Published in final edited form as:

Nat Genet. ; 44(3): 237–238. doi:10.1038/ng.2201.

Self fertilization sweeps up variation in the worm genome

Patrick C. Phillips

Institute of Ecology and Evolution, University of Oregon, Eugene, OR 97403-5289 USA

Patrick C. Phillips: pphil@uoregon.edu

Abstract

A new study reports a comprehensive survey of genetic diversity in natural populations of the nematode *C. elegans*. They suggest that recent chromosome-scale selective sweeps have reduced *C. elegans* genetic diversity worldwide and strongly structured genetic variation across its genome.

When creating a new model system for biological research, “plays well with humans” would probably be at the top of the list of desired properties. A model organism should be easy to culture in a human-dominated habitat while having the capacity to endure harsh and variable environments; attributes that contribute to survival at the hands of busy graduate students. The majority of plant and animal species studied in the lab (e.g., mice, rats, fruit flies, *Arabidopsis*) do well in disturbed areas in close proximity to humans. On page 285 of this issue, Leonid Kruglyak and colleagues¹ provide a survey of natural genetic variation in *Caenorhabditis elegans* that confirms that these attributes likely also apply to this model nematode system. Their comprehensive analysis of nearly every known *C. elegans* natural isolate suggests that strong natural selection, undoubtedly driven by human dispersion, has led to extremely reduced levels of genetic variation in isolates from across the globe, especially in genomic regions with infrequent genetic recombination. This study provides unique insights into the role that this species unusual mating system plays within natural populations, and suggests new ways to capitalize on the power of *C. elegans* genetics for understanding ecological and evolutionary processes outside of the lab.

A model system in the wild

C. elegans is one of the best genetically characterized organisms, but we still know very little about its natural history. Previous studies had suggested surprisingly limited genetic variation among individual strains from different parts of the world, but this was based on at most a few dozen genes at a time². Anderson et al.¹ now make effective use of recently developed high throughput sequencing approaches³ to analyze sequence variation in selected genomic regions within a worldwide collection of 200 wild strains, including 58 collection locations on six continents, representing most of the known *C. elegans* isolates. This analysis identified over 41,000 SNPs spread across the genome. Surprisingly, they identified only 47 unique haplotypes among the isolates that they analyzed. They also found extensive long range LD, including between chromosomes. Variation was unevenly distributed across the genome, with most chromosomes displaying very little genetic

variation in their centers, but much higher genetic variation near their tips. Explanations for these observations can be found by considering the *C. elegans* mating system.

Selfing and genetic variation

C. elegans exist primarily as self-fertilizing hermaphrodites, a trait that makes it preferable to mice and fruit flies for many genetic analyses. In the lab, mutant lines can very easily be generated and maintained in a homozygous state because selfing leads to rapid inbreeding, effectively homogenizing the entire genetic background of a given strain. In nature, self-fertilization reduces genetic variation, but by itself is not predicted to reduce variation at any given gene by more than 50%⁴. The more important consequence here is that, under selfing, variants at multiple genetic loci tend to be inherited together. Thus, while sex and crossing-over between chromosomes occur normally, the efficacy of recombination in remixing chromosome segments is reduced because recombination frequently occurs between two genetically identical chromosomes. When a new advantageous mutation arises in the population, the low effective recombination rate tends to drag whole regions of the genome along with the new mutation as natural selection increases its frequency⁵. This hitchhiking tends to suppress variation across the whole genome as opposed to just the region directly surrounding the selected mutation (Fig. 1). This is exactly the pattern found by Anderson et al., in which genetic variation in the worldwide sample is close to zero in the centers of chromosomes, where recombination in *C. elegans* is low, and substantially higher at chromosome ends, where recombination is also much higher (similar patterns are also predicted under the accumulation of numerous deleterious mutations⁶).

The good news about this limited level of genetic variation and long range LD that extends across chromosomes, is the increased potential for association mapping. If one can identify an interesting phenotype that differs among natural isolates, then it should be relatively easy to pinpoint its genetic basis, as has been accomplished in several previous cases⁹. Anderson et al.¹ report here examples for association mapping for two quantitative traits, resistance to abamectin, and aversion to the human pathogen *Pseudomonas aeruginosa*. The bad news is that the identified genetic variants are likely to be fairly limited relative to the full complement of possible causal genetic variants segregating in a comparable outcrossing population and, more importantly, the rampant hitchhiking will make it difficult to infer anything about the evolutionary history of any particular genetic polymorphism (Fig. 1).

Evolutionary history

Earlier work on *C. elegans* natural populations using more limited sampling approaches suggested very few haplotypes and limited population structure among natural isolates², which led to the speculation that *C. elegans* may have been recently spread around the world in association with humans⁷. Such a suggestion is made plausible by the fact that *C. elegans* (and other species from this group) are frequently isolated from human dominated environments such as compost heaps and fruit orchards⁸. Anderson et al.'s high resolution population genetic analyses now suggests that this distribution of sequence variation was caused by a series of selective sweeps, the most recent of which is likely to have occurred within the last few hundred years. Only extreme genetic linkage combined with very strong

selection and recent worldwide migration could generate such a pattern. The mating system of the worms provides the genetic linkage, but only humans could be expected to generate selection and migration strong enough to span six continents.

One still has a sense that the original population of *C. elegans* might still be somewhere out in the world, maybe in Southeast Asia. Or perhaps the self-fertilizing mating system of *C. elegans* simply generates a constant churn of selective sweeps and population turnover that has obliterated any genetic signal of *C. elegans*' deeper evolutionary past. Either way, thanks to Anderson et al.'s comprehensive analysis of genetic variation in existing natural isolates of *C. elegans*, it now appears that these worms are more dependent on us than could have been imagined when they first began to be used as a model system some 50 years ago.

References

1. Anderson EC, et al. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat Genet.* 2012; 44:285–290. [PubMed: 22286215]
2. Cutter AD, Dey A, Murray RL. Evolution of the *Caenorhabditis elegans* genome. *Molecular Biology and Evolution.* 2009; 26:1199–1234. [PubMed: 19289596]
3. Baird NA, et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE.* 2008; 3:e3376. [PubMed: 18852878]
4. Anderson JL, Morran LT, Phillips PC. Outcrossing and the maintenance of males within *C. elegans* populations. *J Hered.* 2010; 101 (Suppl 1):S62–74. [PubMed: 20212008]
5. Barton NH. Genetic hitchhiking. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 2000; 355:1553–1562.
6. Rockman MV, Skrovanek SS, Kruglyak L. Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science.* 2010; 330:372–376. [PubMed: 20947766]
7. Phillips PC. One perfect worm. *Trends Genet.* 2006; 22:405–407. [PubMed: 16806564]
8. Kiontke KC, et al. A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new species from rotting fruits. *BMC Evol Biol.* 2011; 11:339. [PubMed: 22103856]
9. Gaertner BE, Phillips PC. *Caenorhabditis elegans* as a platform for molecular quantitative genetics and the systems biology of natural variation. *Genet Res (Camb).* 2010; 92:331–348. [PubMed: 21429266]

Figure 1.

Influence of a self-fertilizing mating system on the patterns of genetic variation across the genome. When a new advantageous mutation (red) arises against a backdrop of already pre-existing genetic variation on the rest of the chromosome (other colors; A), it will increase in frequency under natural selection (B). Under outcrossing (left), recombination during sexual reproduction will tend to mix genetic markers at other sites until the genetic backgrounds are more uniformly distributed (B). Under self fertilization (right), recombination still occurs, but is less effective because the paired chromosomes tend to be genetically identical to one another (B). Thus, under outcrossing, scanning the pattern of sequence variation across the genome would be expected to identify localized regions of low sequence variation at sites undergoing selective sweeps (C, left), whereas the a similar analysis under self fertilization would be expected to identify an overall pattern of reduced genetic variation (C, right), providing less information about the specific genetic target of natural selection.