JAMES RENNIE BEQUEST

REPORT ON CONFERENCE

Conference Title: Genetics – Understanding Living Systems; XX International Congress of Genetics

Travel Dates: 12th July 2008 – 18th July 2008

Location: International Congress Centre, Berlin, Germany

Group Member(s): Simon Aeschbacher

Aims: a) To present the poster "Contrasting Observed and Simulated Genetic Structure of Bottlenecked Alpine Ibex Populations Reveals Evidence of Gene Flow"
b) To learn about new results in the field of population genetics, connect with people, and profit for my own work
c) To get an overview on the recent development in genetics research

---

OUTCOME:

Showing the four letters of the genetic code, a photograph of two chromosomes and the picture of a flying bird, the cover of the program booklet picked up the central theme of the 20th International Congress of Genetics (ICG): Understanding living systems. While most questions in biology arise from what we observe on the level of organisms and their interactions, we must dig more deeply down to the cellular and molecular levels to find answers. As we learn on this top-down journey, we become able to make predictions on how changes on the level of genes may affect organisms and to do experiments. This combination of top-down and bottom-up approaches leads to a full understanding of the mechanisms. Over the last decades, genetic research has profited from enormous technical and methodological improvements. The purpose of the congress was to address the latest development in theory and application in this field.

Held in Berlin, Germany, from 12th to 17th July 2008, the ICG was a big international conference with about 2000 participants, one Keynote Symposium, nine Plenary Lectures, 54 Concurrent Symposia, two Poster Sessions and 280 of the most prominent genetics researches speaking. The International Congress Centre (ICC) in Berlin provided a location with the necessary dimensions, but was not a very inviting place apart from that.

The Keynote Symposium launched the congress on Friday afternoon. The following five congress days all had the same structure, starting with a Plenary Lecture in the morning, followed by Concurrent Symposia. The two Poster Sessions were during lunchtime and there was a second window with Concurrent Symposia in the afternoon. The day was closed with another Plenary Lecture.

I presented my poster on population differentiation and gene flow in Alpine Ibex on Sunday and Monday. The poster walls were placed such that they created small niches and people needed to get into these niches to read the poster. The niches became crowded quickly. I found this arrangement not very helpful for presentation and discussion. The interest in my poster seemed to be limited. I had a couple of discussions, mainly arising from questions about the context of the project or the understanding of the methods I used. People

interested in the application of methods said they found my poster too theoretical for them to understand. These were good opportunities to explain what I had done. I would have appreciated more feedback from theoreticians, but would most likely have had to make a bigger effort to ask them specifically.
The authors were supposed to be present close to their posters, and so there was little chance to get around and visit other posters during the same session. Nevertheless, three posters stood out for me: Danielle Jones, a student of John Wakeley, suggested a method to distinguish between recombination and gene conversion in a model with three linked neutral loci. Investigating the proportional contribution of the different absorbing states to the overall probability of identity-by-descent, she found that one absorbing state might be informative to distinguish the two processes. By designating one of the loci as a beneficial mutation, she then showed that the pattern of linkage disequilibrium depends on the ordering of the beneficial mutation among the three loci and the amount of recombination. Ülo Väli, in collaboration with Hans Ellegren, addressed the question if there is a concordance in estimates of genetic diversity based on single nucleotide polymorphisms (SNPs) and estimates based on microsatellites. Comparing SNP diversity and microsatellite variability in populations of four European carnivore species he found significant correlations between the diversity estimates among populations, but no correlations on the level of individuals within populations. The range of putatively neutral SNP diversity was much larger then that of microsatellite heterozygosity. Väli argued that using SNPs has the advantage of avoiding the ascertainment bias associated with the choice of polymorphic microsatellites. Claudia Junge, PhD student of Craig Primmer and Asbjörn Vøllestad, investigated population genetic structure and gene flow in recently established populations of European grayling in a lake in southern Norway. The lake and its tributaries had been colonised in the 1880s and more than 20 demes have since established in the tributaries. Genetic differentiation increased despite the fact that all demes are potentially connected via the lake and gene flow may mainly have happened along the tributaries, but not among them. This drives the attention to fundamental questions of adaptive radiation and sympatric differentiation. Junge points out that this river system could be used to experimentally test for divergent natural selection.

In the second session on Wednesday and Thursday, Vitor Sousa, a PhD student supervised by Lounes Chikhi, had a poster on Approximate Bayesian Computation (ABC) without summary statistics. He analysed a simple admixture model in which two parental populations mixed at a certain time in the past to create a hybrid population. The three populations then evolved under pure genetic drift. The parameters of interest here are the proportions to which the two parental populations contributed to the hybrid population. The simplicity of this model and further assumptions on the number of alleles per locus and the initial distribution of allele frequencies allowed him to calculate results with a full-likelihood method. Sousa then compared the performance of two types of ABC procedures to the likelihood estimates. In the first type, he based the rejection step on the full allelic distribution (thus 'without' summary statistics), in the second he used the summary statistics expected heterozygosity and Fst. Both types of ABC procedures provided good approximations to the full-likelihood method, but the method without summary statistics was noisier. Although this study showed that it is possible to use ABC without summary statistics for inference, a thorough comparison for more complex models and a better quantification of the relative performance of the methods is needed.
In her poster on population connectivity in patchy and temporally unstable deep-sea habitats, Asta Audzijonyte and collaborators addressed an interesting question: Do existing data from allozyme and mitochondrial DNA markers allow us to infer population connectivity? Using data from 20 invertebrate species and simulations under different demographic scenarios, the authors evaluated the power of available data sets to distinguish among different demographic scenarios. The results show that in the context of a metapopulation model, it takes many generations to reach drift-mutation equilibrium. Further, the dispersal characteristics of the species influence population structure and diversity. Audzijonyte concluded that in realistic populations not having reached drift-mutation equilibrium, current genetic data can detect genetic structure only under very low levels of migration. By arbitrarily removing populations from her sample, she also showed that this increased the likeliness of finding population structure. This illustrates the bias created by missing samples from single populations. For me, a very stimulating poster was the one called Approximate Genealogical Inference by Chris Hallsworth, PhD student of Gil McVean. Genetic variation data only provide limited information on the genealogical history of a sample and recombination between different loci makes it even more difficult. One way of working around that problem is to treat the genealogy as missing data and use sampling methods to integrate over it. This method is currently only applicable to smaller data sets. As an alternative, Hallsworth presented a novel approach in which he approximates the genealogical distribution and showed an algorithm to draw from this distribution. The algorithm produces trees with properties consistent with known results such as the distribution of the time to the most recent common ancestor.

Given the large number of talks I can only pick out a few of them. In her opening keynote lecture, Christiane Nüsslein-Volhard, winner of the Nobel Prize in Medicine in 1995, gave an chronologic overview on the development of different methods of mutagenesis and the milestones in establishing the fruit fly, *C. elegans*, zebrafish and mouse as model organisms. Oliver Smithies then traced the progression from gels to genes, and from knocking out genes to altering their levels of expression in a very entertaining talk, using mainly excerpts from his lab book for illustration. In the third keynote lecture, Mario Capecchi focussed on the mouse as a model organism for human cancer. He pointed out how important it is not only to use a germ line model in mouse, but models for somatic processes. Both, Smithies and Capecchi, were awarded with the Nobel Prize in Medicine in 2007. In a very passionate plenary lecture on Monday afternoon, Eric Lander presented implications of the Human Genome Sequencing project for Biology and Medicine. He showed how genome wide association studies are being used to relate inherited genetic variation to common diseases, how investigation of somatic genetic variation and gene expression may reveal mechanisms of cancer or how genetic variation on the level of populations can be used to search for signs of adaptation in human evolution. The talk also revealed the enormous amount of money, data and resources needed in the field of human genomics.

In the Population Genetics session, Wolfgang Stephan's talk illustrated the shift from using a limited number of genetic markers towards approaches using data on polymorphism in the entire genome. This has been facilitated by recent genome sequencing projects in species like *Drosophila melanogaster* and, in parallel, by new statistical tests that infer the demographic history and localise potential targets of selection. Comparing X-chromosomal diversity of African and European populations of *D. melanogaster*, he illustrated the estimation of parameters in simple models of expansion and showed signs of adaptation to new habitat types in Europe. He also contrasted the variation on the X chromosome to autosomal data between African and European flies and argued that the ratio of female to male effective population size is much smaller in the derived European populations. The scale of data available has also exploded for human genetic variation in the last few years. In a review-like talk in the session on the Evolution of Humans, Gil McVean summarised past studies on natural selection and asked what can be learned from whole-genome scans. Different genes show different footprints of selective histories, such as high amounts of polymorphism for the MHC, or, in the other extreme, very low diversity in the surroundings of the lactase gene, characteristic for the sweep of a beneficial mutation (lactase persistence). Although we see signals of selection and adaptation, McVean said that in recent genome wide scans signs of strong selective sweeps and ancient balancing selection are rare. On the other hand, strong population differentiation seems to be common. A general statement on the rates of adaptation in humans cannot be made. McVean also pointed out that there is little evidence for a correlation between levels of polymorphism and recombination, leading to what he called the paradox of the missing sweeps: Although there is evidence for widespread local selection, there is little evidence for many strong selective sweeps in the human genome. Is this just a question of the power to detect them or due to a more fundamental reason such as ancient standing polymorphisms?

The large scale genome-wide association studies seem to have reduced the importance of pedigree information and linkage analysis. Nevertheless, in his talk in the session on Human Genetics, Augustine Kong demonstrated that the recent high density SNP data for large numbers of individuals can be use to exploit pedigree relationships in a new manner. Particularly, he showed that algorithms for the reconstruction of haplotypes can be much more efficient if they make use of known haplotypes of other individuals in the pedigree. These do not need to be the parents of the individual of interest, but can have much lower familial relationship. They only need to be genetically similar. Kong illustrated this concept of 'surrogate parenthood' on data from the human population of Iceland and proposed that large sequencing projects may be improved by such new phasing methods.

The scientific program closed with a Plenary Lecture by Antoine Kremer, who asked if forest trees are ill-equipped to face climatic changes. In a large-scale study on oak data from Europe, he investigated the dynamics of colonisation after the last ice-age, the amount of differentiation over space, and the rate of evolution. Much of the contemporary differentiation must have arisen after the last glacial period, suggesting that rates of dispersal and gene flow have been high. However, Kremer doubts if those rates will be high enough to face the rate of climatic change. With a clear undertone, he finished by saying that at any rate we would better ask the oaks to carry out a study on mankind’s potential to cope with climatic change instead of being concerned about the fate of the oaks.

Summarised, the ICG 2008 gave an impressive overview on the current state of genetics research, namely the deep-sequencing projects, the recovered potential of stem cells, the growing insight into the role of different types of RNA, increasing knowledge on functional mechanisms in neurogenetics and attempts to quantify rates of evolution in ecological genetics. It also showed that genetics brings along responsibility in ethics, society and regulations. Third, the field is getting more interdisciplinary, which is a challenge for education and calls for more engineering to bridge gaps between formerly isolated areas. Related to my work, it was good to meet Vitor Sousa and exchange ideas about Approximate Bayesian Computation. The talks and posters using coalescence theory were stimulating for my plans on future projects.