Evolution of cancer genomes

Cancer cells evolve from normal cells by somatic mutations and natural selection. Comparing the evolution of cancer cells and that of organisms can elucidate the genetic basis of cancer. Analyzing somatic mutations in >400 cancer genomes, we found that the frequency of somatic single-nucleotide variations (SNV) increases with replication time during the S phase much more rapidly than germ-line SNV. The ratio of nonsynonymous to synonymous SNV is higher for cancer cells than for germ-line cells, suggesting weaker purifying selection against somatic mutations. Among genes with recurrent mutations only cancer driver genes show evidence of strong positive selection, and late-replicating regions are depleted of cancer driver genes. Thus, replication timing plays a prominent role in shaping the SNV landscape of cancer cells. Moreover, we used large-scale human protein interaction network (PIN) data to explore the relationship among network topology, somatic mutation, and evolutionary age of cancer proteins. Cancer genes tend to be subjected to stronger purifying selection than non-cancer genes and to be ancient, likely originated in early metazoan, although they are younger than Mendelian disease genes and essential genes. Finally, protein age positively correlates with protein connectivity in the human PIN.

Symposium (S-1 – S-5)


Revealing the time-dependent evolution of the hepatitis B virus from a chain of sequentially infected chronic carriers

The evolution rates of the hepatitis B virus (HBV) estimated using contemporary sequences are $10^9$-$10^4$ times faster than those derived from archaeological and genetic evidence. To evaluate whether the dual demands (i.e., adaptation within hosts and colonization between hosts) of the viral life cycle affect this conundrum, the HBV quasispecies dynamics within and among hosts from a chain of sequentially infected transmission was examined. We observed that viruses colonized to the subsequent hosts did not directly originate from those adapted to previous hosts, suggesting that various strains excel at distinct infection stages. Continual switching between colonization and adaptation resulted in the rapid accumulation of mutations at a limited number of positions, which quickly became saturated, whereas substitutions at the remaining regions occurred at a much slower rate. Thus, the HBV substitution rate was not constant, but decreased as the divergence time increased. Our study explains the >1000-fold difference in HBV substitution rates reported in the literature and provides new insight into the origin of the virus.

$ S $ SHI, Suhua$^1$ ($^1$State Key Lab. Biocontrol, Grad. Sch. Life Sci.s, Sun Yat-Sen Univ.)

Living on the edge--Mangrove genomes reveal severe adaptative strategies and hint at an uncertain future

Fewer than 100 species of woody plants, collectively known as mangroves, inhabit the interface between terrestrial and marine environments. Understanding how mangroves adapted to the habitats, which have experienced frequent sea level changes, may inform about their future. We report the sequencing of the whole genomes of species in three mangrove genera (Avicennia, Rhizophora and Sonneratia) and the transcriptomes of 14 others. The data show that the three species invaded new habitats independently within a relatively small window of time (43-58 million years ago), suggesting that climates favorable for invasion have occurred infrequently. In adapting to new habitats, mangroves, instead of expanding their genomes as might have been expected, have undergone extensive convergent reductions in both repetitive sequences and gene families. In the process of adaptation, ~300,000 amino acids have been substituted in each genome, close to the limit imposed by the "cost of selection". All species of mangroves show very low heterozygosity as a consequence of small populations occupying shifting habitats under continual sea level changes. Mangrove genomes suggest a precarious existence throughout the history. Hence, their ability to withstand the double assaults of rising sea levels and increasing human perturbations is hardly assured.

Population history before and after the transition from hunter-gatherer to agriculturist in Africa and Southeast Asia

Since the rise of agriculture in sub-Saharan African ~4,000 years ago, Bantu agriculturists spread to the African continent and displaced most of hunter-gatherer groups. The Khoisan people, living in southern Africa, have maintained ancient lifestyles as hunter-gatherers up to modern times though little else is known about their early history. We sequenced the complete genomes of five Khoisan individuals and one Bantu-speaker from southern Africa. Population genetic analyses using SNP datasets from worldwide individuals showed that two genomes from the Ju/'hoansi (Northern Khoisan) population contain exclusive Khoisan ancestry. A coalescent analysis was applied to the two genomes and revealed that the Khoisan and their ancestors have maintained their large effective population size since their split from non-Khoisans ~100-150 kya. In contrast, the ancestor populations of non-Khoisans dramatically declined after the split and lost more than half of their genetic diversity. This is in stark contrast to the current census size of the Khoisans, which is drastically smaller than that of the Bantu-speakers.

In Asia, the Austronesian agriculturists dispersed from East Asia to Southeast Asia ~5,000 years ago. They are currently major populations in Indonesia. Our study, using a 500K SNP dataset of 508 individuals from Indonesia, revealed the history of migration and admixture of two distinct populations in Indonesia.
Big-Data analysis used NGS and study of evolution

Next Generation Sequencing (NGS) technology makes a big step of the life science study in the many varieties of the field including environment, populations, and comparative genomics. Especially metagenomics study gives important and useful information on the interaction between species or among species. The approaches of the combination of metagenomics data and any additional environmental information by using multiple analytical approaches are expected to give useful and new insight of the dynamic biological systems and environments. Such a information will also be interested to understand evolution. For the purpose, we need to use large sequence data from NGS and additional information for annotate. We are now developing the data analysis system for the NGS data with simple interface for data analysis. Here, I will introduce the recent trial by using meta genomic approaches to understand the dynamics of land and marine biological systems and also the usage of our NGS data analysis system named Maser.

Ser7 phosphorylation of RNAPII-CTD ensures on-chromatin retention of nascent ncRNAs, facilitating establishment and epigenetic inheritance of RNAi-dependent heterochromatin formation

Multiple phosphorylation of C-terminal domain (CTD) of the largest subunit of RNA polymerase II (RNAPII) links transcription and co-transcriptional events such as RNA processing and active histone modifications. Phosphorylation at Ser2 (Ser2P) and Ser5 (Ser5P) are well studied, but conserved function of phosphorylation at Ser7 (Ser7P) is poorly understood. Here we show that in fission yeast, Ser2P and Ser7P regulate silencing of centromeric ncRNAs. Ser2P does not affect heterochromatin structure, whereas Ser7P is involved in RNAi-dependent heterochromatin formation. Lack of Ser7P causes failure of siRNA generation and deposition of H3K9 methylation at pericentromeric heterochromatin. Ser7P and RNA binding activity of chromodomain of Chp1 (Chp1-CD) stimulate on-chromatin retention of heterochromatic ncRNA. Furthermore, Ser7P is important for physical interaction between RNAPII and Chp1. Finally, we found Ser7P is required for establishment and epigenetic inheritance of heterochromatin, cooperatively with Chp1-CD. Our findings reveal that a CTD code of Ser7P regulates a silent histone code of H3K9 methylation by linking ncRNA dynamics to recruitment of an epigenetic modifier.

The H3K27 demethylase, Jmd3, is essential for activation of pax6 during eye development

H3K27 methylation is a gene-silencing epigenetic mark that is dynamically regulated during vertebrate development. In this study, we analyzed expression and roles of jmd3.L and jmd3.S, the two homeologous H3K27 demethylase genes of X. laevis. While expression of jmd3.S was obvious in the axial mesoderm, neural plate and surrounding preplacodal ectoderm of early neurula embryos, jmd3.L expression was evident only in the preplacodal ectoderm, suggesting ongoing sub-functionalization in this homeolog pair. Embryos injected with an antisense morpholino oligonucleotide for jmd3.S (jmd3.S-morphants) failed the gastrulation, whereas jmd3.L-morphants showed a microphthalmia phenotype. Gene expression analysis revealed that the jmd3.L-morphants showed normal expression of rx in the retina field and dlx5 in the preplacodal ectoderm, but failed to express pax6 and six3 in the lens field (part of the preplacodal ectoderm) at the early neurula stages. These analyses first revealed the essential role for jmd3 in eye development, by selectively inhibiting the sub-functionalized X. laevis homeologs.

Ser7 phosphorylation of RNAPII-CTD ensures on-chromatin retention of nascent ncRNAs, facilitating establishment and epigenetic inheritance of RNAi-dependent heterochromatin formation

Multiple phosphorylation of C-terminal domain (CTD) of the largest subunit of RNA polymerase II (RNAPII) links transcription and co-transcriptional events such as RNA processing and active histone modifications. Phosphorylation at Ser2 (Ser2P) and Ser5 (Ser5P) are well studied, but conserved function of phosphorylation at Ser7 (Ser7P) is poorly understood. Here we show that in fission yeast, Ser2P and Ser7P regulate silencing of centromeric ncRNAs. Ser2P does not affect heterochromatin structure, whereas Ser7P is involved in RNAi-dependent heterochromatin formation. Lack of Ser7P causes failure of siRNA generation and deposition of H3K9 methylation at pericentromeric heterochromatin. Ser7P and RNA binding activity of chromodomain of Chp1 (Chp1-CD) stimulate on-chromatin retention of heterochromatic ncRNA. Furthermore, Ser7P is important for physical interaction between RNAPII and Chp1. Finally, we found Ser7P is required for establishment and epigenetic inheritance of heterochromatin, cooperatively with Chp1-CD. Our findings reveal that a CTD code of Ser7P regulates a silent histone code of H3K9 methylation by linking ncRNA dynamics to recruitment of an epigenetic modifier.