SMBE Satellite Meeting / NIG International Symposium:  
THE CAUSES OF GENOME EVOLUTION

March 14-17, 2014

Toray Conference Center and National Institute of Genetics  
Mishima, Japan

Sponsorship and Support

National Institute of Genetics, Mishima, Japan  
Society for Molecular Biology and Evolution  
General Incorporated Association, Tokyo Club  
Kyushu University, Department of Biology  
Inoue Foundation  
Japan Association for DNA Testing
The coupling of recent progress in high-throughput experimental methods with developments in theoretical population genetics have led to major advances in our understanding of genome evolution. We hope that this conference of about 60 participants including established and young investigators will promote exchanges and collaborative projects that will accelerate the advance of genome science. The focus of the symposium will be the interplay of molecular and evolutionary mechanisms underlying genome diversity. This meeting will also include a special session to honor the pioneering work of Tomoko Ohta on the occasion of her 80th birthday. Dr. Ohta was born September 7, 1933 and has made seminal contributions to the population genetics of weak selection as well as gene family evolution over nearly five decades. Her proposal of weak selection as a key factor in evolution (published in 1973) has gained wide acceptance and “near neutrality” is considered a predominant mode of evolution of proteins, regulatory regions and other classes of coding and non-coding DNA. The interaction of genetic drift and natural selection may be critical for understanding the evolution of phenotypic complexity in multicellular eukaryotes. We are looking forward to hearing about leading edge research in evolutionary genomics from all over the world and hope that the content and format of the meeting will help to educate and inspire the next generation of biologists.

Organizing Committee Members

Hiroshi Akashi  
chair  
National Institute of Genetics, Japan

Hidenori Tachida  
deputy chair  
Kyushu University, Japan

Takashi Gojobori  
KAUST, Saudi Arabia

Manyuan Long  
University of Chicago, USA

Naoki Osada  
National Institute of Genetics, Japan

Hideki Innan  
Graduate University for Advanced Studies, Japan

Aya Takahashi  
Tokyo Metropolitan University, Japan

Staff
- Mizue Sei-Macfhearchair, NIG
- Yoko Ueda, NIG
- Asuka Kutsuma, NIG
Toray Conference Center

Seminar Room No. 81&82
(Oral Presentation)

Multi-function room No. 21&22
(Welcome mixer and Poster/Dinner)

Center Building 8F

Center Building 2F
National Institute of Genetics (NIG)

URL: http://www.nig.ac.jp/
TRAVEL GUIDE

Access to the Conference Venue

Nearest Station: MISHIMA station (Japan Railway TOKAI Super Express "Shinkansen" Station)

*FROM NARITA AIRPORT (International Airport):
About 2-2.5 hours from NARITA-Airport to Mishima (pronounced MI-SHI-MA) (The JR "NARITA Express" train from NARITA to TOKYO & JR Super Express "HIKARI or KODAMA" from TOKYO to MISHIMA).

*FROM HANEDA AIRPORT (Domestic Airport, a transit from KANSAI International Airport):
About 1.5 hours (A KEI-KYU train from HANEDA airport to SHINAGAWA & JR Super Express "HIKARI or KODAMA" from SHINAGAWA to MISHIMA)

About Mishima City

"Mishima is located in the east of Shizuoka Prefecture, at the entrance of the Izu-Haokone National Park. Mishima is about 11.1 kilometers (6.9 miles) east to west and about 13.2 kilometers (8.2 miles) north to south. The total area is 62.13 square kilometers (38.6 square miles). Two-thirds of the area is mountainous and hilly districts. The base of the Ashitaka Massif is to the west and the Hakone mountain ranges to the east. To the north of the city we have a fine view of Mt. Fuji (3,766m/12,388ft) and the thawing water from Mt. Fuji gushes out into many ponds and streams around Mishima. People in Mishima are blessed with a beautiful natural environment and a mild climate, and City of Mishima is tackling to establish a “City filled with fresh green, pure water and cultural fragrance”"

(cited from Mishima City Guide)
March 14 (Fri)

Place: Toray Conference Center

17:00~20:00 Welcome mixer (Toray Multi-function Room)

March 15 (Sat)

Place: Toray Conference Center & National Institute of Genetics

8:45~9:15 Coffee/Tea & Registration (Toray Center Building 8F)

Session 1 (Toray Seminar Room)

- 9:15~9:30 Hiroshi Akashi and Tomoko Ohta: Welcome Remarks
- 9:30~10:00 (T1) Laurent Duret: On the Evolution of Recombination Hotspots in Modern and Archaic Humans: Biased Gene Conversion Leads the Race
- 10:00~10:30 (T2) Mizuki Ohno: 8-Oxoguanine Causes Spontaneous de novo Germline Mutations : A Study from the Mutator Mouse Line
- 10:30~11:00 Coffee break
- 11:00~11:30 (T3) Ziheng Yang: On the Estimation of Human Mutation Rate
- 11:30~12:00 (T4) Guy Sella: The Impact of Recent Human Demography on Deleterious Mutation Load and the Genetic Architecture of Disease Susceptibility

12:00~13:30 Lunch

Session 2 (Toray Seminar Room)

- 13:30~14:00 (T5) Junko Kusumi: A Model of Compensatory Molecular Evolution with Indirect Compensation
- 14:00~14:30 (T6) Naoki Osada: Finding Evidence for Weak Compensatory Evolution
14:30~15:00  Coffee break

• 15:00~15:30  (T7) Deepa Agashe: The Evolutionary Dynamics of Dealing with Deleterious Synonymous Mutations
• 15:30~16:00  (T8) Howard Ochman: A New Selective Force Operating in Bacterial Genomes

16:00~16:30  Coffee break

• 16:30~17:00  (T9) Hidenori Tachida: Evolution in Conifer Genome
• 17:30~17:30  (T10) Takashi Gojobori: Comparative Metagenomics: Diversity of Marine Microorganisms in Japan

Move to NIG Auditorium

• 18:15~21:30  Poster Session (P1~P16) / Dinner (NIG Auditorium)

March 16 (Sun)

Place: Toray Conference Center
8:45~9:00  Coffee/Tea

Session 3 (Toray Seminar Room)

• 9:00~9:30  (T11) Toshiyuki Takano-Shimizu: Gene Expression Profiling of Testis in Drosophila melanogaster
• 9:30~10:00  (T12) Chau-Ti Ting: Genetic Basis of Behavioral Isolation Revealed by Comparative Genomic Approach
• 10:00~10:30  (T13) Nobuhiko Tokuriki: Is Enzyme Evolution Reversible? - Exploring Fitness Landscapes by Laboratory Evolution
PROGRAM (CONTINUED)

10:30~11:00  Coffee break

• 11:00~11:30  (T14) D. Allan Drummond: A Nutrient-driven tRNA Modification Alters Translational Fidelity and Genome-wide Protein Encoding across the Drosophilids
• 11:30~12:00  (T15) Aya Takahashi: Pleiotropic Effect and Conflict Resolution in cis-regulatory Mutations Elucidated by Genome-wide Allelic Gene Expression Profiles from a Natural Population of *Drosophila melanogaster*

12:00~15:45  Lunch & Excursion

Session 4 (Toray Seminar Room)

• 16:00~16:30  (T16) Dmitri Petrov: Pervasive, Rapid, and Strong Adaptation in Drosophila
• 16:30~17:00  (T17) Wolfgang Stephan: Genomic Signatures of Positive Selection at QTL for Cold Tolerance in *Drosophila melanogaster*

17:00~17:20  Coffee break

• 17:20~17:50  (T18) Charles H. Langley: Natural selection in the interactions of DNA polymorphism with nucleosome structure and function
• 17:50~18:20  (T19) Naruya Saitou: Neutral Mutations and Purifying Selection Dominate Genome Evolution

18:30~21:30  Poster Session (P17~P32) / Dinner (Toray Multi-function Room)

March 17 (Mon)

Place: Toray Conference Center
8:45~9:00  Coffee/Tea
Session 5 (Toray Seminar Room)

- 9:00~9:30  (T20) **Soojin Yi**: DNA Methylation and Evolution of Duplicate Gene
- 9:30~10:00 (T21) **Doris Bachtrog**: The Epigenome of Evolving Sex Chromosomes: Dosage Compensation and Heterochromatin Formation

10:00~10:30  **Coffee break**

- 10:30~11:00 (T22) **Manyuan Long**: An Emerging Paradigm: New Genes Drive Phenotypic Evolution
- 11:00~11:30 (T23) **Jeff Fawcett**: The Role of Gene Conversion in Preserving Rearrangement Hotspots in the Human Genome

Move to NIG Auditorium

12:15~13:45  **Lunch at NIG Guest House**

Session 6 (NIG Auditorium)

- 13:45~14:00  **Hiroshi Akashi**: Welcome Remarks
- 14:00~14:45 (S1) **Andrew G. Clark**: Positive and Purifying Selection on the Drosophila Y chromosome
- 14:45~15:30 (S2) **Michael Lynch**: Understanding Features of Recombination and Historical Demography from Linkage Disequilibrium in Single Individuals

15:30~16:00  **Coffee break**

- 16:00~16:45 (S3) **Nancy A. Moran**: Genomic Erosion in Symbiotic Bacteria
- 16:45~17:00  **Hiroshi Akashi**: Introduction of Tomoko Ohta
- 17:00~17:30 (S4) **Tomoko Ohta**: Weak Selection in Evolution of Gene Regulation

18:30~21:00  **Dinner at Restaurant**
On the Evolution of Recombination Hotspots in Modern and Archaic Humans: Biased Gene Conversion Leads the Race

In many eukaryotes, the number and location of crossover events has to be tightly controlled to ensure the proper segregation of chromosomes during meiosis. In human and mice, crossovers cluster in narrow hotspots, whose location is determined by the PRDM9 protein, through its DNA binding domain. This domain evolves very rapidly, as a consequence of a positive selective pressure for PRDM9 to change of targets. However, the causes for this selective pressure, and hence the constraints that govern the evolution of recombination hotspots location, are not understood yet. One hypothesis is that this selective pressure might result from the loss of PRDM9 genomic targets via the self-destructive process of biased gene conversion (BGC). To test this hypothesis we investigated the evolution of human historical hotspot loci and of the major human PRDM9 target motif, both in modern and archaic human lineages. We show that human historical hotspots started to be active at most shortly before the split between Denisovans and modern humans, about 400,000-800,000 years ago. Moreover, we show that Denisovan recombination hotspots did not overlap with modern human ones, which indicates that the turnover of recombination hotspots can be extremely rapid. Finally, we quantified the strength of the BGC drive acting on HM motifs in extant human populations. Our analyses indicate that the life expectancy of human recombination hotspots is very short, in agreement with the hypothesis that the gradual degradation of hotspots caused by BGC could be the direct cause of the selective pressure for PRDM9 to change of targets.
Spontaneous germline mutation (GM) generates genetic variation and is regarded as a driving force of molecular evolution. Elucidation of the causes and mechanisms of spontaneous GM may bring us closer to a better understanding of genome evolution in mammals.

We focus on the oxidative DNA lesions as an endogenous cause of spontaneous GM in mammals. Among the four bases, guanine is the most susceptible to oxidation, and 8-oxoguanine (8-oxoG) is a major form of oxidized guanine, which is spontaneously and constantly generated by reactive oxygen species in vivo. 8-OxoG is known as a potent pre-mutagenic lesion, because it can pair with adenine as well as cytosine during DNA replication and cause a G:C to T:A transversion mutation. In E. coli, inactivation of any of the genes; MutM, MutT, and MutY, in the system for preventing 8-oxoG-induced mutation leads to a mutator phenotype.

To assess the contribution of 8-oxoG in de novo spontaneous GM in mouse, we generated Mth1/Ogg1/Mutyh triple knockout (TOY-KO) mouse deficient in 8-oxoG-induced mutation avoidance system. To expect the accumulation of GMs in the progeny, TOY-KO mice were maintained by the intragenerational cross until generation eighth. In this mouse line, we found higher tumor incidence and shorter life span than the control. Moreover, as generation proceeded, we observed the decreased litter size and increased frequency of congenital phenotypic abnormality. By whole exome sequencing analysis, we successfully detected accumulated GMs in the offspring’s genome, which occurred in the mice in ancestral generations. It's notable that most of mutations were G to T transversion, which were caused by 8-oxoG. The mutation rate per generation was increased approximately 18-fold in this mutator mouse line. These results suggest that 8-oxoG potentially induces de novo spontaneous GMs, and its repair system is effectively suppressing those mutations in wild-type animals to maintain their genome and phenotype stable.
Reliable estimation of hominoid mutation rates is of great importance to interpretation of genomic data and in particular to testing hypotheses of human origin and migration. Phylogenetic estimates of the mutation rate typically assume 5-6 million years for the human-chimpanzee divergence, and, with a difference of 1.2% between the two species, give an estimate that is about 1 mutation per kb per million years. Recently genomes from families or pedigrees are sequenced. Unexpectedly direct estimates from counting mutations over generations produced mutation rate estimates that are only about half as large. I will discuss various sources of errors and uncertainties that may account for this discrepancy. One major factor appears to be the interpretation of the fossil evidence for use in molecular clock dating. We are modeling the process of morphological character evolution in both extant and extinct species, to generate better summaries of the fossil evidence to calibrate the molecular clock. By incorporating uncertainties in the fossil data and accommodating polymorphisms in the genomes of ancestral species, one may arrive at phylogenetic estimates of the mutation rate that are close to the direct estimates.
The Impact of Recent Human Demography on Deleterious Mutation Load and the Genetic Architecture of Disease Susceptibility

Yuval Simons\textsuperscript{1}, Michael C. Turchin\textsuperscript{2}, Jonathan K. Pritchard\textsuperscript{3,4*} and Guy Sella\textsuperscript{1}; 1. Department of Biological Sciences, Columbia University, 2. Department of Human Genetics, University of Chicago, 3. Department of Biology, Stanford University, Stanford, 4. Howard Hughes Medical Institute, Stanford University.

Human populations have undergone dramatic changes in population size in the past 100,000 years, including a severe bottleneck of non-African populations and recent explosive population growth. There is currently great interest in how these demographic events may have affected the burden of deleterious mutations in individuals and the allele frequency spectrum of disease mutations in populations. Here we use population genetic models to show that – contrary to previous conjectures – recent human demography has likely had very little impact on the average burden of deleterious mutations carried by individuals. This prediction is supported by exome sequence data showing that African American and European American individuals carry very similar burdens of damaging mutations. We next consider whether recent population growth has increased the importance of very rare mutations in complex traits. Our analysis predicts that for most classes of disease variants, rare alleles are unlikely to contribute a large fraction of the total genetic variance, and that the impact of recent growth is likely to be modest. However, for diseases that have a direct impact on fitness, strongly deleterious rare mutations likely do play important roles, and the impact of very rare mutations will be far greater as a result of recent growth. In summary, demographic history has dramatically impacted patterns of variation in different human populations, but these changes have likely had little impact on either genetic load or on the importance of rare variants for most complex traits.

Reference:
A Model of Compensatory Molecular Evolution with Indirect Compensation

Junko Kusumi1, Motoshi Ichinose2, Masasuke Takefuji, Wolfgang Stephan4 and Masaru Iizuka5; 1. Department of Environmental Changes, Faculty of Social and Cultural Studies, Kyushu University Kyushu University, 2. Department of Contemporary Liberal Arts, Junior College, Chikushi Jogakuen University, 3. Center for Comprehensive Community Medicine, Faculty of Medicine, Saga University, 4. Department of Biology II, Section of Evolutionary Biology, Ludwig-Maximilian-University, 5. Division of Mathematics, Kyushu Dental University

If there is an extra pair of sites that can reduce the deleterious effects of single mutants at directly interacted pair of sites, this reduction is called indirect compensation. This idea was motivated by compensatory molecular evolution of t-RNA secondary structure. A previous study reported that a mutation that destroyed a Watson-Crick pair in a t-RNA stem structure could be compensated by creation of a new pair in the neighborhood. This phenomenon appears in molecular evolution of t-RNAs in mammalian lineages repeatedly. This observation indicates that compensatory evolution could occur not only by a mutation at directly interacted pair of sites but also a mutation at extra sites neighboring the original pair sites by forming a new bond. Here, we considered the effects of “indirect compensation” on the rate of compensatory molecular evolution by introducing a simple mathematical model. Computer simulations demonstrate that the effect of indirect compensation effectively decrease the fixation time of double mutants of direct compensatory sites if mutation is weak (4Nu = 0.01, 0.1) and deleterious effect of selection is strong (4Ns > 3). The degree of reduction of the fixation time depends on the relative ratio of mutation rate to selection intensity. Further, we investigate properties of sample paths by classifying them to, direct pathway, indirect pathway I and indirect pathway II. Each pathway corresponds to no change on indirect compensatory site, the first change on direct compensatory sites and the second change on indirect compensatory site, and the first change on indirect compensatory site and the second change on direct compensatory sites. It is shown that most of the double mutants appear through the indirect pathway II when selection is strong. This pathway enables individuals to complete compensatory molecular evolution without reducing fitness if a mutation at indirect compensation site compensates completely the deleterious effects of single mutants at directly interacted pair of sites. Our results suggest that the compensatory evolution with indirect compensation may be responsible to the molecular evolution of t-RNA secondary structure.

Reference:
Oral Presentation: T6

**Finding Evidence for Weak Compensatory Evolution**

Naoki Osada and Hiroshi Akashi; Division of Evolutionary Genetics, Department of Population Genetics, National Institute of Genetics

Lineage-specific accelerations of protein evolution working in mitochondria have been reported among mammals, which may reflect Darwinian co-adaptation between mammalian mitochondrial and nuclear genomes for efficient energy production. However, mitochondria have several specific genetic features such as lacking recombination, small effective population size, and high mutation rate, which may promote the accumulation of deleterious mutations. Here, we hypothesize that adaptive evolution in nuclear-encoded proteins working in mitochondria may compensate weakly deleterious substitutions fixed in mitochondrial-encoded proteins. This model of compensatory weak selection, which was motivated by the work of Tomoko Ohta, has a different evolutionary trajectory from that under compensatory neutral model, where two compensatory mutations fix together. Distinguishing these two different types of compensatory evolution processes is important to understand the process of molecular evolution and effect of weak selection on genomes.

To test the idea, we investigated the relationship between the pattern of molecular evolution and the protein structure of the primate cytochrome c oxidase (COX) complex, which is composed by both nuclear- and mitochondrial-encoded subunits. Occurrences of coupled amino acids changes on same branches at sites in physically close distance were elevated between mitochondrial- and nuclear-encoded COX subunits. The pattern was not observed between two mitochondrial subunits and between two nuclear subunits. Furthermore, the compensatory model predicts that, following a deleterious fixation in mitochondria, some compensatory substitutions in nuclear may also occur after splitting events in the phylogeny. We show that physically close pairs of codons had more nuclear substitutions preceded by mitochondrial substitutions than the distant pairs, indicating compensatory evolution accounted for the accelerated evolution of nuclear-encoded COX proteins in primates.

**Reference:**

Recent evidence indicates that synonymous codon changes may sometimes face quite strong selection, although it remains difficult to derive general patterns about the nature and strength of such selection. In previous work with synonymous variants of an enzyme-encoding gene of Methylobacterium extorquens, we showed that altering codons could be extremely deleterious. Although the exact physiological mechanism likely depends on the specific sequence, the fitness disadvantage arose from insufficient enzyme production.

We now show that during laboratory evolution, these synonymous variants rapidly regain fitness, often via repeatable and mutant-specific beneficial point mutations in the N-terminal region of the gene. Interestingly, none of the SNPs (some of which were synonymous) caused a reversion to the wild-type codon, but all of them increased focal gene and protein expression. Other putatively beneficial compensatory mutations in the genome that were associated with increased enzyme production did not involve tRNA genes. Our results thus suggest that co-evolutionary dynamics between tRNA copy number and codon use may be unlikely in the short-term, potentially because of the existence of multiple fitness peaks. Instead, bacteria may find diverse evolutionary solutions to the immediate physiological problems caused by accumulated deleterious synonymous mutations. Subsequent tRNA gene copy number changes may fine-tune global protein production in the long term.

Reference:
Nucleotide base composition of bacterial genomes was among the first molecular characters considered to reflect a purely neutral process, and continued research during the past 50 years has been unable to resolve an adaptive basis for the variation in base composition among bacterial groups. Relatively recent evidence, based on nucleotide sequence comparisons, suggests that the mutational input alone cannot produce observed base compositions, implying a role for natural selection. Because bacterial genomes have very high coding densities, forces that act on the G+C contents of genes could shape the base composition of entire genomes. We tested the hypothesis that differences in genomic base compositions are actually due to genic selection by testing growth rates of bacterial strains that expressed genes engineered to have different nucleotide contents. These experiments revealed that bacteria expressing GC-rich versions of genes display higher growth rates than those expressing the identical protein from AT-rich versions, again showing that selection favors higher genic G+C in a genome where mutations are biased towards A+T. Disruptions of ribosome binding sites largely eliminated these differences, suggesting that selection favoring higher G+C content depends on features of mRNA that interact with ribosomes. This selection is sufficiently intense, on each base change, to govern genomic base composition. Moreover, we have tested these gene variants in phylogenetically divergent genomes of widely different base compositions, and we are currently searching for the substrate upon which selection is acting.
Conifers constitute the largest group of gymnosperms. They have a larger genome size than most other angiosperms, a long generation time, and have undergone few chromosome duplications during their evolution. Because of these characteristics, conifers are interesting targets of molecular evolutionary studies. Nonetheless, there have been only a few studies regarding their genome structure and the levels and patterns of diversity in non-coding regions and they are mostly limited to Pinaceae. Because conifers families are phylogenetically separated by a few hundred million years, obtaining such information from the other families than Pinaceae is highly desirable. *Cryptomeria japonica*, belongs to a conifer family, Cupressaceae, and is the most important timber tree in Japan, making investigating the structure and diversity of its genome both interesting and worthwhile. We have been studying this species from this standpoint and I report some of the results here.

First, we analyzed the sequences of several Bacterial Artificial Chromosome (BAC) clones from *C. japonica* and compared them with comparable sequences from other model plants. From this analysis, we identified several features of the *C. japonica* genome. First, the genome of *C. japonica* has many repetitive sequences, and they are divergent. Second, we found a putative protein-coding gene with a very long intron (approximately 70 kb) despite previous plant genome studies that showed that plant genes have mostly short introns. Finally, CpG deficiency was far greater in conifers than in other plants.

Next, we studied patterns of genetic diversity and levels of linkage disequilibrium in genomic regions using the BAC clone sequences. Each region spanned up to approximately 100 kb and was mostly non-coding. The average nucleotide diversity in these regions was 0.35%, comparable to the 0.44% observed at silent sites in the coding regions previously studied in this species. Neutrality statistics did not deviate significantly from the expectations of the standard neutrality model, again showing similarity to those in the coding regions. However, linkage disequilibrium was extensive and did not decay even at a distance of 100 kb. The average estimate of the population recombination rate per base pair was less than 1/30 of that in the coding regions. Thus, these regions seemed to evolve similarly to those in the coding regions, except for their recombination rates.

I will discuss these features of genome structure, diversity and linkage disequilibrium in *C. japonica* in the context of conifer evolution.
Environmental metagenomics is a genomic approach in which genomic fragments of any species contained in environmental samples such as a bottle of sea water and a cup of land soil are sequenced without morphological identification of those species in order to observe ecological features of a diversity of microorganisms. When this kind of metagenomics is applied for studies of marine microorganism diversity in particular, species composition of microorganisms obtained can be used for environmental evaluation of the sea water in addition to understanding of dynamic features of marine microorganism diversity.

Here, we present our on-going projects in which environmental metagenomics approach has been for understanding of marine microorganism diversity in the sea surrounding the Japan islands. In particular, we would see if it is feasible to predict emergence of the so-called red tide phenomena toward prevention from them. Our preliminary results show that a species distribution of microorganisms manifests significantly enough changes with time to conduct comparative analysis. When a given set of major species of microorganisms are paid keen attention to, they can be a clear indicator of possible red tide phenomena.

Thus, comparative studies of marine metagenomics with time-monitoring may be useful for not only the understanding of marine microorganism diversity but also the applications to environmental monitoring of the sea.
Although meiotic recombination (crossover and gene conversion) plays an important role in numerous fundamental issues in evolutionary biology (e.g., the distribution of standing nucleotide diversity), we lack a refined understanding of the evolutionary and genetic factors that regulate their occurrence. Despite recent progress on human and some classic model organisms such as Drosophila and Arabidopsis, the recombination pattern in other species is poorly understood. This study investigates the meiotic recombination patterns in the freshwater microcrustacean *Daphnia pulex*. With whole genome sequences of 96 Daphnia isolates, we analyzed the genome-wide pattern of linkage disequilibrium to estimate the rate and distribution of crossover events in the Daphnia genome. Furthermore, we used whole genome sequences of single sperm to examine both the crossover and gene conversion events in male Daphnia. Despite being at an early stage of this project, we will present the exciting, preliminary findings about meiotic recombination in the Daphnia genome, as well as the challenges that we are still facing.
The gut microbial communities within great apes have been shown to reflect the phylogenetic history of their hosts, indicating co-diversification between great apes and their gut microbiota over evolutionary timescales. But because the great apes examined to date represent geographically isolated populations whose diets derive from different sources, it is unclear whether this pattern of co-diversification has resulted from a long history of co-adaptation between microbes and hosts (heritable factors) or from the ecological and geographic separation among host species (environmental factors). To evaluate the relative influences of heritable and environmental factors on the evolution of the great ape gut microbiota, we assayed the gut communities of sympatric and allopatric populations of chimpanzees, bonobos and gorillas residing throughout equatorial Africa. Comparisons of these populations revealed that the gut communities of different host species can always be distinguished from one another but that the gut communities of sympatric chimpanzees and gorillas have converged in terms of community composition, sharing on average 53% more bacterial phylotypes than the gut communities of allopatric hosts. Host environment, independent of host genetics and evolutionary history, shaped the distribution of bacterial phylotypes across the Bacteroidetes, Firmicutes, Proteobacteria and Actinobacteria, the four most common phyla of gut bacteria. Moreover, the specific patterns of phylotype sharing among hosts suggest that chimpanzees living in sympatry with gorillas have acquired bacteria from gorillas. These results indicate that geographic isolation between host species has promoted the evolutionary differentiation of great ape gut bacterial communities.
Why are there so many Species with Y-autosome Fusions?

Jun Kitano, Tomotaka Matsumoto, Kohta Yoshida, and the Tree of Sex Consortium; 1. National Institute of Genetics, 2. NESCent

Sex chromosomes turn over rapidly in some taxonomic groups. One of the mechanisms causing the turnover of sex chromosome is a chromosomal fusion between a sex chromosome and an autosome. Although sex chromosomes were thought to be conserved in soma taxa, such as mammals, recent genomic studies have revealed that sex chromosome-autosome fusions have occurred multiple times even in these taxa. Furthermore, our previous studies have demonstrated that sex chromosome-autosome fusions may promote speciation and genomic divergence between incipient stickleback species (see Kohta Yoshida’s poster). Therefore, elucidation of the driving forces of sex chromosome-autosome fusions is essential for a better understanding of the evolutionary mechanisms of sex chromosome evolution and speciation.

Previous theoretical studies suggest that evolutionary forces, such as sexually antagonistic selection and heterozygote advantage, can drive sex chromosome-autosome fusions and both mechanisms likely favor the fixation of Y-autosome fusions (W-autosome fusions) over X-autosome fusions (Z-autosome fusions). Consistent with this prediction, our compilation of cases of sex chromosome-autosome fusions in vertebrates revealed that Y-autosome fusions occur more frequently than X-autosome fusions in fishes and reptiles. However, in mammals, we found that X-autosome fusions are as common as Y-autosome fusions. These differences may result from different patterns of female meiotic drive between taxa. Second, we found that fusions occur more frequently in XY systems than in ZW systems. By using individual-based modeling, we explored the conditions where Y-autosome fusions occur more frequently than W-autosome fusions. We found that either a combination of slightly deleterious effects of fusions and low male effective population size in polygynous mating system or asymmetric sexually antagonistic selection (the presence of autosomal alleles with detrimental effects on females being slightly larger than beneficial effects on males) can favor the fixation of Y-autosome fusions over W-autosome fusions. Thus, our studies demonstrate that XY and ZW systems differ in the frequencies of fusions with autosomes.

Reference:

**Poster Presentation: P4**

**Tissue- and Stage-dependent Dosage Compensation on the Neo-X Chromosome in *Drosophila pseudoobscura***

Masafumi Nozawa; Laboratory for DNA Data Analysis, Center for Information Biology, National Institute of Genetics

Sex chromosome dosage compensation (or DC) is widely accepted in various organisms. This concept is mostly supported by comparisons of gene expression between chromosomes and between sexes. However, genes on the X chromosome and autosomes are mostly not homologous, and the average gene expression level on these chromosomes may not be the same even under DC, which complicates comparisons between chromosomes. Many genes with sex-biased expression also make comparisons between sexes difficult. To overcome these issues, we investigated DC by comparing the expression of neo-X-linked genes in *Drosophila pseudoobscura* with those of their autosomal orthologs in *D. melanogaster*. The ratio of the former to the latter in males would be 1 under DC, whereas it becomes 0.5 without DC. We found that the ratio was ~0.85 for adult whole bodies, indicating that the DC is incomplete on the neo-X chromosome in adults as a whole. The ratio (~0.90) was also significantly less than 1 for adult bodies without gonads, whereas it was ~1.0 for adult heads. These results indicate that DC varies among tissues. Our sliding-window analysis of the ratio also revealed that the upregulation of neo-X-linked genes in males occurred chromosome wide in all tissues analyzed, indicating global upregulation mechanisms. However, we found that gene functions also affected the levels of DC. Furthermore, most of the genes recently moved to the X were already under DC at the larval stage but not at the adult stage. These results suggest that DC in *Drosophila* species operates in a tissue/stage-dependent manner.

Reference:
The chromatin landscape is key for gene regulation, but little is known about how it differs between the sexes or between species. Here, we study the sex-specific chromatin landscape of Drosophila miranda, a species with young sex chromosomes, and compare it to the model organism D. melanogaster. We analyze six histone modifications in male and female larvae of D. miranda and define seven biologically meaningful chromatin states that show different enrichment for transcribed and silent genes, repetitive elements, housekeeping and tissue-specific genes. The genome-wide distribution of both active and repressive chromatin states differs between males and females. In males, active chromatin is enriched on the X, relative to females, due to dosage compensation of the hemizygous X. Furthermore, a larger fraction of the euchromatic portion of the genome is in a repressive chromatin state in females relative to males, consistent with previous work suggesting a role for the Y chromosome in global heterochromatin distribution. However, sex-specific chromatin states appear not to explain sex-biased gene expression. Overall, conservation of chromatin states between male and female D. miranda is comparable to conservation between D. miranda and D. melanogaster, which diverged over 30MY ago and lacks the secondary sex chromosomes of D. miranda. Active chromatin states are more highly conserved across species, while heterochromatin shows very low levels of conservation. Divergence in chromatin profiles contributes to expression divergence between species, with about 26% of genes in different chromatin states in the two species showing species-specific or species-biased expression.
DNA methylation is a common feature of eukaryotic genomes and is especially common in noncoding regions of plants. Protein coding regions of plants are often methylated also, but the extent, function, and evolutionary consequences of gene body methylation remain unclear. Here we investigate gene body methylation using an explicit comparative evolutionary approach. We generated bisulfite sequencing data from two tissues of *Brachypodium distachyon* and compared genic methylation patterns to those of rice (*Oryza sativa* ssp. *japonica*). Gene body methylation was strongly conserved between orthologs of the two species and affected a biased subset of long, slowly evolving genes. Because gene body methylation is conserved over evolutionary time, it shapes important features of plant genome evolution, such as the bimodality of G+C content among grass genes. Our results superficially contradict previous observations of high cytosine methylation polymorphism within *Arabidopsis thaliana* genes, but reanalyses of these data are consistent with conservation of methylation within gene regions. Overall, our results indicate that the methylation level is a long-term property of individual genes and therefore of evolutionary consequence.
The number of N-glycosylation sites (NGS) as well as the positive charge (+charge) of hemagglutinin (HA) are known to have increased after influenza A virus subtype H3N2 (H3N2 virus) entered the human population in 1968. Experimentally, it has been shown that N-glycans attached to NGS around the antigenic sites (AS) block the binding of antibodies (Ab) to AS and increases in the +charge enhance the receptor-binding avidity of HA. Therefore, these were considered to have contributed to the immune escape of the virus. However, since N-glycans may impair the receptor-binding avidity by covering the receptor-binding pocket and increases in the +charge may inhibit the release of progeny virus from infected cells, evolutionary mechanisms for these phenomena remain an enigma. To clarify the role of NGS in HA, we examined natural selection operating at AS before and after gains of NGS. Positive selection was detected before but not after gains of NGS, supporting the hypothesis that NGS generated in HA contributed to blocking the binding of Ab to AS. In addition, by designing a single-substitution analysis method for detecting episodic natural selection, we demonstrate that gains of N-linked glycosylation sites in HA during evolution of H3N2 virus were subject to positive selection. Although gains of NGS possibly reduced the receptor-binding avidity of HA, we observed that gains of NGS occurred almost coincidentally with increases in the +charge. These results suggest that increases in the +charge of HA, which enhance the receptor-binding avidity, may have occurred to compensate for the reduced receptor-binding avidity caused by gains of NGS during evolution of H3N2 virus.

References:

Clade Replacement and Molecular Evolution of Dengue Serotype II Viruses in the Philippines

Kent D. Kawashima¹, Hannah Karen M. Labayo², Veni R. Liles², David C. Klinzing², Filipinas F. Natividad², Ronald R. Matias²; ¹. National Institute of Genetics, ². Research and Biotechnology Division, St. Luke’s Medical Center

Dengue viruses are arthropod-borne viruses from the Flavivirus genus that cause systemic viral infection in humans. The global disease burden of Dengue infections is estimated to be about 390 million infections per year; most infections occur in tropical regions where all four Dengue serotypes are endemic, and the Aedes mosquito vector is present. Compared to other regions affected by Dengue, and despite the growing threat to public health, there is an acute lack of information about Dengue viruses found in the Philippines.

We used a modified sequence-independent single primer amplification protocol and next-generation sequencing to deep sequence Dengue serotype 2 viruses (DENV-2) isolated from 1995 to 2010 in the Philippines. Phylogenetic analysis of DENV-2 sequences showed samples segregated to either Asian II or Cosmopolitan subtypes. All samples isolated before 1998 classified with the Asian II subtype while recent samples were all members of the Cosmopolitan subtype. Between 1998 and 2002, the Asian II subtype was less frequently isolated while the Cosmopolitan subtype became more commonly sampled. The Asian II subtype is one of the subtypes originating from Southeast Asia and is currently most prevalent in South and Southeast Asia while the Cosmopolitan subtype is not associated with any particular region but is frequently isolated in urban centers across Asia, and Central and South America. Comparing structural and non-structural genes of these two subtypes, as well as other DENV-2 subtypes, we found all genes of all DENV-2 subtypes were under varying degrees of strong purifying selection (ω<0.1) suggesting a strong restriction to adaptive evolution. As DENV-2 replicates both in humans and mosquitoes, the virus needs to strike a balance between two very different organisms. Perhaps because of this – when faced with a changing fitness landscape in terms of herd immunity and abiotic factors – that subtype replacement occurs. Because the dominant subtype is unable to evolve fast enough to adapt to the changing fitness landscape, the minority subtype, which has become more fit due shift in the fitness landscape, progressively replaces it.

References:

The transfer of organelle DNA fragments to the nuclear genome is frequently observed in eukaryotes. These transfers are thought to play an important role in gene and genome evolution of eukaryotes. In plants, such transfers occur from plastid to nuclear [nuclear plastid DNAs (NUPTs)] and mitochondrial to nuclear [nuclear mitochondrial DNAs (NUMTs)] genomes. The amount and genomic organization of organelle DNA fragments have been studied in model plant species such as *Arabidopsis thaliana* and rice. At present, publicly available genomic data can be used to conduct such studies in non-model plants. In this study, we analyzed the amount and genomic organization of NUPTs in 17 plant species for which genome sequences are available. The amount and distribution of NUPTs varied among the species. We also estimated the distribution of NUPTs according to the time of integration (relative age) by conducting sequence similarity analysis between NUPTs and the plastid genome. The age distributions suggested that the present genomic constitutions of NUPTs could be explained by the combination of the rapidly eliminated deleterious parts and few but constantly existing less deleterious parts. After the integration to nuclear genome, NUPTs tend to have GC to AT biased mutation. However, the biases differ by age of NUPTs. Old NUPTs have high ratio of AT to GC change compared to young NUPTs. This tendency was observed commonly in analyzed species. The mutation bias would be changed related to deleterious effect of NUPTs.

Reference:

- Yoshida, Furihata, and Kawabe (2014) DNA Research
Characterization of ESTs Collected from *Cryptomeria japonica*

Miho Tamura¹, Kouichi Teshima², Kentarou Mishima³, Taiichi Iki¹, Mine Nose³, Miyoko Tshubomura³, Manabu Kurita³, Atshushi Watanabe¹; 1. Depertment of Forest Enviromental Science,Faculty of Agriculture,Kyushu University, 2. Deptment of Biology, Faculty of Sciences, Kyushu University, 3. Forest Tree breeding Center, Forestry and Forest Products Research Institute

Conifers are the largest group of gymnosperms. Recently, a draft genome of conifer species, *Picea abies*, belonging to Pinaceae, has been published, and it revealed that repetitive sequences are dispersed throughout genome. Although conifer genome is an interesting target of molecular evolution, it is not easily to access conifer genome study because of large genome size as compared to most angiosperm. Conifer have been divided into two groups, conifer I includes only Pinaceae and conifer II includes other five families excluding Pinaceae. Pinaceae is a center of conifer evolutionary and genome study. To understand conifer evolution well need to obtain more information from conifer II. However, it is too time- and resource-consuming, even with next generation sequencing, to perform on conifer species. Instead of genome, we obtained ESTs from several organs including tree-top, shoot with needle, cambium, root and male flower in *Cryptomeria japonica* following different developmental stages and seasonal changes. *C. japonica* that belongs to Cupressaceae, conifer II, is separated from Pinaceae phylogenetically for more than 200,000,000 years. Up to date, we collected about 1.8 million reads (707.8 Mbp) using Roche 454 and they were assembled into 22,250 isotigs by Newbler, with average length 1469.62 bp, and average depth was 13.22 reads per site. Comparing to CDS of other plant, *Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa*, *Picea abies*, *Selaginella moellendorffii* and *Physcomitrella patens*. The average length of *C. japonica* is similar to those of *A. thaliana* and *O. sativa* and is longer than those of *P. abies* and *P. tricocarpa*. Therefor, these isotigs might be nearly full-length cDNA. When reads from each organ were mapped to isotigs, a half of isotigs was shared by all organs. However some isotigs were recognized to specific organ, ESTs derived from cambium were different from other organs, especially. Most group of transcription factor (TF) ware detected in *C. japonica* and number of TF sequence were similar to that of other model plants. The amount of repetitive sequences in *C. japonica* was not significantly different from other model plant, but more transposons like sequence in *C. japonica* than others. ESTs collected from *C. japonica* provide variable informations to understand gene evolution in conifer.
Expanding and Contraction of the Olfactory Receptor Universe: Orthology among 13 placental mammals highlights the evolutionary fate of genes

Yoshihito Niimura1,2, Atsushi Matsui1,2, Kazushige Touhara1,2; 1. Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 2. ERATO Touhara Chemosensory Signal Project, JST, The University of Tokyo

Olfaction is essential for the survival of most animals. It is used for finding foods, recognizing mates and offspring, and avoiding preys and dangers. Diverse odor molecules in the environment are detected by olfactory receptors (ORs) expressed in the sensory neurons of the olfactory epithelium. OR genes form the largest multigene family in mammals. There are ~400 OR genes in the human genome, while the number is >1,000 for mice or rats. Extensive bioinformatic analyses using the whole genomes of various organisms showed that the numbers of OR genes extensively vary among species, reflecting the diversity of species’ living environments1. We also reported that OR genes are characterized by extremely frequent gene duplications and losses2.

However, whether the extent of gene gains/losses differs among individual ORs and what might generate such variation is unknown. To answer these questions, we classified >10,000 intact OR genes in 13 placental mammal species into orthologous gene groups (OGGs) and compared among these OGGs. We found that OR genes experienced less gene duplication were under stronger purifying selection and that ORs with broader ligand specificity tended to generate more descendants. Class 2 OR genes evolved more dynamically than Class 1 OR genes did. Only three OGGs showed complete one-to-one orthology without any gene gains/losses, and these OGGs are also the most highly conserved in amino acid sequences. This study provides a basis for inferring OR function from evolutionary trajectory.

References:
Sensing the temperature of the surrounding environment should be an important traits for organisms in terms of adaptation. Several members of transient receptor potential (TRP) superfamily are known to work as thermo sensors. These genes are called as thermoTRPs. TRPs are ion channels which have six transmembrane domain. ThermoTRPs open their channel by the heat or cold stimuli, but the number of thermoTRPs members or the temperature they sense is frequently changed during evolution. I conducted analysis to find out the candidates of TRPs in two echinoderm genomes, sea urchin and starfish. The larvae of them have been shown to have thermotaxis and the thermoTRPs of them should play a great role. Based on HMM search using known TRPs, I found that sea urchin and starfish both potentially have almost the same number of TRP genes as vertebrates. I also found that sea urchin has more than 10 TRPA genes, while fruit fly have four and human have only one. The unusual expansion of TRPA gene in sea urchin may have functionally important for its thermotaxis and adaptation.
Population Genetic Analysis of *Distylium racemosum*, a Climax Tree Species in Japanese Evergreen Broadleaf Forests

Kosuke M. Teshima¹, Eri Sato¹, Atsuya Kamei¹, Setsuko Suzuki², Kyoko Sugai³, Atsushi Sakai², Nobuyuki Inomata⁴, Shinichi Yamamoto⁵, Hidenori Tachida¹; 1. Kyushu University, 2. Forestry and Forest Products Research Institute, 3. University of the Ryukyus, 4. Fukuoka Women’s University, 5. Okayama University

*Distylium racemosum* is one of the most dominant species in primeval old-growth evergreen broad-leaved forest in western Japan. We studied the pattern of genetic variation of *D. racemosum* and its closely related species *D. lepidotum*, which is endemic to Ogasawara Islands. Leaf samples of *D. racemosum* were collected from 6 populations in Kyushu and Shikoku. Eight individuals were analyzed for each population. Samples of *D. lepidotum* were collected from Chichi-jima and Haha-jima. 12 individuals from Chichi-jima and 11 individuals from Haha-jima were analyzed. Sequences of 42 protein coding genes were obtained using Roche 454 system. Nucleotide diversity was about 0.3%, which is in good agreement with the previous study using Sanger sequencer. Relatively low level of population differentiation was observed for *D. racemosum* ($F_{ST}<0.1$). In contrast, high level of differentiation ($F_{ST}>0.2$) was observed between Chichi-jima and Haha-jima populations of *D. lepidotum*, and between the two species. Although the true reason is still unknown, differences in genetic compositions of the founders of these species or adaptation to the local environment might have caused differentiation in some genes.
Most copy number variations (CNVs) are neutral, but some are deleterious and associated with various human diseases. CNVs are distributed non-randomly in vertebrate genomes, and it was recently reported that ohnologs, which are duplicated genes derived from whole genome duplication, are refractory to CNVs (Makino and McLysaght PNAS 2010). However, it is unclear what genomic factors affect the deleterious effects of CNVs and the biological significance of the biased genomic distribution of CNVs remains poorly understood. Here we show that non-ohnologs neighboring ohnologs are unlikely to have CNVs, resulting in ohnolog-rich regions (ORRs) in vertebrate genomes being CNV deserts (Makino et al. Nature Commun. 2013). Similarly, probable dosage-sensitive singletons that are unduplicated in all vertebrate lineages also repress CNVs of their immediate neighbours. In addition, long CNVs, prone to overlap genes, are less frequently observed near ohnologs. Our results suggest that the genomic location of ohnologs is a determining factor in the retention of CNVs and that dosage-balanced ohnologs are likely to cause the deleterious effects of CNVs in these regions. We propose that investigating CNV of genes in regions that are typically CNV deserts is an efficient means to find disease-related CNVs.

References:
- Makino, T and McLysaght, A. (2010) Ohnologs in the human genome are dosage balanced and frequently associated with disease. PNAS 107(20), 9270-9274
A Bioinformatics Method Suitable for Molecular Evolutionary Studies in the Era of Big Data: BLSOM

We have developed a bioinformatics tool "BLSOM (Batch-Learning Self-Organizing-Map)", which can analyze more than one million genomic sequences simultaneously, and therefore, allow us to gain efficiently a wide range of knowledge from the big sequence data derived from a wide range of species [1-3]. BLSOM with oligonucleotide (e.g. tetranucleotide) composition, which can cluster genomic sequence fragments (e.g. 1-kb sequences) according to phylotype by using only the oligonucleotide composition, has been successfully applied to the phylogenetic classification of a large number of metagenomic sequences [4,5]. Oligonucleotides such as tetra - heptanucleotides often represent motif sequences responsible for sequence-specific protein binding (e.g. transcription factor binding). Occurrences of such motif oligonucleotides should differ from the occurrences expected from the mononucleotide composition in the respective genome and may differ among genomic portions within a single genome. Actually, we have recently found that a pentanucleotide-BLSOM for the human genome can detect characteristic enrichment of many transcription-factor-binding motifs in pericentric heterochromatin regions [6].

Influenza virus is one of zoonotic viruses and shows clear host tropism. Important issues for evolutionary studies of influenza viruses are prediction of genomic sequence changes in the near future and surveillance of potentially hazardous strains. To characterize sequence changes in influenza virus genomes after invasion into humans from other animal hosts, we applied BLSOMs to analyses of oligonucleotide compositions in all genome sequences of influenza A and B viruses and found clear host-dependent clustering of the sequences [7]. Retrospective time-dependent directional changes of oligonucleotide compositions, which were visualized for human strains on BLSOMs, could provide predictive information about sequence changes in newly invaded viruses from other animal hosts. The strong visualization power of BLSOM also provides surveillance strategies for efficiently detecting potential precursors to pandemic viruses [8].

When we constructed BLSOMs for oligonucleotide compositions in fragment sequences (e.g. 100 kb) from a wide range of vertebrates, including coelacanth, we found that the sequences were clustered primarily according to species without species information. The characteristic oligonucleotide composition found for coelacanth was connected with the highest CG suppression among fishes, which was rather equivalent to that of tetrapods. This evident CG suppression in coelacanth should reflect molecular evolutionary processes of epigenetic systems including DNA methylation during vertebrate evolution. Sequence of a de novo DNA methylase (Dntm3a) of coelacanth was found to be more closely related to that of tetrapods than that of other fishes.

References:
2) Abe et al. (2006) Self-Organizing Map (SOM) unveils and visualizes hidden sequence characteristics of a wide range of eukaryote genomes. Gene 365, 27-34.
3) Abe et al. (2005) Novel Phylogenetic Studies of Genomic Sequence Fragments Derived from Uncultured Microbe Mixtures in Environmental and Clinical Samples. DNA Res. 12, 281-290.
5) Nakao et al. (2013) A novel approach, based on BLSOMs, to the microbiome analysis of ticks. ISME J. 7, 1003-1015.
8) Iwasaki et al. (2013) Nobel bioinformatics strategies for prediction of directional sequence changes in influenza virus genomes and for surveillance of potentially hazardous strains, BMC Infec Dis. 21, 386.
Next generation sequencer (NGS) technology is a powerful and useful for the study of molecular biology including evolution and genomics. The difficulty of the use of this technology is mostly for the part of data analysis process because of the size of output and wide range of applications. The cost for NGS is going down quick and the size of output going-up quick, and also new technologies are also coming up now.

The Cell Innovation Program aims to understand the cellular/life program by utilizing the technologies of next-generation sequencing and cellular imaging etc. by the use of the results and platform from the "Genome Network Project (GNP)" finished in FY2008. For this purpose, the Program works to establish "Sequencing Center" operating with high-speed sequencers and "Data Analysis Center" operating with the vast amount of diverse data, and promotes "Leading Research Projects" to investigate cellular functions making good use of these platforms and to develop innovative technologies. The National Institute of Genetics (NIG) as Data Analysis Center the promotion of the research activities that utilized a large quantity of genome information through development and maintenance of the data analysis platform for NGS. Now, we open the data analysis system (http://cell-innovation.nig.ac.jp/index_en.html) that is web-based, easy access and user friendly without any particular knowledge of informatics. We show the examples and results produced through this project.
Spermatogenesis defects are frequently observed in interspecific hybrids of Drosophila. As a first step to understand how the hybrid sterility arises, we characterized the expression profile of testes of pure species *Drosophila melanogaster* through combined analysis of transcriptomes by RNA-seq and proteomes by liquid chromatograph mass spectrometer. The obtained results suggest that the testis possesses many noisy transcripts.
Two behavioral races, M (for cosmopolitan) and Z (for Zimbabwe) of *Drosophila melanogaster* provide a great model to study the genetic basis of racial differentiation. When given a choice, females from the Zimbabwe race mate only with males from its congener whereas females from the cosmopolitan race mate readily with males from both races little discriminatorily. A series of genetic analyses showed that the Z/M behavioral isolation is mainly contributed by two major autosomes, and several fragments of the third chromosome are crucial in either male behavior or female preferences. However, very little was known about the genetic locus and the evolution of racial differentiation genes. To address this question, we have generated a reference genome of Z race by deep sequencing. By comparing to the published reference genome (M race), 0.8% of the sites have diverged between the two genomes. In addition, over 100 copy number variations were identified. Of which, we narrow down to around 10 candidate regions that may contribute to the M/Z racial differentiation by analyzing a small sample from the DPGP2 genomes. These results provide a general framework on mapping behavioral genes underlying racial differentiation in *D. melanogaster*. 
The extent to which mutations interact each other or epistasis dictates how protein evolves to new functions. Highly constrained fitness landscapes indicate that protein evolution is deterministic and predictable. Although importance of epistasis has been well recognized in protein as well as organismal evolution, molecular basis underpinning epistasis is poorly studied.

To explore ruggedness of fitness landscapes, we performed extensive laboratory enzyme evolution between two enzymatic activities, phosphotriesterase (PTE) and arylesterase (AE). First, we evolved PTE into AE with 22 rounds of directed evolution, resulting a complete function transition between PTE and AE with $>10^9$ specificity switch. Then, starting the newly evolved AE, we evolved AE back to PTE activity. The phenotype of enzyme (PTE activity) was highly reversible; with 12 rounds of directed evolution, we obtained a new efficient PTE on par with the wild-type PTE. However, on the sequence level, the evolution followed only partially the reversal trajectory; the newly evolved PTE (neoPTE) differ 28 mutations from the wild-type PTE (ancPTE). Intriguingly, many, if not most, mutations differ between neoPTE and ancPTE were deleterious when we tested on top of each background, indicating mutational incompatibility, and neoPTE are located on a distinct, incompatible peak from the original one on the fitness landscape.

Here I discuss molecular basis underpinning the incompatibility and irreversibility of the enzyme evolution. Detailed characterization of crystal structures of various enzymes revealed that the incompatibility is caused by rewiring intra-molecular interaction networks to displace and replace key components in the active sites. The active site configuration that degenerated for PTE activity during the forward evolution was reconciled by a different subset of mutations in the reverse evolution. Overall, our experimental evolution disclosed extensive epistasis between mutations. I also discuss a caveat to characterize mutations in nature; characterization of extant proteins may not lead to identify historically important mutations to switch functions.
We have discovered a large-scale reversal in the relative translational accuracy of codons across twelve fly species. Because the reversal involves pairs of codons which are read by the same genomically encoded tRNAs, we hypothesize, and show experimentally, that a tRNA anticodon modification—from guanosine to queuosine—has coevolved with these genomic changes. This modification is present in most organisms but its function remains unclear. Levels of queuosine modification in flies reflect bioavailability of the precursor nutrient queuine, which eukaryotes scavenge from the tRNAs of bacteria. These results reveal a strikingly direct link between nutrient intake, translational fidelity, and genome-wide exon recoding.
Cis-regulatory changes are shown to have large contribution to interspecific expression divergence. In order to analyze within- and between-population properties of gene regulation, we sampled alleles from inbred strains available as Drosophila Genomic Resource Panel, which was constructed from a natural population in North Carolina. In order to obtain robust allelic expression level estimates, and to focus on variation due to the cis-regulatory effect, following experimental scheme was used. First, each of those strains was crossed to a reference strain from West Africa to obtain F1 heterozygote samples. RNAseq was performed using head and body tissues from 100 adult F1 individuals of each sex. Then, the allelic expression ratios (AERs) of the reference and the sampled alleles against the total expression level in the F1 heterozygotes were calculated for each gene. The comparisons of within-population AER variance among DGRP alleles and between-population allelic expression difference (AER divergence) showed differences between sexes, which suggested that at a genomic level, gene expression is differentially regulated between sexes to some extent. The extent of differential regulation between sexes can be inferred from the level of correlation in AER divergences between male and female. We found that the strength of correlation differed between genes with male-biased expression and those with female-biased expression. Higher percentage of the variation in AER divergence among the genes was explained by intersexual expression differences in female-biased genes than that in male-biased genes. This indicates that conflict between male and female expression regulation is resolved to a certain degree in genes that belong to female-biased expression category. Whereas, such conflict resolution is less prominent in male-biased genes, which have been reported to show higher interspecific expression divergence. A strong positive correlation seen in male-biased genes suggests an abundance of mutations with pleiotropic effects on both sexes. These results collectively elucidate the sex-related properties in transcriptome evolution and suggest that intersexual conflict resolution in cis-regulation, for example through possessing independent modular enhancers, may not be a prerequisite for expression divergence to take place.
I will argue that evolution in Drosophila populations is more dynamic and driven much more by strong positive selection than usually assumed. Drosophila populations appear to be large enough and the rate to adaptive mutations in its populations high enough that rapid adaptation by soft sweeps is common even when generated by de novo mutations. In addition, Drosophila populations maintain abundant balanced genetic variation that allows them to adapt rapidly and cyclically to the seasonal shifts in the environment. I will discuss implications of these results for the understanding of evolutionary dynamics in this and other species with large population sizes.

References:
The fixation of beneficial alleles leaves genomic footprints characterized by a reduction of polymorphism at linked neutral sites and strong genetic differentiation among subpopulations. In contrast, for phenotypic evolution the effects of adaptation on the genes controlling the trait are little understood. Theoretical work on polygenic selection suggests that fixations of beneficial alleles causing selective sweeps are very rare relative to subtle allele frequency shifts among subpopulations (leading to the formation of clines along environmental gradients). We dissected an X-linked QTL underlying variation in chill coma recovery time (CCRT), a proxy for cold tolerance, in *D. melanogaster* from temperate (European) and tropical (African) regions. This reduced the 6-Mb long QTL to 124 kb. Against expectation, however, this region co-localizes with a strong selective sweep in the European population. Of the genes overlapping with this sweep, none seems to affect CCRT. However, one gene (brinker) in the immediate neighborhood (but outside) of the sweep showed significant differences in gene expression between the African and European populations. A possible cause of this variation may be a polymorphic indel upstream of brinker forming a latitudinal cline.
Eukaryotic genomic DNA is assembled into nucleosomes. Nucleosome associated sequences exhibit characteristic $\approx 10.4$ bp periodicities of AA, TT and GC dinucleotides. While these and other sequence properties have structural and in vitro functional interpretations, evidence of phenotypic effects attributable to allelic variants associated with the dinucleotide patterns has been lacking. Biased mutation has remained a tenable explanation.

To investigate the evolutionary consequences of the interactions between genomic polymorphisms and the highly conserved histone core we isolated nucleosomal DNA sequences from *Drosophila melanogaster* early embryos. We then examined the patterns of genomic variation within that species and divergence from its most recent common ancestor with simulans in the alignment nucleosomal regions. Consistent with predictions of models of equilibrium weak systematic selection, mutation and stochastic forces, the nucleosomal patterns of divergence are toward the “preferred” bases. While this could by itself be attributed to biased mutation, our observations of strong periodicities in the site frequency spectra favor a central role for natural selection in the maintenance of these ubiquitous and conserved features.
Neutral Mutations and Purifying Selection Dominate Genome Evolution

Naruya Saitou; Division of Population Genetics, National Institute of Genetics

Ever since proposal of Neutral Evolution Theory by Kimura and King & Jukes, importance of selectively neutral mutations was increasingly clear, and nowadays it is established that majority of mutations accumulated through eons of evolution are neutral ones. Evolutionary changes are thus mostly attributed to selectively neutral mutations and their fixation by chance. How about natural selection? Charles Darwin was essentially wrong in considering "positive" natural selection as main cause of evolutionary changes. However, we also have "negative" selection, and this type of selection is much more important than "positive" one. Existence of this conservative process to keep status quo, also called as "purifying selection", was well known before Charles Darwin and Alfred Russell Wallace proposed "positive" natural selection. I would like to summarize our recent studies on neutral mutations which shaped GC content heterogeneity of vertebrate genomes, and those on purifying selections which show lineage-specific characteristics possibly through lineage-specific conserved non-coding sequences in eukaryote genomes as well as introduction of my recently published book (Saitou 2014).

Reference:
Co-evolution among Ribosomal Constituents Sustained by Assembly Constraints

Saurav Mallik⁴, Hiroshi Akashi⁵, Sudip Kundu⁶; 1. Center of Excellence, TEQIP Phase-II, University of Calcutta, 2. National Institute of Genetics

Ribosome biogenesis is a central and essential cellular process, expensing a significant fraction of the entire cellular energy budget⁴. In prokaryotes, the cytoplasmic assembly of ribosomal subunits requires the coordinated association of 55 ribosomal proteins and three ribosomal RNAs (rRNAs), facilitated by about 20 cofactors⁵. Cooperativity between macromolecular constituents is an essential factor that helps driving the formation of macromolecular complexes⁴. It is acquired either through the mutually supportive physical interaction between the two constituents, or through two different conformations of their common binding partner⁴. During ribosome assembly, both kinds of cooperativities are observed. On the other hand, co-evolution among biomolecules or individual residue pairs refers to the correlated sequence changes for maintaining functional interaction⁶. Identifying the co-evolving residue-pairs of the entire small ribosomal subunit, followed by a network analysis, here we show that protein pairs under kinetic cooperativity preferably co-evolve with each other. Furthermore, stronger co-evolutionary preference is observed in cases of direct cooperativity between proteins (primary to secondary and secondary to tertiary) compared to that of the indirect cooperativity (primary to tertiary). This leads us to conclude that protein-rRNA structural segments constrained by assembly process are associated with strong co-evolution amongst themselves.

References:
Evolutionary Dynamics of Conserved Noncoding Sequences

Babarine Isaac Adeyemi\textsuperscript{1,2} and Naruya Saitou\textsuperscript{1,2}; 1. Department of Genetics, SOKENDAI, 2. Division of Population Genetics, National Institute of Genetics

Conserved noncoding sequences (CNSs) of vertebrates are considered to be closely linked with protein-coding gene regulatory functions. Since gene regulation has been suggested to play a major role in morphological evolution, I examined the abundance and genomic distribution of CNSs in four mammalian orders: primates, rodents, carnivores, and cetartiodactyls. Using conservation level of coding genes to determine the functionality, I set two identity thresholds for CNSs with at least 100 base pairs. I then investigated the dynamics of evolution of CNSs across lineages and ages. The abundance of CNSs retrieved varied among lineages, with primates and rodents having highest and lowest number of CNSs, respectively, whereas carnivores and cetartiodactyls had intermediate values. Evolution of new CNSs as well as retention of ancestral CNSs contribute to the differences in abundance. These CNSs cover 1.3–5.5% of the mammalian genomes and have signatures of purifying selection. Comparisons of CNSs of different ages reveals that older CNSs tend to have stronger signatures of selective constraints than the younger ones. The genomic distribution of CNSs is dynamic with higher proportions of rodent and primate CNSs located in the introns compared with carnivores and cetartiodactyls. In fact, 19% of orthologous single-copy CNSs between human and dog are located in different genomic regions. If CNSs can be considered as candidates of gene expression regulatory sequences, heterogeneity of CNSs among the four mammalian orders may have played an important role in creating the order-specific phenotypes. Fewer CNSs in rodents suggest that rodent diversity is related to lower regulatory conservation. With CNSs shown to cluster around genes involved in nervous systems and the higher number of primate CNSs, our result suggests that CNSs may be involved in the higher complexity of the primate nervous system. Taken together, this study gives further credence to the hypothesis that gene regulation plays an important role in morphological evolution.

Reference:
Disease susceptibility can arise as a consequence of adaptation to infectious disease. Recent findings have suggested that higher rates of chronic kidney disease (CKD) in individuals with recent African ancestry might be attributed to two risk alleles (G1 and G2) at the serum-resistance-associated (SRA)-interacting-domain-encoding region of APOL1. These two alleles appear to have arisen adaptively, possibly as a result of their protective effects against human African trypanosomiasis (HAT). In order to explore the distribution of potential functional variation at APOL1, we studied nucleotide variation in 187 individuals across ten geographically and genetically diverse African ethnic groups with exposure to two Trypanosoma brucei subspecies that cause HAT. We observed unusually high levels of nonsynonymous polymorphism in the regions encoding the functional domains that are required for lysing parasites. Whereas allele frequencies of G2 were similar across all populations (3%–8%), the G1 allele was only common in the Yoruba (39%). Additionally, we identified a haplotype (termed G3) that contains a nonsynonymous change at the membrane-addressing-domain-encoding region of APOL1 and is present in all populations except for the Yoruba. Analyses of long-range patterns of linkage disequilibrium indicate evidence of recent selection acting on the G3 haplotype in Fulani from Cameroon. Our results indicate that the G1 and G2 variants in APOL1 are geographically restricted and that there might be other functional variants that could play a role in HAT resistance and CKD.

References:

Poster Presentation: P20

Aging Effects and Sexual Differences in Human Gene Expression in Comparison with Non-coding Genes

Maiko Narahara; Graduate School of Medicine, Kyoto University

Background
Males and females are different physically and behaviorally in human; and importantly, nearly all diseases have different prevalence and course between the sexes. Previous studies for sexual differences have primarily focused on contributions of gonadal tissues and genes on the sex-chromosomes. Recent studies, however, suggested that autosomal genes in non-gonadal tissues also contribute to sexual dimorphisms [1,2].

Aging phenotypes also widely range, and aging is also important for disease susceptibility and course as well. Several studies were carried out for aging gene expression in human. However, further study is still worthwhile because a large-scale study focused on female samples [3], and also because it is still controversial whether or not aging gene expression is tissue specific.

In addition, previous studies focused on coding genes; and influences of aging and sex on non-coding RNAs have not been studied.

Methods
We obtained gene expression profiles of whole blood from 298 Japanese individuals (102 males and 196 females) by using microarray (Agilent). Our gene expression profiles include both mRNAs and lincRNAs. We analyzed for changes with age by using a non-parametric regression and sexual differences with correction for age, and compared the influence of these factors on mRNAs to lincRNAs.

Results
lincRNAs were more variable than mRNA when variances of lincRNAs and mRNAs with the same median expression levels were compared. Aging influenced the global expressions (75th percentile) of mRNA and lincRNA to the opposite directions; when mRNA increases, lincRNA decreases, and when mRNA decreases, lincRNA increases.

We identified 2 and 3,910 aging transcripts for males and females, respectively, at FDR < 5%. We performed clustering of the aging transcripts for females, and identified two clusters. Cluster 1 was comprised of 1,083 transcripts increasing with age; and 83.9% of aging lincRNAs were categorized into this cluster. Cluster 2 was comprised of 2,827 transcripts decreasing with age; and 93.7% of transcripts in this cluster were mRNAs.

At FDR < 5%, we identified 5,888 transcripts differentially expressed between the sexes, 44% of which were expressed at the higher levels for males (male-biased). Differentially expressed mRNAs were almost equally often male-biased or female-biased (2,270 vs. 2,101), whereas differentially expressed lincRNAs were predominantly female-biased (97 vs. 663).

Conclusion
We found that 10.6% of tested transcripts significantly changed with age for females, and that 16.0% were differentially expressed between the sexes. Our analyses suggest that mRNAs and lincRNAs tend to have the opposite characteristics in many aspects.

References:
Lineage-specific Genome Evolution in the *Drosophila melanogaster* subgroup

Neha Mishra and Hiroshi Akashi; National Institute of Genetics

Evolutionary forces such as mutation, genetic drift, natural selection, and recombination contribute to long-term evolution but the time-scale and magnitude of their variation is not well understood. We examine lineage-specific genome evolution in the *Drosophila melanogaster* subgroup and test its causes. Patterns of synonymous codon usage bias can be used as a model to study the effects various evolutionary forces acting on a genome. The “major codon preference” model of codon usage bias suggests that selection favors major codons whereas mutational pressure and genetic drift allow minor codons to persist. This balance between weak forces makes codon bias sensitive to changes in evolutionary parameters. Hence, we examined the changes in the codon bias patterns in >5000 genes from the 7 *Drosophila melanogaster* subgroup species. We employed existing genome data for five species and added data for two of the *Drosophila melanogaster* subgroup species. We performed Next-Generation RNA sequencing on the *Drosophila tessieri* and *yakuba* transcriptomes. We describe a protocol for gene annotation of the RNA-seq data using the available data from the sequenced species. We inferred ancestral states using maximum likelihood approaches that account for both base composition bias and non-stationarity and assigned substitutions to 10 lineages. Several lineages showed strong departures from the equilibrium codon bias. *Drosophila sechellia* has an island population with little detected polymorphism. The substitutions that occurred in the lineage leading to *Drosophila sechellia* were biased towards unpreferred fixations. Variability among genes is consistent with the decline in selection intensity (selection coefficient scaled to population size) and no change in the mutation bias. Another species, *Drosophila orena*, which may also have experienced historically small population sizes, showed patterns consistent with a decline in the selection intensity but also a change in background substitution patterns. These findings suggest that the magnitudes of forces governing base composition at synonymous sites may have varied frequently in a lineage-specific manner in the *Drosophila melanogaster* subgroup and may need to be taken into account when testing evolutionary mechanisms at other classes of sites.
Effects of Unrealistic Model Usage on Accuracy of Ancestral Inference

Tomotaka Matsumoto¹, Ziheng Yang², Hiroshi Akashi¹; ¹. National Institute of Genetics, 2. University College London

Reliable ancestral inference is critical in many molecular evolution approaches. Previous work has shown that maximum parsimony can cause considerable error in the inference of past codon usage bias in the *Drosophila melanogaster* subgroup, and even maximum likelihood approaches (under the HKY85 substitution model) can be strongly biased. In the present study, we employed more general General Time Reversible (GTR) substitution models including some that allow non-homogeneous base composition. We examined whether increased generality in the assumed model increases the accuracy of ancestral inference. We conducted computer simulations to obtain nucleotide sequences that emulate the tree topology and distances among six species in *D. melanogaster* subgroup. In these scenarios, mutation bias and natural selection are determinants of codon usage bias and parameter values can differ among lineages. Using simulated sequence data, we compared the accuracy of ancestral inference under several models implemented in PAML.

The GTR model with non-homogeneous base composition performed well when the simulation scenario exactly matched the substitution model. However, compared to a stationary HKY85 model, the “more realistic” stationary GTR and non-stationary HKY85 models did not improve inference accuracy and, in some cases, showed greater inference error.

We also examined the accuracy of the parameter estimation and found that even when the inconsistency between assumed model and simulation parameterization appears to be small, estimates of branch length and equilibrium base composition can show considerable bias. In addition, if the assumed model exactly matches the simulation parameterization, small sample sizes can cause erroneous inference because of the larger number of parameters in non-homogeneous models. These results demonstrate the trade-off between model generality and sample size; inference reliability is reduced for both overly simplistic models and for complex models applied to limited sample sizes.
Slightly Deleterious and Adaptive Evolution in mtDNA of Cichlid Fishes in East Africa

Kazumasa Shirai, Nobuyuki Inomata, Shinji Mizoiri, Mitsuto Aibara, Yohey Terai, Norihiro Okada, and Hidenori Tachida; 1. Graduate School of Systems Life Sciences, Kyushu University, 2. International College of Arts and Sciences, Fukuoka Women’s University, 3. Sekisui Chemical Tanzania Ltd., 4. Foundation for Advancement of International Science, 5. The Graduate University for Advanced Studies, 6. Department of Life Sciences, National Cheng Kung University, 7. Department of Biology, Faculty of Sciences, Kyushu University

The nearly neutral theory asserts that fixations of slightly deleterious mutations occur ubiquitously in molecular evolution. When population size is large, slightly deleterious mutations are removed from the population effectively by natural selection while slightly deleterious mutations may be fixed in the population by genetic drift when population size is small. Because nonsynonymous substitutions are more strongly affected by natural selection than synonymous substitutions, fixations of slightly deleterious mutations are expected to increase the nonsynonymous substitution rate. Indeed several studies have revealed that the dN/dS ratios in the species with smaller population size are higher than those in the species with larger population size. In the present study, we examine this prediction of the nearly neutral theory by studying fishes belonging to Cichlidae in African lakes. Cichlidae is known as one of the most species-rich family in vertebrates. Especially, there are a large number of species of cichlid fishes in East African rivers and three African Great Lakes, Lakes Victoria, Tanganyika and Malawi. It has been suggested that Lake Victoria and East African rivers were seriously affected by past desiccations. Such aridity might have decreased population sizes of cichlid fishes in Lake Victoria and East African rivers, thereby fixing slightly deleterious mutations in the cichlid populations living there. On the other hand, geographical studies have shown that Lake Tanganyika and Lake Malawi have kept fairly constant water levels. The stable environments in these two lakes are considered to have kept the population size of cichlid fishes living there large enough to remove slightly deleterious mutations. Therefore, the difference in the stability of population size could have lead to differences of $\omega$ between the cichlid fishes in Lake Victoria and rivers on one hand and those in Lakes Tanganyika and Malawi on the other. We estimated $\omega$ and compared them at five nuclear and 13 mitochondrial protein coding genes of cichlid fishes in the three lakes and a river fish using maximum likelihood methods. We found the lineages leading to cichlids in Lake Victoria and the river had significantly higher $\omega$ at the mitochondrial genes. Moreover, positive selection was suggested at several codons in mtDNA genes in the lineage leading to the cichlids of Lake Victoria. Our results indicated somewhat unusual molecular evolution has taken place in the cichlids of Lake Victoria and the river at mtDNA genes, whose nonsynonymous sites are generally conserved.

References:
The dynamics of immune system evolution are impacted by temporal and spatial fluctuations in the microbial and parasitic environment. Accordingly, it is a prime system for a detailed investigation of both long-term evolution and current population adaptation. Leveraging full sequence data from 84 inbred *Drosophila melanogaster* lines, we dissect patterns of evolution along the *D. melanogaster* lineage and compare it to spatially variable evolution within five extant populations (Beijing, New York, the Netherlands, Tasmania and Zimbabwe). We have assembled a set of 348 immune genes, matched with a set of over 1200 control genes, allowing us to conservatively control for the effects of demography and local genetic environment. Applying a suite of statistical tests, we find evidence for strong spatially variable selection within some individual genes. These genes serve a range of functions including microbe recognition (NimC3, Sr-CIII), melanization (proPO59, Alk), and viral defense (CHKov1, CHKov2). In some instances, there is evidence of selection in only a single population while other genes show signs of directional selection on a more global scale. In addition, we have tested for weak polygenic selection whose effects may be detected in aggregate across a pathway even when no single gene presents a significant signature of selection. Immune pathways are well characterized in *D. melanogaster* and are known to differ in types of stimuli, breadth of effect, tissue-specificity, and degree of constitutive activity. By comparing evolutionary signatures within these pathways as a whole, we have inferred the strength and direction of selection that is acting on specific compartmentalized immune functions. On both long and short timescales, we find evidence for polygenic selection, suggesting that immune adaptation can result from many changes throughout a single pathway. For example, certain pathways show evidence of heightened pairwise $F_{st}$ (melanization, epithelial defenses) or higher rates of adaptive evolution (parasitoid defense, viral defense). Overall, this multi-level examination of immune function suggests that viruses and parasitoid wasps exert disproportionally strong selection pressures on fly immune defenses on both short and long timescales.
Inference of Population Structure in *Taxodium distichum*, a Coniferous Tree in North America, Based on Amplicon Sequence Analysis

Yuka Ikezaki, Yoshihisa Suyama, Beth A. Middleton, Yoshihiko Tsumura, Kousuke Teshima, Hidenori Tachida and Junko Kusumi; 1. Laboratory of Evolutionary Genetics, Department of Systems Life Sciences, Graduate school of systems life sciences, Kyushu university, 2. Field Science Center, Graduate School of Agricultural Science, Tohoku University, 3. USGS National Wetlands Research Center, 4. Forestry and Forest Products Research Institute, 5. Department of Biology, Faculty of Sciences, Kyushu University, 6. Department of Environmental Changes, Faculty of Social and Cultural Studies, Kyushu University

*Taxodium distichum* is a long-lived coniferous tree widely distributed in southeast North America. It prefers wetlands and has two varieties, bald-cypress and pond-cypress in the United States. These two varieties have clearly different habitats and morphological characteristics. From the results of previous studies, both geographical (between the Florida and Mississippi River regions) and variety specific differentiations were suggested (Tsumura et al., 1999; Lickey and Walker, 2002; Kusumi et al., 2010; Tanaka et al., 2012). The distributions of the two varieties overlapped in the Florida and southern Mississippi River regions, hence the geographical and variety specific differentiations are related in a complex way. Recently, the Next Generation Sequencing (NGS) approach along with amplicon sequence analyses enables us to obtain sequences of many individuals at many loci fairly easily. To infer the detail of population structure of *Taxodium*, we collected 96 individuals from the two varieties in the Mississippi River, Texas and Florida regions and determined their sequences at 48 nuclear loci using this approach with a GS FLX 454 sequencer. To examine the level of differentiation at neutral loci, we randomly selected genes from an annotated Expressed Sequence Tag (EST) library of Taxodium. The results of a Bayesian clustering analysis and estimated $F_{ST}$ values suggested that the populations in the Mississippi River and Texas regions had similar genetic constitution and they are differentiated from the populations in the Florida region though the level of the differentiation was low. The variety specific differentiation was also found, however it was attributed to five loci showing much higher $F_{ST}$ values than the other loci. After removing these loci, the Bayesian clustering analysis showed only geographical differentiation. Finally, analyses of the geographical and variety subdivisions assuming the isolation with migration model suggested that the divergence time of the variety subdivision was much older than that of the geographical subdivision (4.6 MYA and 0.2 MYA, respectively). However, a high level of migration from the bald-cypress populations to the pond-cypress populations ($2Nu = 10.0$) has reduced the level of the variety specific differentiation. On the other hand, the level of migration between the Mississippi-Texas region and Florida region was low (from Mississippi-Texas populations to Florida populations: 0.28; from Florida populations to Mississippi-Texas populations: 0.30), so the level of geographical differentiation seemed to have been increasing.

References:

Sex chromosomes turn over rapidly in some taxonomic groups, where closely related species have different sex chromosomes. Although there are many examples of sex chromosome turnover, we know little about the functional roles of sex chromosome turnover in genomic evolution. The sympatric pair of Japanese threespine stickleback (*Gasterosteus aculeatus*) provides an excellent system to address these questions: the Japan Sea species has a neo-sex chromosome system resulting from a fusion between an ancestral Y chromosome and an autosome, while the sympatric Pacific Ocean species has a simple XY sex chromosome system. Furthermore, previous quantitative trait locus (QTL) mapping demonstrated that the Japan Sea neo-X chromosome contributes to phenotypic divergence and reproductive isolation between these sympatric species. To investigate the genomic evolution involved in sex chromosome turnover, we conducted whole genome sequencing of males and females of both the Japan Sea and the Pacific Ocean species. No substantial degeneration has yet occurred on the neo-Y chromosome, but the nucleotide sequence of the neo-X and the neo-Y has started to diverge, particularly at regions near the fusion. Non-synonymous substitutions have also accumulated on the neo-X and/or neo-Y chromosomes. Transcriptome analysis indicated that the neo-sex chromosomes also harbor an excess of genes with sex-biased expression, which is correlated with nucleotide divergence between the neo-X and the neo-Y. These results suggest that the divergence between neo-X and the neo-Y chromosomes might promote functional evolution of genes located near the fusion. In contrast to the divergence between the neo-X and the neo-Y, genomic regions of higher sequence divergence between species, genes with divergent expression between species, and QTL for inter-species phenotypic differences were found not only at the regions near the fusion, but also at other regions along the neo-X chromosome. Recombination reduction might explain this pattern of rapid genetic divergence between the species in the neo-X chromosomes. Overall, our results show that sex chromosome turnover can promote accumulation of substitutions causing genomic divergence between species.
Detecting the genetic targets of natural selection is a fundamental goal in evolutionary genetics. Population genomic data offers the opportunity to identify numerous outlier loci based on genome wide patterns, and these loci may provide insights into the genetic basis of adaptive evolution. Here we perform three genome scans for selection on more than 100 fully sequenced genomes of *Drosophila melanogaster*, mainly from sub-Saharan (ancestral range) populations. First, we search for selective sweeps within a single population sample from Rwanda. Second, we compare variation in Rwanda to a population from France and look for loci that may have been involved in the adaptation of this species to temperate non-African environments. Finally, we identify loci that show elevated genetic differentiation among African populations that were sampled from a variety of environments. Our work contributes toward a genomic and geographic atlas of recent positive selection in the *D. melanogaster* genome.

Reference:
Recent studies have shown that adaptation from \textit{de novo} mutation often produces so-called soft selective sweeps, where multiple adaptive mutations of independent origin sweep through the population at the same time. Population genetic theory predicts that soft sweeps should be likely if the product of the population size and the mutation rate towards the adaptive allele is sufficiently large, such that multiple adaptive mutations can establish before one has reached fixation. However, it remains unclear how demographic processes affect the probability of observing soft sweeps. Here we extend the theory of soft selective sweeps to realistic demographic scenarios that allow for changes in population size over time. We first show that population bottlenecks can lead to the removal of all but one adaptive lineage from an initially soft selective sweep. The parameter regime under which such 'hardening' is likely is determined by a simple heuristic condition. We further develop a generalized analytical framework, based on an extension of the coalescent process, for calculating the probability of soft sweeps under arbitrary demographic scenarios. Two important limits emerge within this analytical framework: If population size fluctuations are fast compared to the duration of the sweep, the likelihood of soft sweeps is determined by the harmonic mean of the variance effective population size, estimated over the duration of the sweep. By contrast, in the slow fluctuation limit the likelihood of soft sweeps is determined primarily by the instantaneous variance effective population size at the onset of the sweep. We show that as a consequence of this finding the probability of observing soft sweeps becomes a function of the strength of selection. Specifically, in species with sharply fluctuating population size, strong selection is more likely to produce soft sweeps than weak selection. Our results highlight the importance of accurate demographic estimates over short evolutionary timescales for understanding the population genetics of adaptation from \textit{de novo} mutation.
The pattern of molecular evolution of imprinted genes is controversial and the entire picture is still to be unveiled. Recently, a relationship between the formation of imprinted genes and gene duplication was reported in genome-wide survey of imprinted genes in Arabidopsis thaliana. Because gene duplications influence the molecular evolution of the duplicated gene family, it is necessary to investigate both the pattern of molecular evolution and the possible relationship between gene duplication and genomic imprinting for a better understanding of evolutionary aspects of imprinted genes. In this study, we investigated the evolutionary changes of type I MADS-box genes that include imprinted genes by using relative species of Arabidopsis thaliana (two subspecies of A. lyrata and three subspecies of A. halleri). A duplicated gene family enables us to compare DNA sequences between imprinted genes and its homologs. We found an increased number of gene duplications within species in clades containing the imprinted genes, further supporting the hypothesis that local gene duplication is one of the driving forces for the formation of imprinted genes. Moreover, data obtained by phylogenetic analysis suggested “rapid evolution” of not only imprinted genes but also its closely related orthologous genes, which implies the effect of gene duplication on molecular evolution of imprinted genes.

Reference:
• PLOS ONE (2013) 5;8(9):e73588. doi:10.1371/journal.pone.0073588
Understanding how genetic variation is translated into phenotypic variation, and how this translation depends on the environment is fundamental to our understanding of complex traits and the evolution, and has enormous practical implications. In particular, there has been much speculation about the existence of buffering, or “canalizing” effects that serve to mask the phenotypic effects of genetic variation under “normal” environmental conditions, but reveal them under abnormal conditions. We are trying to reveal the buffering mechanisms using a systematic approach with multi-layered omics data.

As a model study, we are focusing on buffering mechanisms of flowering time of Arabidopsis thaliana. We collected several phenotypes of 143 Swedish strains as well as RNAseq and DNA methylation data as intermediate phenotypes under two growth temperatures (10, 16°C). Under 16°C, variance of the phenotype across strains was obvious and heritability of the phenotype was high, although the variance under 10°C was much smaller than that of 16°C. To figure out the causal loci regulating G x E, we conducted genome wide association study (GWAS) and network analysis based on correlation between the flowering time phenotype and intermediate phenotypes. GWAS detected polymorphisms on well-known flowering genes such as FLOWERING LOCUS C and VERNALIZATION INSENSITIVE 3 that have dominant effects on flowering time. Furthermore, network analysis predicted minor genetic effects and unmapped effects that affect flowering time under warmer temperature. We propose the model of buffering mechanisms of flowering time in this presentation.
Gene Expression Profiling in Nervous System of Sea Urchin Larvae and Insights into the Evolution of Central Nervous System

Sonoko Kinjo¹, T. Okayama², K. Ikeo¹, and T. Gojobori¹; ¹. National Institute of Genetics, ². BITS Co., Ltd, ³. KAUST

With the aim of understanding an evolutionary process of central nervous system (CNS) from the viewpoint of gene expression, we conducted a RNA Tag-seq analysis of the apical organ in sea urchin larvae. Echinoderms including sea urchin that are the closest relatives to chordates, lack an apparent organ or tissue of CNS. Instead, they are said to have a distinctive “pentaradial” nervous system. The molecular developmental studies suggested that gene expression patterns in development of vertebrate brain is remarkably conserved in arthropods, implying that such gene expression patterns may have already existed in the last common ancestor of all bilaterians. This indicates that the expression of those genes were probably lost or considerably changed after the split of echinoderm from chordates, along with the drastic changes in morphology. Echinoderm larvae, however, retain a bilateral body plan with large ganglia at the anterior end of the body, called as apical organ. Indeed, the morphological homology between apical organ and chordate brain has been discussed repeatedly for years. Collectively, this structure is a strong candidate for a trace of central nervous system.

In this study, to understand an evolutionary relationship between apical organ and CNS, we conducted a RNA Tag-seq analysis and investigated expressions of CNS patterning genes in sea urchin larvae. We obtained more than 16 million sequence tags from two tissues with and without apical organ, and then computed the difference in gene expression level between the tissues. In the tissue with apical organ, forebrain-specific homeobox genes Six3, Nkx2.1, and their downstream genes are highly expressed as well as retinal homeobox (Rx) required for eye formation. In addition, among orthologs of CNS patterning genes, forebrain patterning genes are expressed at high level in apical organ-containing tissue while midbrain and hindbrain patterning genes are not. These results indicate that echinoderm apical organ could be comparable to forebrain in the viewpoint of gene expression. Thus, we conclude that the expression of midbrain/hindbrain patterning genes may have been lost or reduced in echinoderm larvae, although other possible scenarios should also be considered.
The coleoid cephalopods such as squids and octopuses are known to possess an elaborated brain, by far the most highly evolved among invertebrates. It appears that the brain confers on the cephalopods unexpectedly superb learning abilities such as tactile reflex and visual discrimination. In fact, a number of neurons in the brain exceed 30-100 million that are almost equal to that of a small mammal such as mice. These facts lead to an immediate question of how their brain was evolutionarily attained and what kind of genetic changes have taken place during evolution.

To answer the above-mentioned questions of significance, we have chosen the Japanese pygmy squid (Iodosepius paradoxus) that has an advantage of the smallest genome (2.2G bases) among the cephalopods. We also studied Nautilus pompilius, that is one of the oldest cephalopods and has a simpler brain (with 13 discernible lobes) than in squids (33-37 lobes). However, the nervous system is vastly more complex than that of any non-cephalopod molluscs and shows intermediate state of the brain evolution.

In practice, we conducted RNA-Seq analysis of tissues and embryonic samples of the two cephalopods. Using the RNA-Seq data and the genome sequence, we determined complete gene set of the cephalopods. Our comparative analysis indicated the following four important points of evolutionary features: (1) Genes of the cephalopods that are conserved across animal phyla tended to be expressed constitutively throughout the developmental stages, (2) on the other hand, cephalopod-specific genes, particularly their duplicated genes, are expressed exclusively at the latest embryonic stage, (3) we also found that those late stage-specific genes showed a brain-specific pattern of gene expression, and (4) some gene families such as vasopressin/neurophysin have already duplicated in Nautilus, suggesting the genes emerged before their divergence that is dating back to early Paleozoic era.

From these observations, we proposed that the tissue- and stage-specific genes have played crucial roles in evolutionary formation of novel types of neural circuits in the cephalopod brain. It indicates that genomic and transcriptomic information of the cephalopods will provide us with fundamental platform of evodevo studies of the broad animal phyla in the coming years.
DNA Methylation and Evolution of Duplicate Gene

Thomas E. Keller and Soojin V. Yi; Georgia Institute of Technology

The evolutionary mechanisms underlying duplicate gene maintenance and divergence remain highly debated. Epigenetic modifications, such as DNA methylation, may contribute to duplicate gene evolution by facilitating tissue-specific regulation. However, the role of epigenetic divergence on duplicate gene evolution remains little understood. Here we show, using comprehensive data across ten diverse human tissues, that DNA methylation plays critical roles in several aspects of duplicate gene evolution. We first demonstrate that duplicate genes are initially heavily methylated, before gradually losing DNA methylation as they age. Within each pair, DNA methylation divergence between duplicate partners increases with evolutionary age. Importantly, tissue-specific DNA methylation of duplicates correlates with tissue-specific expression, implicating DNA methylation as a causative factor for functional divergence of duplicate genes. These patterns are apparent in promoters but not in gene bodies, in accord with the complex relationship between gene body DNA methylation and transcription. Remarkably, many duplicate gene pairs exhibit consistent division of DNA methylation across multiple, divergent tissues: for the majority (73%) of duplicate gene pairs, one partner is always hyper-methylated compared to the other. This is indicative of a common underlying determinant of DNA methylation. The division of DNA methylation is also consistent with their chromatin accessibility profiles. Moreover, at least two sequence motifs known to interact with the Sp1 transcription factor mark promoters of more hypo-methylated duplicate partners. These results demonstrate critical roles of DNA methylation, as well as complex interaction between genome and epigenome, on duplicate gene evolution.
Sex chromosomes differ from autosomes at their genome, transcriptome, and epigenome, yet the X and Y share a common evolutionary origin. The Drosophila Y chromosome is gene poor and associated with silencing heterochromatin. The X, in contrast, is enriched with activating chromatin marks and hyper-transcribed (i.e. dosage compensated), which is thought to be an adaptive response to gene decay and silencing on the Y. While ultimately heterochromatin formation and dosage compensation have opposite effects on gene expression, their formation has intriguing parallels. Both are initiated at specific DNA sequences by selective recruitment of chromatin-modifying proteins, spread along the chromosome and direct the modification of histones, allowing large genomic regions to adopt either a transcriptionally silent (heterochromatin) or hyper-transcribed (dosage compensation) chromatin structure. How sex chromosomes have evolved to become that way, and what genomic changes drive their dramatically different epigenetic makeup, however, has remained a mystery. By studying the genome, epigenome and transcriptome of a species with a very recently evolved sex chromosome (the neo-X and neo-Y of *D. miranda*), we recapitulate how both dosage compensation and heterochromatin formation evolve in Drosophila and establish several novel and important principles governing the evolution of sex chromosomes.

References:

The genetic basis governing phenotypic evolution has been an important problem to attract wide attention. It has been found that genomes in various organisms acquire novel genetic elements as sources of functional and phenotypic diversity, including new genes that originated in recent evolution. Substantial progress has been made recently in understanding the evolution and phenotypic effects of new genes, leading to an emerging picture that new genes can rapidly evolve indispensable roles in fundamental biological processes, including development, reproduction, brain function and behavior.

An inescapable inference is that the genetic systems in control of these important phenotypes in various organisms are, per se, species-specific or lineage-specific, despite previous observed conservative genetic elements. Furthermore, significant progress has been also made to understand the evolution of gene functions through ancient functions-based models or functional innovation by neofunctionalization. New genes provided evidence of functional innovation, suggesting that gene functions have been evolving, besides the reuse or co-option of ancient functions. The recent studies of new genes yielded fresh insights into our broad understanding of biological diversity at refined resolution.

References:
Hotspots of non-allelic homologous recombination (NAHR) have a crucial role in creating genetic diversity and are also associated with dozens of genomic disorders. Recent studies suggest that many human NAHR hotspots have been preserved throughout the evolution of primates. NAHR hotspots are likely to remain active as long as the segmental duplications promoting NAHR retain sufficient similarity. Here, we will show that an evolutionary model of SDs that incorporates the effect of gene conversion best explains the data on copy number variants (CNVs) and genomic disorders, and discuss the implications of this finding.
Patterns of molecular evolutionary change on the Y chromosome provide unique opportunities to dissect the forces that have acted on this unique genetic element. The partial or complete lack of recombination and male-limited transmission of the Y set it apart from the X and autosomes. Male-limited transmission has consequences for both the effective population size of the Y and the efficacy of selection on male-beneficial mutations on the Y relative to elsewhere in the genome. Together, these defining features of the Y chromosome are expected to influence rates and patterns of molecular evolution on the Y as compared with X-linked or autosomal loci. Sequence data and analysis from eleven genes in nine Drosophila species reveal insights into the efficacy of natural selection on the Drosophila Y relative to the rest of the genome. Drosophila is an ideal system for assessing the consequences of Y-linkage for molecular evolution in part because the gene content of Drosophila Y chromosomes is highly dynamic. Gene features, such as codon usage bias, that are generally under weak selection, can be strongly impacted by gene movements from the Y to X (or autosomes) and back. Not surprisingly, results indicate that the efficacy of natural selection at weakly selected sites is reduced on the Y chromosome. In contrast, purifying selection on the Y chromosome for strongly deleterious mutations does not appear to be compromised. In addition, we find evidence of recurrent positive selection for four of the 11 genes studied here. Our results thus highlight the variable nature of the mode and impact of drift and natural selection on the Drosophila Y chromosome.
Oral Presentation: S2

Understanding Features of Recombination and Historical Demography from Linkage Disequilibrium in Single Individuals

Michael Lynch; Department of Biology, Indiana University

Although the analysis of linkage disequilibrium (LD) plays a central role in many areas of population genetics, the sampling variance of LD is known to be very large both with respect to numbers of nucleotide sites and individuals sampled. It will be shown, however, that a simple genome-wide analysis of the distribution of heterozygous sites within a single diploid genome can yield highly informative patterns of LD as a function of physical distance. Provided the individual is drawn from a random-mating population, the proposed statistic, the correlation of zygosity, is closely related to the conventional population-level measure of LD, D2, but is agnostic with respect to allele frequencies. Application of the method to sixteen vertebrate species leads to the conclusion that >90% of recombination events are typically gene-conversion-like processes unaccompanied by crossovers, with the average lengths of conversion tracts being on the order of one to two kb in length. Contrary to common assumptions, the recombination rate between sites does not scale linearly with distance, even up to distances of 100 kb. In addition, the amount of LD between sites separated by <0.2 kb is typically much greater than can be explained by the conventional neutral model, possibly because of the nonindependent origin of mutations within this spatial scale. It will also be shown how single-individual LD profiles can be used to infer historical changes in effective population sizes. The formulations used in all of these inferences are closely related to expressions first derived by Tomoko Ohta and Motoo Kimura.

References:
Intimate symbiotic associations between bacteria and animals have evolved often, driven by effects of symbionts on hosts, including nutritional provisioning, protection from enemies, and manipulation of host reproduction. Although obligate symbionts typically cannot be cultured, genomic studies have illuminated their functions and have given insight into general processes of genome evolution. Many symbionts are strictly maternally transmitted and clonal, a system that minimizes evolutionary conflicts of interest between symbionts and hosts (at least female hosts).

While strict maternal transmission and clonality lowers the impact of cheaters and pathogenic tendencies in symbionts, it maximizes the opportunity for the accumulation of deleterious mutations. Obligate symbionts experience effective population sizes dependent on host population sizes and much smaller than those of environmental bacteria. Under strict maternal transmission, symbionts are clonal, unlike their sexual hosts, making symbionts especially prone to fixation of deleterious mutations and genomic erosion. One result is the accumulation of mutations that lower functionality, as reflected in protein stability, and mutations that inactivate and eliminate nonessential genes. Small genome symbionts show little or no homologous recombination or uptake of foreign genes, and thus cannot restore lost genes or acquire new capabilities. Furthermore, selection on hosts acts to compensate for inadequacies of mutualistic symbionts, enabling ever more extreme degeneration of symbiont genomes. As a result, symbiotic bacteria in insects possess the smallest known cellular genomes, and have undergone other extreme changes including extremely rapid protein evolution and codon reassignments associated with loss of translational release factors.

In most systems, symbionts retain genes underlying central elements of replication, transcription and translation but lose genes generally considered to be essential for production of cell envelope components. These losses appear to depend on compensatory coadaptation by hosts, sometimes involving host acquisition of bacterial genes, with most acquired genes originating from past symbionts. Another outcome in cases of insects dependent on symbionts with highly eroded genomes is the acquisition of a novel symbiont that has more robust gene sets and less degenerate proteins. This symbiont may replace an original symbiont with an eroded genome, or both may persist due to complementarity in their contributions to hosts.

In other bacterial-animal symbioses, symbionts are horizontally exchanged and recombine, preventing or slowing genomic erosion but increasing the opportunity for "selfish" tendencies. These cases include intracellular symbionts that are usually maternally but occasionally horizontally or paternally transmitted as well as gut symbionts that show more frequent horizontal transmission.
Prevalence of weak selection of protein evolution has now gained much support. The evolution of gene regulation seems to be more complicated and problematic. Recent progress on molecular mechanisms of gene regulation and epigenetics is reviewed, that suggests importance of weak selection. In connecting genotypes with phenotypes, numerous interaction systems are involved. Effects of transcription factor binding at DNA regulatory regions range from very small to quite large. Many factors like RNAs and protein kinases participate. Another important topic, function of chromatin structure in relation to epigenetics can not be slighted, i.e., DNA methylation and histone modifications have significant effects on gene expression in higher organisms. Gene families encoding such chromatin modifying factors seem to be under weak selection with pleiotropic effects. Chromatin component protein families also show characteristic patterns in evolution that suggest prevalence of weak selection. Some of the examples like H1 histone and high mobility group protein are presented. All systems involving such dynamically evolving protein families need to be incorporated for our understanding of evolution of gene regulation. Weak selection is thought to be prevalent at various levels of the regulation systems.
## List of Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution and Address</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babarinde Isaac Adeyemi</td>
<td>Division of Population Genetics, NIG, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan</td>
<td><a href="mailto:ibabarin@nig.ac.jp">ibabarin@nig.ac.jp</a></td>
</tr>
<tr>
<td>Hiroshi Akashi</td>
<td>Division of Evolutionary Genetics, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan</td>
<td><a href="mailto:hiakashi@nig.ac.jp">hiakashi@nig.ac.jp</a></td>
</tr>
<tr>
<td>Deepa Agashe</td>
<td>National Centre for Biological Sciences (NCBS), GKVK, Bellary Road, Bangalore, India 560065</td>
<td><a href="mailto:dagashe@ncbs.res.in">dagashe@ncbs.res.in</a></td>
</tr>
<tr>
<td>Doris Baachtrog</td>
<td>University of California, Berkeley, Department of Integrative Biology, Center for Theoretical Evolutionary Genomics, 3060 VLSB, Berkeley, CA 94720-3140</td>
<td><a href="mailto:dbaachtrog@berkeley.edu">dbaachtrog@berkeley.edu</a></td>
</tr>
<tr>
<td>Emily Brown</td>
<td>Department of Integrative Biology, University of California, 3040 Valley Life Sciences Building, University of California, Berkeley, Berkeley, CA, 94720, USA</td>
<td><a href="mailto:emilybrown@berkeley.edu">emilybrown@berkeley.edu</a></td>
</tr>
<tr>
<td>Andrew G. Clark</td>
<td>Department of Molecular Biology and Genetics, Cornell University, Campus Road, 107 Biotechnology Building, Cornell University, Ithaca, NY 14853-2703 USA</td>
<td><a href="mailto:ac347@cornell.edu">ac347@cornell.edu</a></td>
</tr>
<tr>
<td>Laurent Duret</td>
<td>Laboratoire Biométrie et Biologie Evolutive, UMR CNRS 5558, Université Lyon 1, 43 Bld du 11 Novembre 1918, 69622 Villeurbanne cedex, France</td>
<td><a href="mailto:Laurent.Duret@univ-lyon1.fr">Laurent.Duret@univ-lyon1.fr</a></td>
</tr>
<tr>
<td>D. Allan Drummond</td>
<td>Department of Biochemistry and Molecular Biology, University of Chicago, GCIS W234, 929 E 57th St, Chicago, IL 60637 USA</td>
<td><a href="mailto:dadrummond@uchicago.edu">dadrummond@uchicago.edu</a></td>
</tr>
<tr>
<td>Angela Early</td>
<td>Department of Molecular Biology and Genetics, Cornell University, 227 Biotechnology Building, Cornell University, Ithaca, NY 14853</td>
<td><a href="mailto:ame54@cornell.edu">ame54@cornell.edu</a></td>
</tr>
<tr>
<td>Toshimichi Ikemura</td>
<td>Nagahama Institute of Bio-Science and Technology, 1266 Tamura-Chō, NAGAHAMA, Shiga-pref, 526-0829 JAPAN</td>
<td><a href="mailto:t_ikemura@nagahama-i-bio.ac.jp">t_ikemura@nagahama-i-bio.ac.jp</a></td>
</tr>
<tr>
<td>Jun Gojobori</td>
<td>Graduate University for Advanced Studies (SOKENDAI), Kaniyamaguchi 1560-35, Hayama, Kanagawa 240-0193, Japan</td>
<td><a href="mailto:gojobori_jun@soken.ac.jp">gojobori_jun@soken.ac.jp</a></td>
</tr>
<tr>
<td>Kazuho Ikeo</td>
<td>Laboratory for DNA data analysis, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan</td>
<td><a href="mailto:kikeo@nig.ac.jp">kikeo@nig.ac.jp</a></td>
</tr>
<tr>
<td>Yuka Ikekaki</td>
<td>Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, 812-8581, Japan</td>
<td><a href="mailto:matrosika@gmail.com">matrosika@gmail.com</a></td>
</tr>
<tr>
<td>Hideki Innan</td>
<td>Graduate University for Advanced Studies (SOKENDAI), Kaniyamaguchi 1560-35, Hayama, Kanagawa 240-0193, Japan</td>
<td><a href="mailto:innan_hideki@soken.ac.jp">innan_hideki@soken.ac.jp</a></td>
</tr>
<tr>
<td>Jeff Fawcet</td>
<td>Graduate University for Advanced Studies (SOKENDAI), Kaniyamaguchi 1560-35, Hayama, Kanagawa 240-0193, Japan</td>
<td><a href="mailto:fawcett@soken.ac.jp">fawcett@soken.ac.jp</a></td>
</tr>
<tr>
<td>Akira Kawabe</td>
<td>Department of Bioresource and Environmental Sciences, Kyoto Sangyo University, Kamigamo Motoyama Kita-ku, Kyoto 603-8555, Japan</td>
<td><a href="mailto:akiraka@cc.kyoto-su.ac.jp">akiraka@cc.kyoto-su.ac.jp</a></td>
</tr>
<tr>
<td>Kent D. Kawashima</td>
<td>National Institute of Genetics, 1111 Yata, Mishima, Shizuoka, 411-8540 Japan</td>
<td><a href="mailto:kentkawashima@nig.ac.jp">kentkawashima@nig.ac.jp</a></td>
</tr>
<tr>
<td>Sonoko Kinjo</td>
<td>Laboratory for DNA data analysis, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan</td>
<td><a href="mailto:skinjo@nig.ac.jp">skinjo@nig.ac.jp</a></td>
</tr>
<tr>
<td>Jun Kitano</td>
<td>Ecological Genetics Laboratory, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan</td>
<td><a href="mailto:jkitano@nig.ac.jp">jkitano@nig.ac.jp</a></td>
</tr>
<tr>
<td>Wen-Ya Ko</td>
<td>Center for Genomic Medicine, Graduate school of Medicine, Kyoto University, 606-8507, Japan</td>
<td><a href="mailto:wenyako@gmail.com">wenyako@gmail.com</a></td>
</tr>
<tr>
<td>Junko Kusumi</td>
<td>Department of Environmental Changes Faculty of Social and Cultural Studies, Kyushu University, 744 Motooka Nishi-ku Fukuoka, 819-0395, Japan</td>
<td><a href="mailto:jkusumi@ses.kyushu.ac.jp">jkusumi@ses.kyushu.ac.jp</a></td>
</tr>
</tbody>
</table>
List of Participants

Charles H. Langley
Evolution and Ecology, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA
chlangley@ucdavis.edu

Manyuan Long
The University of Chicago, 1101 E 57th Street, Chicago IL 60637, USA
mlong@uchicago.edu

Michael Lynch
Department of Biology, Indiana University, Bloomington, Indiana 47405, USA
milynch@indiana.edu

Takashi Makino
Department of Environmental Life Sciences, Tohoku University, 6-3 Aoba, Aramaki, Aoba-ku, Sendai, 980-8578, Japan
tamakino@m.tohoku.ac.jp

Saurav Mallik
Department of Biophysics, Molecular Biology and Bioinformatics, University of Calcutta, 92, APC Road, Kolkata, India
saurav.bmb@gmail.com

Tomotaka Matsumoto
Division of Evolutionary Genetics, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
tmatsumoto@nig.ac.jp

Phillip W. Messer
Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA 94305 USA
messer@stanford.edu

Neha Mishra
Division of Evolutionary Genetics, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
nmishra@nig.ac.jp

Tomoko Ohta
Division of Evolutionary Genetics, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
tohta@nig.ac.jp

Andrew Moeller
Department of Ecology and Evolutionary Biology, Yale University, 4305 Duval Street, Apt. 120, Austin, TX 78751, USA
andrew.moeller@yale.edu

Naoki Osada
Division of Evolutionary Genetics, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
nosada@nig.ac.jp

Nancy A. Moran
Department of Integrative Biology, University of Texas, NMS Building A5000, 2506 Speedway, Austin TX 78712, USA
nancy.moran@austin.utexas.edu

Dmitri Petrov
Department of Biology, Stanford University, 371 Serra St, Stanford, CA 94305, USA
dpetrov@stanford.edu

Yoshihito Niimura
Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan
aniimura@mail.ecc.u-tokyo.ac.jp

Guy Sella
Department of Biological Sciences, Columbia University, 740A Mudd M.C. 2435, New York, NY 10027
gs2747@columbia.edu

Masafumi Nozawa
Laboratory for DNA data analysis, 1111 Yata, Mishima, Shizuoka 411-8540, Japan
mnozawa@nig.ac.jp

Eriko Sasaki
Gregor Mendel Institute of Molecular Plant Biology, Dr. Bohr-Gasse 3, 1030 Vienna, Austria
eriko.sasaki@gmi.oeaw.ac.at

Howard Ochman
Department of Integrative Biology, 2506 Speedway A5000, NMS 4.110, University of Texas, Austin TX 78712, USA
howard.ochman@austin.utexas.edu

Naruya Saitou
Division of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizuoku 411-8540, Japan
saitounr@nig.ac.jp

Miuzuki Ohno
Department of Medical Biophysics and Radiation Biology, Faculty of Medical Sciences, Kyushu University
mohno@biophys.med.kyushu-u.ac.jp

Kazumasa Shirai
Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan
blackparade31109@yahoo.co.jp
LIST OF PARTICIPANTS

Wolfgang Stephan
Fakultät für Biologie, Ludwig-Maximilians-Universität München, LMU München, Biozentrum Martinsried, Großhaderner Straße 2, D-82152 Martinsried
stephan@bio.lmu.de

Ryuichi Sugino
Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, 812-8581, Japan
rpsugino@gmail.com

Hidenori Tachida
Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, 812-8581, Japan
htachida@kyudai.jp

Toshiyuki Takano-Shimizu
Drosophila Genetic Resource Center, Kyoto Institute of Technology, Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
fruitfly@kit.ac.jp

Aya Takahashi
Department of Biological Sciences, Tokyo Metropolitan University, Tokyo, Japan
ayat@tmu.ac.jp

Shohei Takuno
Graduate University for Advanced Studies (SOKENDAI), Kamiyamaguchi 1560-35, Hayama, Kanagawa 240-0193, Japan
takuno_shohei@soken.ac.jp

Miho Tamura
Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, 812-8581, Japan
mtamura@kyudai.jp

Kosuke M. Teshima
Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, 812-8581, Japan
kteshscb@kyushu-u.org

Chau-Ti Ting
Department of Life Science, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, Taiwan 10617, ROC
ctting@ntu.edu.tw

Nobuhiko Tokuriki
Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC Canada
V6T 1Z4, Vancouver, BC Canada
tokuriki@msl.ubc.ca

Sen Xu
Department of Biology, Indiana University, 1001 E Third St, Bloomington, IN, 47405, USA
senxu@indiana.edu

Soojin Yi
School of Biology, Georgia Institute of Technology, 310 Ferst Drue, Atlanta, GA 30332 USA
soojinyi@gatech.edu

Kohta Yoshida
Ecological Genetics Laboratory, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
koyoshid@nig.ac.jp

Masa-aki Yoshida
Laboratory for DNA data analysis, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken, 411-8540, Japan
myoshida@nig.ac.jp

Takanori Yoshida
Department of Bioresource and Environmental Sciences, Kyoto Sangyo University, Kamigamo Motoyama Kita-ku, Kyoto 603-8555, Japan
yoshida.takanori@cc.kyoto-su.ac.jp