Science Summit 2023 September 20

Genomics of Biodiversity and Evolution
Aula Magna, Stockholm

Speaker abstracts

Foto: Love Dalén

SciLifeLab

09:45-10:15
Elinor Karlsson
UMass Chan Medical School and Broad Institute, USA

Investigating mammalian evolution and human disease through comparative genomics in hundreds of species

A major challenge in genomics is discerning which bases among billions alter organismal phenotypes and affect health and disease risk. The Zoonomia project compared 240 different placental mammal species to detect which individual bases in the genome are exceptionally conserved (constrained) and likely to be functionally important. Eighty percent of the most constrained bases are outside protein-coding exons, and half have no functional annotations in the ENCODE project. By pairing Zoonomia's genomic resources with phenotype annotations, we find genomic elements associated with phenotypes that differ between species, including olfaction, hibernation, brain size, and vocal learning. Comparative genomics is advancing human health today by identifying functionally important in both coding and noncoding regions. Exploring the genomic basis of phenotype diversity that has emerged across 100 millions of years of placental mammal evolution is a powerful tool for discovering the next generation of biotechnological advances.

10:15-10:40
Tom van der Valk
Swedish Museum of Natural History

A Million-Year-Old Journey: Exploring Mammoth Speciation and Adaptive Evolution through time

In this presentation, I will share our efforts over the past years in recovering genome-wide data from woolly mammoths and discuss how the genomic insights have advanced our understanding of evolutionary processes such as speciation and long-term adaptive evolution. Our research involved sequencing specimens from the Early and Middle Pleistocene subepochs, including some of the last surviving woolly mammoths. This work led to the identification of a previously unknown mammoth lineage, evidence of hybridization between different mammoth species, and the finding that most protein-coding changes linked to cold adaptation in woolly mammoths were already present one million years ago. I will show that the woolly mammoths had acquired a diverse array of positively selected genes associated with among others hair and skin development and fat storage at the time of its origin. Our research also identified genes that underwent recent positive selection, including those related to skeletal morphology, body size, and a gene potentially responsible for the small ear size observed in Late Quaternary woolly mammoths. Overall, our findings highlight the potential of palaeogenomics in enriching our understanding of speciation and long-term adaptive evolution.

11:20-11:55
Richard Durbin
University of Cambridge, UK

Insights from high quality genome sequencing across the tree of life

I will discuss progress in scaling up the generation of high quality reference genome sequences, and some of the insights that we have obtained from them. Using high accuracy PacBio CCS reads and paired end Illumina Hi-C data we are now obtaining essentially complete, high contiguity chromosomal assemblies at an increasing rate and decreasing cost, with more than a thousand assemblies completed within the Tree of Life programme at the Wellcome Sanger Institute. Because in most cases even heterochromatic repeat DNA is well assembled, we have been able to characterise turnover of centromere-associated repeates which are some of the most rapidly evolving seqence in the genome, observing repeated transitions in plants between satellite tandem repeats and transposon cluster patterns. Furthermore, by selection of very high confidence bases from the individual CCS sequencing reads, we are able to identify signatures of somatic mutations across a very wide range of species, suggesting previously unobserved mutational processes.

11:55-12:20
Karin Norén
Stockholm University

Conservation genomics of the Scandinavian Arctic fox

Over the past decade, whole genome sequencing has generated important insights about key processes in conservation genetics, e.g. inbreeding depression, genomic erosion and genetic rescue in small and threatened populations. The Scandinavian Arctic fox (Vulpes lagopus) was on the verge of extinction in the late 1990s. In response to efficient conservation actions, the population has increased, but studies have documented that the bottleneck, geographic fragmentation and long-term low population size resulted inbreeding depression and loss of genetic variation. To explore the dynamics of inbreeding, accumulation of deleterious genetic variation and genetic rescue, we assembled a draft reference genome and resequencing data of >80 complete Arctic fox genomes. We found alternating levels of genomic inbreeding and a short-term genetic rescue effect, operating over different time scales. An immigration event resulted in an increased mutational load where immigrant offspring displayed higher proportion of loss of function mutations compared to native individuals. Further, we established a link between deleterious genetic variation and individual fitness. The results from these studies play a fundamental role for making informed decisions in conservation of this particular population, but also make important contributions for conservation of other small and threatened populations and species.

13:40-14:15
Tom Gilbert
University of Copenhagen, DK

Domestication hologenomics – what are we missing without taking this approach?

Analyses that compare the genomes of contemporary domestic animals and plants with those of their wild relatives have provided a wealth of insights into not only when and where our ancestors started the process, but also what specific genetic variants are key to modern phenotypes. Furthermore, once coupled to palaeogenomic data, such datasets can also even reveal the order in which such variants arose, shedding further insights into the process itself. However while there is no doubt that we have learnt much about domestication in general, and indeed for most domestic species we can clearly document the genetic basis of why the end product differs from the start, I argue that there may be certain processes that were involved that have been largely overlooked, in particular related to the so-called hologenome.

14:15-14:40
Tanja Slotte
Stockholm University

Sequencing the supergene that governs Darwin's different forms of flowers

Supergenes are genomic regions containing sets of tightly linked genes that control multi-trait phenotypic polymorphisms. Although supergenes are responsible for a wide variety of balanced polymorphisms in nature, our understanding of the origins and evolution of supergenes remains incomplete. We aim to fully characterize and study evolutionary processes at one of the first described supergenes, the S-locus that governs a floral polymorphism called distyly. We are doing so in Linum, wild flaxseed species, a system where Darwin himself described this floral polymorphism, but where its genetic basis remained unknown. To generate a genomic framework for the study of distyly, we assembled high-quality genomes of a diverse set of Linum species. We then identified the distyly supergene and showed that it harbors indel variation and not inversions, which is typical for many other supergenes. Our results have important implications for the evolution and breakdown of distyly supergenes, and shed light on the genetic architecture and evolution of the classic supergene that governs Darwin's "different forms of flowers".

15:45-16:20
Leif Andersson
Uppsala University

The remarkable population structure of Atlantic and Baltic herring - selection, selection, selection

Initial genetic studies with a handful of neutral markers revealed no genetic differentiation among populations of Atlantic herring, not even between Atlantic and Baltic herring classified as distinct subspecies by Linnaeus. Whole genome sequencing has totally changed the picture. Atlantic herring can now be divided into many subpopulations (ecotypes). It is an adaptive radiation with incomplete reproductive isolation between ecotypes. We find strong genetic differentiation at loci under selection but minute genetic differentiation at neutral loci. The explanation for this is the huge population sizes minimizing drift, high fecundity allowing effective selection, a homing behavior but with gene flow. Herring is a broadcast spawner and it is probably no strong selection for prezygotic reproductive isolation. Ecological adaptation in Atlantic herring is related to a diversity of environmental conditions including salinity, temperature and light conditions as well as behavioral traits such as timing of reproduction and migration. Recently we used SNP-chip analysis to explore the population structure of Baltic herring in the Bothnian Sea. This analysis revealed two major subgroups: spring- and autumn-spawning Baltic herring. However, a third genetically distinct population was present among the spring-spawners. This ecotype had a three-fold larger body size than other populations spawning in the same area at the same time implying a marked difference in feeding behavior. This illustrates the adaptive differentiation occurring in the herring. The results have important implications for fishery management of the herring in the Baltic Sea and elsewhere.

16:20-16:45
Karin Rengefors
Lund University

Why cyanobacterial blooms produce toxins – unravelling the underlying genetic diversity in the microcystin gene cluster

Cyanobacterial blooms are a global threat to freshwater ecosystems since they produce cyanotoxins that are poisonous to humans, wildlife, and livestock. Microcystin is the most common and toxic among the cyanotoxins and induces liver failure and tumor promotion in mammals. However, its function in cyanobacteria is still contested. In Microcystis spp., the most common microcystin-producing genus, microcystin-producing and non microcystin-producing strains co-exist within populations. Moreover the proportions of the strain types vary in time and space. Microcystin is biosynthesized by the constitutively expressed mcy gene cluster consisting of ten modular genes (mcyA-J). Previous studies have suggested that non-producers lack the entire cluster. In the field, toxigenic strains are quantified using quantitative PCR targeting one of the genes in the microcystin gene cluster (usually mcyB or E). However, the number of gene copies are not always correlated with microcystin in the water. To better understand the underlying genotypes of microcystin producers and non-producers we sequenced the genomes of strains of Microcystis isolated from a single bloom. At the same time, the strains were phenotyped to determine microcystin variants. Unexpectedly, non microcystin-producing strains displayed a range of genotypes yet with a common pattern of mostly lacking mcyF, G, and J. Other genes in the cluster were either present or had partial hits against our

custom-made microcystin gene database. We suggest that non microcystin-producing Microcystis are genotypically and phenotypically diverse and that genotype composition varies among populations. I will also discuss potential causes of mcy gene-loss and the way forward to unravel the function of microcystin.