AWC Projects 2012

Theme: 1. Genomics & Biomathematics
Activity: A. Developing cutting-edge genomic, bioinformatic & biomathematical technologies for New Zealand

Investigator	Project	Stakeholders
Bryant	**SNP based phylogeographic and demographic inference** **Team members**: David Bryant, Steffen Klaere, Remco Bouckaert Single Nucleotide Polymorphisms (SNPs) are locations in a genome where individuals within a population (or closely related populations) differ by a single base-pair mutation. With SNP-chip and next generation sequencing it is now feasible to cheaply obtain tens of thousands of SNPs from hundreds of individuals, a development which is increasingly being applied to biogeography and ecological genetics. For this project we are developing novel and sophisticated mathematical techniques for inferring phylogenetic and demographic history from SNP data, with the aim of making previously intractable inference problems feasible	• Molecular ecologists using genetic data to infer population demographics • Agricultural and horticultural researchers
Matisoo-Smith	**Development and application of new bioinformatic tools for understanding genetic change in a population through time** **Team members**: Post Docs – Michael Knapp (AWC funded), and Ann Horsburgh; PhD students: Stefan Prost (AWC funded), Karen Greig We are working on several projects involving data from ancient and modern animal and human populations. In addition to applying standard Bayesian and population genetic modeling tools, we are working with colleagues in the US and Europe to develop new methods of analysis for population data through time.	• Archaeologists & other colleagues dealing with historical research • Pacific Island communities who are the descendants of the samples we are working with
Steel & Semple	**Development of network-based evolutionary analysis** **Team members***:* Charles Semple and Mike Steel Phylogenetic networks are an essential tool for representing evolutionary relationships when evolution involves reticulate (non-tree-like) processes, such as (i) the formation of hybrid species, (ii) lateral gene transfer within microorganisms, (iii) the ancestry of a diploid population subject to recombination. In research to date, we have provided the first mathematical treatment of a model which biologists have used to try to quantify the extent of lateral gene transfer. We have also explored different ways to define a 'phylogenetic tree' when reticulate processes are at play, and investigated fundamental algorithmic questions that arise in phylogenetic networks. In 2012, we will consider some mathematical and algorithmic approaches to pedigree graphs, which arise in considering population history (process (iii) above).	
Steel & Bryant	**Development of tools for genomic phylogenetics** **Team members:** Mike Steel and David Bryant One of the challenges arising from new sequencing technologies is how to use the large amount of patchy genomic data that is generated to learn more about evolutionary relationships. In our research to date, we have provided the first mathematical analysis of how much species coverage is required in order to have a high probability of being able to reconstruct a tree with confidence. Our results show that many large-scale data sets prove to be incapable of resolving certain evolutionary relationships. We have also described new approaches to estimating ancestral sequences in trees that provide sufficient taxon coverage, and developed new approaches for tree comparison. We plan to develop these and other related themes in 2012.	

Theme: 1. Genomics & Biomathematics
Activity: B. Using new genetic knowledge for molecular ecology & evolution

Investigator	Project	Stakeholders
Bryant	**Improved Network Based Inference Tools** **Team members:** David Bryant, Daniel Huson (International Advisory Panel) Phylogenetic trees are like family trees for different species. Split networks are like phylogenetic trees on steroids: they represent more complex relationships between species, and so can be used to make inferences when nature is too complicated for simple trees. AWC members past and present have made key contributions to the development of these methodologies, however there is much more to be done. In this project we are developing new network methods and software for both metagenomic and population level data, working with collaborators Daniel Huson and Naruya Saitou.	• Evolutionary biologists • Medical researchers
Buckley	**Transcriptome evolution in New Zealand invertebrates** **Team members**: Thomas Buckley, Alice Dennis, Luke Dunning, Victoria Twort, Richard Newcomb We are using transcriptome data from New Zealand stick insects and weta to identify candidate genes involved in cold tolerance and reproductive isolation. The study of cold tolerance is directed at stick species from a range of habitats from the alpine zone to relatively warm, northern forests. Genetic variation at these candidate loci is being screened across these populations and species. Neutral markers obtained are also being used to obtain a robust phylogeny for the New Zealand species and their overseas relatives. The transcriptome sequencing in weta is targeted at reproductive tissues, and focused on highly threatened species, especially giant weta. These reproductive genes will be used to examine gene flow between species and populations.	• Department of Conservation (Primary)
Buckley	**Speciation and adaptation in New Zealand Invertebrates** **Team members:** Thomas Buckley, Richard Newcomb and Howard Ross We are using genome scale sequence data coupled with RAD tag data to test hypotheses on stick insect and giant weta population genetics and speciation. Genome scale assemblies will be used as a template for mapping RAD tag data for detecting loci under selection. Population genetic processes will also be identified and used to inform conservation genetic strategies. The target species include island endemics, relictual taxa and more widespread species. Each of these comparisons will give novel insights into processes of adaptation.	• Ngatiwai, Department of Conservation (both (Primary)
Daugherty	**Genomics of iconic species** **Team members:** Hilary Miller (AWC postdoc), Charles Daugherty, Nicky Nelson We aim to undertake preliminary sequencing of the tuatara genome, focusing on two areas: (1) characterization of a partial transcriptome for tuatara, and (2) characterization of the Major Histocompatibility Complex (MHC) region. This project will enable us to assess the feasibility of assembling and annotating next-generation sequencing data from an evolutionarily distinct organism. For (1) we have used Illumina sequencing to obtain mRNA sequences from an early stage tuatara embryo, and are currently working on assembling and annotating these sequences to produce an Expressed Sequence Tag database for tuatara. In (2) we are collaborating with Scott Edwards at Harvard University to map the structure of the MHC, a genomic region which plays a central role in the immune system and provides a classic example of adaptive evolution. We have identified clones from a tuatara genomic library that span the core MHC region and are currently assembling sequence from these clones produced by GS-FLX 454 sequencing. Comparison of gene content in the tuatara MHC with that of other vertebrates will provide insight into the evolutionary history of this important genomic region.	
French	**Cows, starlings and *Campylobacter* in New Zealand: unifying phylogeny, genealogy and epidemiology to gain insight into pathogen evolution** **Team members:** Nigel French, Barbara Holland, Murray Cox In this study we use multilocus sequence typing and full genome analysis to discover how often, and how much, genetic material is exchanged between natural populations of the important zoonotic pathogens *Campylobacter jejuni* and *C. coli*. We examine how important recombination is relative to mutation for the emergence of new strains; and in which host species these events are most	• MAF-food safety • International infectious disease / pathogen evolution research community • Postgraduate students – training graduate students from NZ and internationally

Investigator	Project	Stakeholders
	likely to occur. Ultimately we can learn how and why *C. jejuni* emerged to become such a prominent human pathogen; anticipate further evolution and restrict the emergence and spread of new strains.	
Matisoo-Smith	**Development and Application of Next Generation Sequencing Technology to Highly Degraded Samples** **Team members**: Post Docs – Michael Knapp (AWC funded) and Ann Horsburgh; PhD students: Stefan Prost (AWC funded.), Karen Greig Much of the current work in our ancient DNA laboratory involves the development and refinement of next generation sequencing technology for use with degraded ancient samples, such as those recovered directly from archaeological sites in the Pacific (Knapp) and Africa (Horsburgh) or from museum collections. While next generation sequencing technology is well developed for generation of data using fresh tissue samples, there are significant problems inherent in the analysis of highly degraded samples, particularly when those samples derive from modern human remains. The combination of high levels of contamination by museum staff and archaeologists and extremely poor preservation require novel approaches for identifying authentic DNA sequences from archaeological remains.	• Colleagues working with ancient DNA – within the AWC, nationally and internationally • students – training graduate students from NZ and internationally • communities who are connected to the samples we are analysing (NZ & worldwide) • Archaeologists and other historic researchers • Museum staff (NZ & worldwide)
Rainey	Evolutionary implications of REPINs and their associated RAYTs **Team members:** Paul Rainey, Frederic Bertels (AWC student) and Xue-Xian Zhange (NZIAS) DNA sequences that copy themselves throughout genomes and make no specific contribution to reproductive success are by definition 'selfish'. Such DNA is a feature of the genomes of all organisms and evident by virtue of its repetitive nature. In bacteria the predominant repetitive sequences are short (~20 bp), extragenic and palindromic. These so-called REP sequences may occur many hundreds of times per genome, but their origins and means of dissemination have been a longstanding mystery. Recent work (with Frederic Bertels) has shown that REPs are components of higher order replicative entities termed REPINs that are themselves derived from the REP sequences that flanked an ancestral autonomous selfish element. In this ancestral state the REP sequences were critical for the movement of the selfish element but were devoid of any capacity to replicate independently. REPINs, on the other hand, have evolved to have a life of their own, albeit one that exploits – even enslaves – a genetic element upon which their existence depends. Current work seeks to understand the evolutionary origins and consequences of REPIN movement through experimental studies.	
Ritchie	**The Genetic Stock Structure of the Australasian School Shark (*Galeorhinus galeus*)** **Team members:** Sebastian Hernandez and Peter Ritchie School sharks are a bentho-pelagic species that typically inhabit the continental and insular coasts around the world. This species has been commercially fished in New Zealand since the early 1940s and has been considered a single biological stock. However, School Sharks have a high dispersal power and often found in widespread populations that are thought to be genetically homogeneous. In this project, we investigate the population genetic structure of School Sharks in the Australasian region, assessing whether this species comprises a single or multiple stocks in the Australasian region.. Given the commercial importance of *G. galeus* in Australasian region, and the concern regarding population levels (trans-Tasmanian and local scale), the knowledge of stock structure is essential for the effective management of fisheries. This project is being conducted in collaboration with Malcolm Francis at NIWA.	
Ritchie	**How Much Genetic Variation can a Wrasse Hold?** **Team members:** Xinkang (Jack) Du and Peter Ritchie The wrasses (family Labridae) are one of the most specious groups of fish. Two species are commonly found in shallow New Zealand waters: the spotty (*Notolabrus celidotus*) and the banded wrasse (*Notolabrus fucicola*). The spotty is endemic to New Zealand coastal areas and it is often found in high abundances in harbours, particularly near wharfs, and among the kelp on rocky reefs. We are using mitochondrial and microsatellite DNA markers to investigate the population genetic structure of the common spotty in New Zealand waters, and determine whether there is genetic differentiation among regions. Spotties are a protogynous hermaphrodite. All individuals begin life as females and only dominant ones in a	

Investigator	Project	Stakeholders
	group will change to males. We are also investigating the effect this sex-change system has on the effective population size.	
Ritchie (Pete)	**Detecting the Population Genetic Structure of Orange Roughy (*Hoplostethus atlanticus*) at Local and Global Scales** **Team members:** Andrea Varela and Peter Ritchie Orange roughy is one of New Zealand most important wild-capture fisheries. Orange roughy grow slowly, live longer (some over 100 years) and mature later than most other marine fishes (~30 years). Fecundity is low and populations have relatively low productivity. All these characteristics make orange roughy vulnerable to over-exploitation. Until now, all studies about the genetic differentiation of orange roughy populations have used different markers and the majority of them studied only specific areas. It is necessary to broaden the study of the population differentiation to a global scale. Our study incorporates samples from the complete geographic distribution of orange roughy: New Zealand, Australia, Tasmania, South Africa, the North Atlantic and possible Chile. We are using three techniques to investigate the genetic connectivity among populations: sequences of the mitochondrial DNA gene Cytochrome Oxidase I (COI), microsatellite DNA markers and candidate genes for selection. This project is being conducted in collaboration with Peter Smith at NIWA.	
Ross	**Estimation of the levels of confidence at which species of birds found in New Zealand may be identified by DNA Barcoding** **Team Members:** Dr Howard Ross, Dr Craig Millar, Selena Patel, Thomas Garden (summer student) Previously (Ross et al 2008 Systematic Biology 57:216-230) the relationship between the amount of genetic variation in a species and the reliability with which it could be identified by genetic methods was determined by simulation. We are applying this relationship to the DNA sequence data being generated as part of the Barcoding of the Birds of New Zealand to gain estimates of the ease with which New Zealand birds can be identified in this manner.	**Primary:** **Conservationists** and others documenting or monitoring biodiversity, because it will enable them to quantify the reliability of species identifications made using DNA-based methods.
Ross	**Re-examination of the frequency of species-level paraphyly in animals** In a highly cited review article, Funk & Omland (2003 Annual Review of Ecology, Evolution, and Systematics 34:397-423) reported a high incidence (23%) of non-monophyly in animal species, on the basis of a review of published studies. There is the potential for this estimate to be biased upwards as a result of the publishing process. I am using the Barcode of Life Database (BOLD) of cytochrome c oxidase subunit I (COI) to reassess this result. These sequences have been collected as a part of the international initiative to document diversity and so develop a species identification tool. Consequently the species representation will be influenced by different biases. It has been argued that high levels of non-monophyly seriously diminish the power of the DNA Barcode Initiative. I hope to also assess the impact of non-monophyly on the efficacy of DNA Barcoding in species identification.	**Primary:** **Evolutionary Biologists**, because it will increase our understanding of how biodiversity is patterned. **Conservationists** and others documenting or monitoring biodiversity, because it will increase our understanding of the reliability of DNA-based methods of species identification.
Waters	**Genetic analysis of Little Barrier Island lichen communities** **Team members**: Ben Myles (PhD student), A/Prof Jon Waters, A/Prof Alexei Drummond As part of the Centre's Model Ecosystem strategic initiative on Little Barrier Island, we are using genetic tools to characterize temporal and spatial variation in Little Barrier Island lichen communities across a range of upland sampling stations. This work will contribute substantial new data to enhance the understanding of NZ lichen biodiversity.	

Theme: **1. Genomics & Biomathematics**
Activity: C. Analytic and predictive modeling in ecology and evolution

Investigator	Project	Stakeholders
Rainey	**Evolution of multicellularity** **Team members:** Paul Rainey and Eric Libby (NZIAS post doc) The evolutionary transition from single cells to multicellularity posses significant theoretical and experimental challenges. Experimental analyses currently underway (with Caroline Rose and Katrin Hammerschimdt) have generated new insights that are currently being generalised via the construction of mathematical models and simulations. Of particular interest are frequency dependent coupling and negative feedback cycles between different developmental stages of organisms and their environmental effects.	
Spencer	I am working on several projects in this activity. I am supervising PhD student Jemma Geoghegan whose project concerns models of "population epigenetics," which incorporate epigenetic modifications into the standard models of population genetics. I continue to work on mathematical models that explain why natural populations of almost every species harbour significant levels of genetic variation. Finally, with collaborators Prof. Graeme Wake and Tony Pleasants, I am investigating mathematical models for optimal waiting and development times for plastic responses.	
Steel & Semple	**Modelling in biodiversity conservation and in speciation/extinction processes** **Team members***:* Charles Semple and Mike Steel Quantitative methods are widely used in conservation biology for deciding which collection of species should be conserved. AWC research has provided the first mathematically-rigorous method of selection under a model of biodiversity in which species features arise and disappear. In previous research we have shown that selections can be very different under this more realistic model when compared to using the standard model for which features never disappear. We have also applied ecological constraints to the questions of minimizing and predicting the expected loss of phylogenetic diversity in the near future given the current high-rate of species loss. Future work will include analysing the loss of phylogenetic diversity in more realistic speciation-extinction processes.	