining genome sequence with tissue-specific gene expression data, both Foth *et al.*² and Jex *et al.*³ give valuable insights into how these parasites potentially modulate the host response. Mammals become infected with *Trichuris* when they ingest embryonated eggs, which hatch in the small intestine and undergo larval development into adult worms (Fig. 1). The worms eventually penetrate the intestinal epithelium of the cecum, where they establish intimate contact with their hosts by transforming the epithelial cells into a tunnel-like structure where the stichosome and bacillary band are embedded. Thereafter, the parasite is able to exchange molecules and nutrients directly with the cytoplasm of the transformed cells. Foth *et al.*² and Jex *et al.*³ each identified serine proteases and serine protease inhibitors as being highly overexpressed in tissue-specific gene expression profiles for these specialized organs in both *T. muris* and *T. suis*. The protease inhibitors also appear to comprise protein domains (such as the WAP (whey acidic protein) domain) associated with immunomodulatory functions, indicating a likely role in controlling inflammation during invasion of the intestinal epithelium.

Low-dose infection of certain mouse strains with *T. muris* leads to chronic infection and reproduces the pathology seen in human ulcerative colitis. To investigate the host response during infection, Foth *et al.*² performed transcriptome analyses of mouse cecum and mesenteric lymph nodes from the same infected mice. Both tissues had similar gene expression profiles and, when compared to tissues from naive mice, displayed an upregulation of immune-specific genes consistent with a T_H1 response, typical of chronic worm infections^{9,10}. Worms like *Trichuris* remain remarkably persistent within their vertebrate host, suggesting that they employ immunomodulatory strategies to regulate host immunity in favor of parasite persistence, with minimal induction of immunopathology. The transcripts for two cytokines thought to control tissue damage in chronic infections (interleukin (IL)-10 and IL-22) were seen by Foth *et al.*² to be upregulated in the infected host. Jex *et al.*³, conversely, propose a model of

Trichuris-driven immunomodulation in which the parasite secretes various molecules with relevant properties, including the ability to stimulate anti-inflammatory IL-10-producing T cells.

Parasite products as immunomodulators

Many reports indicate that helminth-derived molecules have potent regulatory or stimulatory effects on the immune system of their mammalian hosts¹¹ and that these effects could be harnessed as immunotherapeutics against diseases considered to primarily affect the developed world, such as type 1 diabetes and multiple sclerosis¹². Jex *et al.*³ predict that more than 600 of the *T. suis* genes encode excretory-secretory proteins. Significantly, whereas these secreted proteins make up 4% of the *T. suis* genes, they constitute 10% of the transcript abundance across all developmental stages. A key finding in both Foth *et al.*² and Jex *et al.*³ is that many of the genes shown to have higher

expression in the anterior region of the worm could have a central role in immunosuppression and in inhibiting inflammation during damage to the epithelium of the gut. Already, *T. suis* ova are licensed as a medicine and are largely considered to be a safe therapy because they are rapidly expelled by the human host¹³. They have been administered to treat Crohn's disease and ulcerative colitis, but their potential use in controlling inflammatory bowel diseases (IBD) recently suffered a setback when clinical studies failed to demonstrate a statistically significant effect¹⁴. In contrast, clinical tests in multiple sclerosis are proving encouraging¹⁵. The analyses by Foth *et al.*² and Jex *et al.*³ of these three *Trichuris* genomes highlight specific molecules that could be considered to be immunomodulators with therapeutic potential for immune-mediated diseases. These are likely to become more popular in future clinical tests than the alternative, which requires the ingestion of a live parasite.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

- Bethony, J. *et al. Lancet* **367**, 1521–1532 (2006).
- Foth, B.J. *et al. Nat. Genet.* **46**, 693–700 (2014).
- Jex, A.R. *et al. Nat. Genet.* **46**, 701–706 (2014).
- Hotez, P.J. *et al. Lancet* **373**, 1570–1575 (2009).
- Klementowicz, J.E., Travis, M.A. & Grecnis, R.K. *Semin. Immunopathol.* **34**, 815–828 (2012).
- Sandborn, W.J. *et al. Aliment. Pharmacol. Ther.* **38**, 255–263 (2013).
- Mitreva, M. *et al. Nat. Genet.* **43**, 228–235 (2011).
- Zhan, B. *et al. Expert Rev. Vaccines* **13**, 321–331 (2014).
- Schopf, L.R. *et al. J. Immunol.* **168**, 2383–2392 (2002).
- D'Elia, R. *et al. J. Immunol.* **182**, 2340–2348 (2009).
- McSorley, H.J., Hewitson, J.P. & Maizels, R.M. *Int. J. Parasitol.* **43**, 301–310 (2013).
- Moroni, L., Bianchi, I. & Lleo, A. *Autoimmun. Rev.* **11**, A386–A392 (2012).
- Beer, R.J. *Res. Vet. Sci.* **20**, 47–54 (1976).
- Garg, S.K., Croft, A.M. & Bager, P. *Cochrane Database Syst. Rev.* **1**, CD009400 (2014).
- Rosche, B. *et al. Trials* **14**, 112 (2013).

The complex domestication history of the common bean

Brandon S Gaut

A new study reports the genome of common bean (*Phaseolus vulgaris*) and genome-wide resequencing data from both wild and domesticated accessions. These data confirm that common bean was domesticated at least twice, in Mesoamerica and South America, and also provide a framework to identify genes that contributed to the phenotypic changes associated with domestication.

Common bean (*P. vulgaris*) comes in a huge assortment of fresh and dried varieties, including string beans, green beans, French beans, kidney beans and pinto beans. Together, these and other varieties provide a third of daily dietary protein in some cultures, particularly in Africa and the Americas. However, until now, there has been no reference genome for this critical crop. On page 707 of this issue, Scott Jackson, Jeremy Schmutz, Phillip McClean and colleagues¹ report the genome of *P. vulgaris* and use it as a reference to study the complex domestication history of the common bean.

Plant domestication is often associated with a suite of morphological changes². In the case of common bean, domestication has led to increases in seed and leaf sizes, as well as to changes in growth habit and other features. Moreover, these morphological

shifts occurred not once but twice, as common bean was domesticated independently in Mesoamerica (probably in what is now Mexico) and the Andes³. These independent domestication events provide a rare opportunity to assess the genetic effects of domestication in what is, in essence, a replicated evolutionary experiment.

The common bean genome

Schmutz *et al.*¹ sequenced a single *P. vulgaris* accession using a variety of sequencing platforms. Sequences were assembled onto 11 chromosomes and ultimately represent ~80% of the 587-Mb genome. The genome contains few surprises in terms of genomic content (it encodes 27,000 genes and has a high proportion of recent transposon insertions), but at least two features are noteworthy. First, 91% of *P. vulgaris* genes are located within soybean (*Glycine max*) synteny blocks; hence, the genome may prove useful not only to breeders of common bean but also to soybean geneticists. An additional benefit is that the common bean genome is less duplicated than the

soybean genome because *G. max* experienced a whole-genome duplication (WGD) event independently of *P. vulgaris*⁴.

Second, because common bean and soybean differ in their WGD history, their comparison will provide an opportunity for evolutionary biologists to study the effects of WGD on genome fractionation and homolog expression⁵. In fact, an evolutionary comparison of the two species has already yielded an interesting observation: Schmutz *et al.*¹ document that rates of synonymous nucleotide substitution are ~1.4× greater in common bean than in soybean. The underlying cause of this difference is not immediately apparent, but it may reflect variation in current or historical generation times⁶.

Parallel domestication events

To investigate the domestication history of common bean, Schmutz *et al.*¹ resequenced the genomes of 160 wild and cultivated accessions, representing both the Mesoamerican and Andean domestication events. With these data, they applied demographic modeling to

Brandon S. Gaut is at the Department of Ecology and Evolutionary Biology, University of California–Irvine, Irvine, California, USA.
e-mail: bgaut@uci.edu