973A

Genome and transcriptome analysis by the next-generation sequencer. **Tadasu Shin-i**[1], Hiroshi Kagoshima[1], Yoshiki Andachi[1], Kazuko Ohishi[1], Atsushi Toyoda[2], Asao Fujiyama[2], Yutaka Suzuki[3], Sumio Sugano[3], Yuji Kohara[1]. 1) Genome Biol Lab., National Inst. Genetics, Mishima, Japan; 2) Comparative Genomics Lab., National Inst. Genetics, Mishima, Japan; 3) Graduate School of Frontier Sciences, Univ. Tokyo, Kashiwa, Japan.

We have been analyzing the gene expression patterns of the nematode *C.elegans* systematically by EST sequencing and whole mount in situ hybridization. All the results are integrated with the genome map based at NEXTDB http://nematode.lab.nig.ac.jp. As an extension of the project we performed whole transcriptome analysis using the SOLEXA sequencer. Poly-A RNA were prepared from 4 samples; embryo, larvae, and adult of wild type hermaphrodites, and mixed stage population of male rich worms and then converted to double-stranded cDNA using the random hexamers primers. The resulting cDNA were subjected to the SOLEXA sequencer, and then the short reads (effective length of 32 bases) were mapped onto the *C.elegans* genome by the ELAND and our newly developed mapping program that allowed a gap to detect the reads from exon-intron junctions. As to the 5' end, splice-leader capped sequences were detected and mapped. These data are integrated in NEXTDB and can be viewed as to exon structure, expression level, developmental stages, sex and so on.

In parallel, we tried the *de novo* assembling of the transcripts using the total 1Gb reads from all the samples by the VELVET assembler which was developed at EBI, and obtained total 36,000 contigs with maximum length of 5 kb and N50 of 450 bases. Although an alternatively spliced region should be assembled to a branch structure, this assembler generates the branches and the trunk as separate contigs, which may cause the relatively short contigs. However, even with such short contigs we have made many corrections on the gene models, e.g., exon structure and splicing patterns.

We are also conducting the genome sequencing of the other nematode *Diploscapter coronatus* in collaboration with Dr. Einhard Schierenberg at Univ. of Koeln; the nematode shows an interesting cell cleavage and arrangement pattern to gastrulation stage. Genomic DNA of the nematode was analyzed by the SOLEXA sequencer and 6Gb paired-end reads (each length of 36 bases) were obtained. They are being assembled in a hybrid manner together with low coverage Sanger shotgun reads. We have already done EST analysis of the nematode and identified 10,000 unique genes in which 5,800 were found to have homologs or orthologs to C.elegans. We will combine these genome and EST data.

974B

Genome-wide evidence for genetic robustness of the alternative splicing machinery in C. elegans. **L. Basten Snoek**[1], Yang Li[2], Rainer Breitling[2], Joost A.G. Riksen[1], Ritsert Jansen[2], Jan E. Kammenga[1]. 1) Nematology, Wageningen University, Wageningen, Netherlands; 2) Groningen Bioinformatics Centre, University of Groningen, Groningen, Netherlands.

Alternative splicing is considered a major mechanism for creating multicellular diversity from a limited repertoire of genes. Here, we studied genetic variation in alternative splicing patterns in a recombinant inbred population of C. elegans, using whole-genome tiling arrays. This experiment allowed us to detect heritable differences in gene expression with exquisite sensitivity and resolution. Nonetheless, we find only a relatively small number of examples of heritable variation in alternative splicing patterns. This is in striking contrast to earlier observations in humans, which showed much less genetic robustness. This observation points to a profound difference in the regulation of the alternative splicing machinery, which parallels the differences in cellular diversity and developmental flexibility in the two species. Nevertheless some striking examples were found of heritable variation in alternative splicing patterns. We detected 382 genes with substantial heritable variation for at least one exon of which the large majority of eQTLs lead to a consistent differential expression across all exons of the affected gene. Some cases show evidence for a necessary refinement of existing gene definitions, predominantly by expanding known exons. We will present genome-wide proof for earlier hypotheses that in C. elegans the alternative splicing machinery exhibits a general genetic robustness, and only a minor fraction of genes shows heritable variation in splicing forms and relative abundance. Furthermore we will show some of these genes showing heritable variation in alternative splicing.

975C

Characterization of synteny blocks and comparative analysis of operons in *Caenorhabditis* species. **Ismael A. Vergara**, Nansheng Chen. Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada.

Accurate detection of synteny blocks is an important task for understanding genome architecture as well as genome expression because these conserved regions could encode essential functionality critical for the fitness of the organism. In this work we apply our newly developed program, OrthoCluster, for identifying and characterizing synteny blocks between the genomes of *Caenorhabditis elegans* and *Caenorhabditis briggsae*. These synteny blocks have enabled us to systematically improve the genome annotation of *C. elegans* and *C. briggsae*, identifying 52 potential novel genes in *C. elegans*, 195 potential new genes in *C. briggsae* as well as 949 novel orthologous relationships. Using this improved dataset, we detected 3,058 perfect synteny blocks between *C. elegans* and *C. briggsae*, with only 288 of these blocks mapping to non-homologous chromosomes in *C. briggsae*. The largest perfect block contains 42 genes and spans 201.2 kbp in chromosome V of *C. elegans*. Additionally, taking advantage of the perfect synteny blocks detected with OrthoCluster, we have examined the conservation of operons between these two genomes. In contrast to previous findings reporting close to 100% (96% and 93.2%) conserved operons between *C. elegans* and *C. briggsae*, we found that only 75.4% of operons are conserved between these two sister species. Operons not conserved are disrupted by different types of genomic structural variations such as translocations, inversions or deletions of one or more orthologous genes in *C. briggsae*. This analysis suggests that operons in *Caenorhabditis* species are more actively evolving than previously thought. This work demonstrates that OrthoCluster can be effectively applied for examining the conservation of any type of eukaryotic genomes and to improve genome annotations based on syntenic information, as long as the correspondence between orthologous genomic features can be established. In the future, this analysis will be extended to include the recently sequenced genomes of *C. remanei*, *C. brenneri* and *C. japonica*. OrthoCluster is accessible at http://genome.sfu.ca/orthoclusterdb/.