PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Caenorhabditis comparative genomics

ArticleInfo		
ArticleID	:	4885
ArticleDOI	:	10.1186/gb-spotlight-20031118-02
ArticleCitationID	:	spotlight-20031118-08
ArticleSequenceNumber	:	237
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	3
ArticleHistory	:	RegistrationDate : 2003–11–18 OnlineDate : 2003–11–18
ArticleCopyright	:	BioMed Central Ltd2003
ArticleGrants	:	
ArticleContext	:	130594411

Cathy Holding
Email: cholding@hgmp.mrc.ac.uk

Comparative genomic analysis of closely related organisms yields considerably more information on the structure and regulation of genes than that garnered by gene prediction programs. In the October *Public Library of Science Biology*, Lincoln Stein at Cold Spring Harbor (CSH) and 35 collaborators from 13 other research centers report the sequencing of the Caenorhabditis briggsae genome. The data, when used in comparison with the C. elegans genome, yield not only a great deal of structural and functional detail, but also unexpected information on the mechanisms and effects of genome evolution (*PLoS Biology* 2003, **1:**166-192).

Jane Rogers, head of sequencing at the Sanger Institute - one of the collaborating centers - explained the choice of sequencing method for the genome. Whole genome shotgun reads combined with a high-resolution physical map were integrated with the previously finished clone-based sequence. "We had only had experience of using a whole genome shotgun approach for relatively small genomes (5-10 Mb). However, Celera had reported a successful assembly of the *Drosophila* genome, and with the reference sequence of *C. elegans* available, we felt confident that a whole genome assembly would be feasible. The project provided a test-bed for Phusion, a new assembler, developed at the Sanger Institute by Jim Mullikin." This technique produced a draft sequence of the *C. briggsae* genome containing 102 Mb of DNA sequence, representing about 98% of the genome.

Sequencing was divided equally between the Genome Sequencing Center (GSC), Washington University at St. Louis, and the Sanger Institute. Bob Waterston, who led the sequencing effort at GSC and was coleader of the project, told us, "The sequence opens up new experimental and computational approaches to understanding the functional elements of the *C. elegans* sequence." Lincoln Stein, coleader of the project and coordinator of the gene analysis for the collaboration, agreed. "On the basis of comparison with *C. briggsae*, we have been able to make corrections to several thousand *C. elegans* gene predictions and possibly to add as many as a thousand new *C. elegans* genes that were not previously appreciated."

This approach revealed 19,500 coding genes with 2169 protein families containing two or more members, with extensive colinearity between the *C. elegans* and *C. briggsae* genomes.

Analysis of the sequence revealed that major evolutionary changes in genomes do not necessarily lead to gross physical changes in the adapted organism. "*C. elegans* and *C. briggsae* diverged 80 to 120 million years ago, somewhat longer ago than human and mouse," Stein wrote in an E-mail tous. "By several measures, the two nematode genomes are very much scrambled relative to each other - far more than human and mouse are. Yet while human and mouse are extremely different in ecological niche, behavior, and anatomy, the two nematode species occupy identical niches, have similar behavior, and are indistinguishable except to experts.

But, despite this 'scrambling,' the size of the two worm genomes is very similar. "Whole genome sequencing indicates that the genome size for *C. briggsae* (104 Mb) is slightly larger than that for *C. elegans* (100 Mb)," confirmed Nansheng Chen, in Stein's lab at CSH, who carried out a large part of the genome analysis. "But the difference is almost entirely due to repetitive sequences, which accounts for 22.4% of the *C. briggsae* genome in contrast to 16.5% of the *C. elegans* genome." Arrangements of operons were highly conserved, with the greatest divergence observed in the chromosomal arms as

opposed to the centers. In the paper, the authors wrote, "The relatively small amount of similarity between the repeat libraries in the two species suggests that most observed dispersed repeat elements postdate the divergence of the two species."

Avril Coghlan at Trinity College Dublin told us, "What was astounding about the *C. briggsae-C. elegans* comparison is that of all the conserved regions that could be found between the two species, only one third lie inside coding regions of genes. This raises the question of what are the other two thirds of conserved regions that lie in introns and intergenic sequences?" In addition, among 60,775 intron orthologues in the two nematode species, 6579 were species-specific introns, suggesting a rate of evolution of intron-exon structure that is greater between the two nematode species than between humans and mice.

"Comparative genomics will be a very important hypothesis-generating tool," concluded Tom Blumenthal of the University of Colorado School of Medicine, whose analysis showed the high conservation of operons. "As we look closer at the comparisons, we will have ideas about how evolution has proceeded and what forces have been at work on the genes and the genomes as a whole."

References

1. Lincoln Stein, [http://www.cshl.org/public/SCIENCE/Stein.html]

2. *PLoS Biology*, [http://biology.plosjournals.org/]

3. Powledge TM: How many genomes are enough? *Genome Biology*, November 18, 2003., [http://genomebiology.com/researchnews/default.asp?arx_id=gb-spotlight-20031118-01]

4. *Caenorhabditis* Genome Sequencing Projects, [http://www.sanger.ac.uk/Projects/C_elegans/]

5. Celera, [http://www.celera.com/]

6. Genome Sequencing Center, [http://genome.wustl.edu/]

7. Tom Blumenthal, [http://www.uchsc.edu/sm/bbgn/blumenthal2.html]