# Doom 2007


## New Zealand's Phylogenetic Meeting

**11th-16th February 2007**
**Mount Ruapehu**

## List of attendees

| | | | |
|---|---|---|---|
| Baele, Guy | Belgium | Martin, William | Germany |
| Beiko, Robert | Canada | Martins, Leo | Japan |
| Bordewich, Magnus | UK | Matsen, Erick | NZ |
| Bromham, Lindell | Australia | Moulton, Vincent | UK |
| Bryant, David | NZ | Pagel, Mark | UK |
| Buckley, Thomas | NZ | Pardi, Fabio | UK |
| Currie, Thomas | UK | Penny, David | NZ |
| Darling, Aaron | USA | Phillips, Matt | NZ |
| Drummond, Alexei | NZ | Ranjard, Louis | NZ |
| Fischer, Mareike | NZ | Rosenberg, Noah | USA |
| Gascuel, Olivier | France | Ross, Howard | NZ |
| Gemmell, Neil | NZ | Rosskopf, Michael | Germany |
| Gernhard, Tanja | Germany | Roy, Scott | USA |
| Gray, Russell | NZ | Schliep, Klaus | NZ |
| Greenhill, Simon | NZ | Semple, Charles | NZ |
| Grunewald, Stefan | China | Shavit, Liat | NZ |
| Hartmann, Klaas | NZ | Simon, Chris | NZ |
| Hasegawa, Masami | Japan | Spillner, Andreas | UK |
| Heath, Tracy | USA | Steel, Mike | NZ |
| Hendy, Mike | NZ | Stevens, Mark I. | NZ |
| Hickey, Glenn | Canada | Thatte, Bhalchandra | NZ |
| Hingston, Melanie | NZ | Umehara, Satoshi | Japan |
| Holland, Barbara | NZ | Voelckel, Claudia | NZ |
| Humphries, Peter | NZ | von Haeseler, Arndt | Austria |
| Jermiin, Lars | Australia | Waddell, Peter | Japan |
| Joly, Simon | NZ | Will, Olivier | NZ |
| Kishino, Hirohisa | Japan | Wong, Dennis | Canada |
| Linz, Simone | Germany | Woodhams, Michael | NZ |
| Lockhart, Peter | NZ | | |

## Organisers

### General
Peter Lockhart
Joy Wood
Richard Carter
Mort Piripi

### Scientific Program
Simon Joly

### Administration
Karen Sinclair
Susan Adams

# Useful Information

## Activities
Tracks and walks – Some of the best walks in the country are in the National Park area, including the famous Tongariro crossing. A variety of time and difficulty levels are available, enquire at the Whakapapa visitors' centre for details. See the end of booklet for "Activities in and around Tongariro National Park"

Information on other activities such as rafting and guided tours can also be obtained at the visitor's centre.

## Dining
Please note that you will need to prepare your own dinner on Sunday, Monday, Tuesday and Wednesday evenings. Kitchen facilities are available - but bring your own food to prepare as groceries are very thin on the Volcanic Plateau (note that there is little freezer space!).

There are a few places where to eat in Whakapapa Village. There are also several eateries in nearby National Park township, some are listed below (it is advisable to phone for bookings, see end of booklet for menus):

Whakapapa Village:
The Skotel Restaurant
The Café at the Bayview Chateau

National Park:
Eivin's Cafe Bar, S.H. 4 – 07 892 2843
Schnapps, just off S.H. 4 – 07 892 2788
Station Café, National Park Railway Station, Station Road 07 892 2881

## Supermarkets
There are no supermarkets at National Park or in Whakapapa Village. Some items can be purchased at the Whakakapa Holiday Park "dairy" (it houses no cows but does sell dairy products), but if you plan to do a big grocery shop do it before you travel or on route to the Volcanic Plateau. Ohakune to the south and Turangi to the north both have supermarkets.

## Banks
There are **NO** banks in either National Park or Whakapapa Village (the closest are in Turangi and Ohakune ). EFTPOS facilities are available in Whakapapa Village: at the pub and cafes, resturants and at the Holiday Park dairy. Master/Visa cards are also accepted at most places.

## Internet Access
Internet access is available at the Skotel. There are 5 internet terminals, and a new "untested" wireless service in the bar. You have to buy time from the reception ($5 for 15 minutes).

## Sunday February 11th

| | |
|---|---|
| 20h00 – 21h30 | **Registration** |

## Monday February 12th

| | |
|---|---|
| 9h00 | **Announcements** |
| **PHYLOGENETIC DIVERSITY** | |
| 9h10 | **Phylogenetic Diversity on Split Systems**<br>*Minh, B.Q., S. Klaere & A. von Haeseler** |
| 9h30 | **Taxon Selection for Maximising Phylogenetic Diversity**<br>*Pardi, F.* & N. Goldman* |
| 9h50 | **Phylogenetics and Biodiversity Conservation**<br>*Hartmann, K.* & M. Steel* |
| 10h10 | **Optimizing Phylogenetic Diversity under Constraints**<br>*Moulton, V.,* C. Semple & M. Steel* |
| 10h30 | **Morning Tea** |
| 11h00 | **Testing the power of genetic methods of species identification**<br>*Ross, H.A.* |
| **HYBRIDIZATION, RECOMBINATION AND LATERAL GENE TRANSFER** | |
| 11h20 | **How much hybridisation is enough?**<br>*Humphries, P.J.* |
| 11h40 | **Using reductions to compute the minimum number of hybridization events**<br>*Bordewich, M., S. Linz,* K. St. John & C. Semple* |
| 12h00 | **Lunch** |
| 14h00 | **Counting coalescent histories**<br>*Rosenberg, N.A.* |
| 14h20 | **Non-monophyletic species: testing for evidence of hybridization**<br>*Joly, S.* & P.J. Lockhart* |
| 14h40 | **Detecting hybridisation in collections of gene trees**<br>*Holland, B.,* P. Lockhart, K. Huber & V. Moulton* |
| 15h00 | **New results on the SPR distance**<br>*Grünewald, S.* |
| 15h20 | **Afternoon Tea** |
| 16h00 | **The Complexity of Unrooted SPR Distance**<br>*Hickey, G.,* F. Dehne, A. Rau-Chaplin, C. Blouin* |
| 16h20 | **Searches through subtree prune-and-regraft space (Don't try this in Fangorn)**<br>*Beiko, R.G.* & N. Hamilton* |
| 16h40 | **Bayesian recombination detection based on distance between trees**<br>*Martins, L.O.,* E.S. Leal & H. Kishino* |
| 17h00 | **End of Activities** |

## Tuesday February 13th

| | |
|---|---|
| 7h00 – 9h00 | **Breakfast at the Chateau** |
| 9h00 | **Announcements** |
| **DATES AND RATES** | |
| 9h10 | **Phylogenies reveal constant rates of speciation**<br>*Pagel, M.,* A. Meade & C. Venditti* |
| 9h30 | **Out of Australia: dating the origin and diversification of the worldwide cicada tribe *Cicadettini* and its connection to the aridification of the Southern Hemisphere (and some interesting things about among-site rate variation)**<br>*Vanderpool, D., D.C. Marshall, K.B.R. Hill, and C. Simon** |
| 9h50 | **One Rate to Rule Them All and at the Ancestor Bind Them**<br>*Drummond. A.J.* & M.A. Suchard* |
| 10h10 | **Glimpses into the strange world of phylogenetic mixtures**<br>*Matsen, F.A.* & M. Steel* |
| 10h30 | **Morning Tea** |
| 11h00 | **How useful are lineages through time plots for examining the tempo of diversification?**<br>*Wong, D.H.J.* & S.B. Heard* |
| 11h20 | **Estimating speciation times under neutral models**<br>*Gernhard, T.* |
| 11h40 | **The biology of the molecular clock**<br>*Bromham, L.* |
| 12h00 | **Lunch** |
| 14h00 | **Lineage Specific Sequence Evolution**<br>*Shavit, L.,* B. Holland, M. Hendy & D. Penny* |
| **ROOTING DIFFICULT TREES** | |
| 14h20 | **Rooting the eutherian tree-The power and pitfalls of phylogenomics**<br>*Hasegawa M.,* H. Nishihara & N. Okada* |
| 14h40 | **Highly Conserved Indels Support Atlantogenata at the Root of Placental Mammals**<br>*Waddell, P.J.,* S. Umehara, K.-C. Griche & H. Kishino* |
| 15h00 | **Quantitative analysis and modelling of indels from a 2mb region in ten primates**<br>*Umehara, S.,* P.J. Waddell & H. Kishino* |
| 15h20 | **Afternoon Tea** |
| **MITOCHONDRIA: HETEROPLASMY AND RECOMBINATION** | |
| 16h00 | **Mathematically Modelling Mitochondrial Mixing**<br>*Hendy, M.D.* & M.D. Woodhams* |
| 16h20 | **Blood of the Matriarchs: The descent of mitochondrial heteroplasmy.**<br>*Woodhams, M.* |
| 16h40 | **Good molecules go bad, but does it matter?**<br>*Gemmell, N.J.* |
| 17h00 | **End of Activities** |

## Wednesday February 14th

| | |
|---|---|
| 9h00 | **Free Activities** |
| 17h00 | **BBQ** |

## Thursday February 15th

| | |
|---|---|
| 9h00 | **Announcements** |
| **GRAPHS, TREES AND DISTANCES** | |
| 9h10 | **Quality of maximum likelihood estimates**<br>*Schliep, K.* |
| 9h30 | **Realizing Phylogenetic Trees with Weighted Quartets**<br>*Grunewald, S., K.T. Huber, V. Moulton & C. Semple** |
| 9h50 | **Computing planar split graphs**<br>*Bryant, D., V. Moulton & A. Spillner** |
| 10h10 | **Is there a star paradox?**<br>*Steel, M.* & E. Matsen* |
| 10h30 | **Morning Tea** |
| 11h00 | **Perfectly Misleading Distances from Ternary Characters**<br>*Fischer, M.* & H.-J. Bandelt* |
| 11h20 | **Some developments on the consistency of Balanced Minimum Evolution Algorithms**<br>*Bordewich, M.,* V. Berry, O. Gascuel, K. Huber & V. Moulton* |
| 11h40 | **The effect of reduced taxon sampling on phylogenetic tree imbalance**<br>*Heath, T.A.,* D.J. Zwickl, J. Kim & D.M. Hillis* |
| 12h00 | **Lunch** |
| **PACIFIC SETTLEMENT** | |
| 14h00 | **Bayesian coalescent analysis of mtDNA diversity reveals major Southern Asian phase in human prehistory**<br>*Atkinson, Q.D., A.J. Drummond & R.D. Gray** |
| 14h20 | **Pacific settlement and Austronesian languages.**<br>*Greenhill, S.J.* & R.D. Gray* |
| 14h40 | **Spatial and temporal distribution of *Rattus exulans* in Near Oceania**<br>*Hingston, M.,* H.A. Ross, E. Matisoo-Smith & J.H. Robins* |
| 15h00 | **The Evolution of Social Stratification in the Pacific: A Phylogenetic Approach**<br>*Currie, T.* |
| 15h20 | **Afternoon Tea** |
| **ORGANISMS** | |
| 16h00 | **Marsupial "herbivore" phylogeny and the failure of non-evolutionary morphological analysis**<br>*Phillips, M.J.* |
| 16h20 | **Meme analysis of bird songs**<br>*Ranjard, L.* & H. Ross* |
| 16h40 | **The Jekyll and Hyde lifestyle of bacteria revealed by genome phylogenies**<br>*Darling, A. E.* & M.A. Ragan* |
| 17h00 | **End of Activities** |

# Friday February 16th

| | |
|---|---|
| 9h00 | **Announcements** |
| **MOLECULAR AND GENOME EVOLUTION** | |
| 9h10 | **Accounting for Exposition and Secondary Structure in Protein Evolution: Models and Gain**<br>*Le, S.Q. & O. Gascuel*[*] |
| 9h30 | **Disentangling selection and mutation in comparative genomics: the greater-than-equal and opposite 'rule'**<br>*Roy, S.W.* |
| 9h50 | **Detecting Asymmetric Markov Processes in Aligned Sequence Data**<br>*Jermiin, L. S.,*[*] *F. Ababneh, C. Ma & J. Robinson* |
| 10h10 | **A maximum-likelihood framework for gene expression evolution models with mutational and non-mutational effects**<br>*Rosskopf, M.*[*] *& A. von Haeseler* |
| 10h30 | **Morning Tea** |
| 11h00 | **Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution**<br>*Dagan, T. & W. Martin*[*] |
| 11h20 | **Modeling context-dependent evolution: the influence of the immediate flanking bases**<br>*Baele G.,*[*] *Y. Van de Peer & S. Vansteelandt* |
| 11h40 | **Is Microevolution sufficient for Macroevolution; birds and mammals.**<br>*David Penny* |
| 12h00 | **Lunch – End of Activities** |

# Phylogenetic diversity

## Phylogenetic Diversity on Split Systems

Minh, B.Q., S. Klaere & A. von Haeseler*

*Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, University of Vienna, Medical University of Vienna, Veterinary University of Vienna*

One of the questions of biodiversity conservation is to select taxa under constraints on the number of taxa. One of the optimization strategies employs phylogenetic diversity (PD), where given a tree one selects the taxa such that the induced subtree is maximized. Therefore the optimal set strongly depends on the selected tree. However, with the increase in sequence data incongruent gene trees are observed and therefore distinct PD sets depend on the selected tree. Thus the choice of the gene will influence conservation strategies. A way to include incongruent information is the cumulation of available information in so-called split systems. However, constructing PD-sets for split systems cannot be accomplished by a greedy strategy. We extend the notion of PD to split systems and provide a dynamic programming algorithm to obtain optimal PD taxon sets for the special class of split systems which imply a circular order of taxa.

## Taxon Selection for Maximising Phylogenetic Diversity

Pardi, F.* & N. Goldman

*European Bioinformatics Institute, Cambridge, UK*

The phylogenetic diversity (PD) of a set of taxonomical units (e.g. genes, individuals, populations, species) is the total length of the evolutionary tree connecting them. This measure is relevant for directing policy decisions in a variety of applications, especially in biodiversity conservation, where it is essential to conserve phylogenetically diverse sets of species, and comparative genomics, where the effectiveness of many methods (e.g. to detect functional elements) is strongly correlated with the total PD of the sequences being compared. I will present a hierarchy of optimisation problems where the aim is to select, from a collection of taxa, a subset with maximum total PD. When this subset is constrained to have a fixed number of taxa, a simple greedy algorithm can be shown to produce optimal solutions. When instead it is the limited availability of an underlying resource (money, time, etc.) that limits selection, and taxon-specific requirements are quantifiable, a dynamic programming algorithm is instead necessary. Our ideas can also be applied to a framework for bioconservation proposed by economists, called the Noah's Ark Problem, which explicitly models species-specific extinction probabilities and the effects that conservation efforts have on them. The objective now is to maximise the (expected) PD of the species that will survive. The algorithms mentioned above can be adapted to efficiently solve some special cases of this problem. Solving the most general case is also possible, albeit at the expense of some efficiency.

## Phylogenetics and Biodiversity Conservation

Hartmann, K.* & Steel, M.

*Allan Wilson Centre for Molecular Ecology and Evolution, University of Canterbury, New Zealand*

Conservation organisations are often faced with the problem of allocating limited funds to the conservation of many needy species. The ultimate goal is to ensure that the expected future biodiversity is as high as possible. Other factors that need to be considered are the costs of conserving different species and the expected increase in their survival probability that conservation would provide them. The Noah's Ark Problem is a mathematical framework that formalises this problem and uses phylogenetic diversity as a measure of biodiversity. In this talk I consider various algorithmic aspects of this problem and provide an extension that permits uncertain survival probabilities to be incorporated. This extension shows that a frequently made assumption about the survival probabilities provides poor solutions.

## Optimizing Phylogenetic Diversity Under Constraints

Moulton, V.[*§], C. Semple[†] & M. Steel[†]

[§] *School of Computing Sciences, University of East Anglia, Norwich, UK;* [†] *Biomathematics Research Centre, University of Canterbury, Christchurch, NZ*

Phylogenetic diversity (PD) is a measure of the extent to which different subsets of taxa span an evolutionary tree, and provides a quantitative tool for studying biodiversity conservation. Recently, it was shown that the problem of finding subsets of taxa of given size to maximize PD can be efficiently solved by a greedy algorithm. In this talk, we describe some extensions of this work, showing that, although the greedy algorithm can be used to solve some special cases of the PD optimization problem, some measures related to PD fail to be optimized by a greedy algorithm, and an extension of the PD optimization problem to a phylogeographic setting is NP-hard.

### Testing the power of genetic methods of species identification

Ross, H.A.

*School of Biological Sciences, University of Auckland, Auckland, New Zealand*

The species identification of unknown specimens using methods based on DNA sequences is becoming increasingly popular. Although several possible methods have been identified, including BLAST, DNA Barcoding and phylogenetic placement, their relative abilities to produce correct identifications are only now being assessed. Datasets of DNA sequences were simulated to resemble a set of sequences from a family of species in GenBank. The methods named were assessed for the frequency with which they gave true positive identifications, for their sensitivity to paraphyly (incomplete lineage sorting), and for the frequency of false positive identification arising from cryptic species or incomplete taxonomic sampling. When all species are represented in the reference database, identifications based on BLAST, DNA Barcoding and phylogenetic placement in a mono-species clade were equally accurate, and equally influenced by paraphyly. A more strict phylogenetic criterion, requiring that the query be within, but not basal to, the mono-species clade produced more conservative results and returned more uncertain identifications. When the true species was not in the reference database, only the method based on the strict phylogenetic criterion was relatively resistant to false positive identification. Method choice can therefore be based on computational efficiency or the desire to minimize false positive identification, according to the completeness of the reference dataset.

## Hybridization, Recombination and Lateral Gene Transfer

### How much hybridisation is enough?

Humphries, P.J.

*Department of Mathematics and Statistics, University of Canterbury, New Zealand*

A commonly used measure for comparing two rooted trees on the same leaf-set is the SPR distance, or equivalently the size of a maximum agreement forest (MAF). More applicable to biological settings is the hybridisation number, which has been shown to equal the size of a maximum acyclic agreement forest (MAAF). We present the results of some recent work on the size of an MAAF and on the size difference between an MAF and an MAAF for a given pair of trees.

### Using reductions to compute the minimum number of hybridization events

Bordewich, M.,[§] S. Linz,[*†] K. St. John[‡] & C. Semple[†]

[§] *Department of Computer Science, Durham University, Durham DH1 3LE, United Kingdom;* [†] *Biomathematics Research Centre, Department of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand;* [‡] *Department of Mathematics and Computer Science, Lehman College, City University of New York, USA*

Since not all groups of taxa (e.g. certain groups of plants and fish) evolve completely tree-like, reticulation events - like hybridization and horizontal gene transfer - play a fundamental role in evolution. As a result of such events, some species are a mixture of genes derived from different ancestors. Assuming that the initial data set is correct and that hybridization is the only cause of incongruence between the gene trees reconstructed for this set, a basic problem for

biologists is to compute the minimum number of hybridization events to explain the evolutionary history of the species under consideration.

In this talk, we present a newly developed algorithm - based on three reduction rules - to compute the minimum number of hybridization events for two rooted binary phylogenetic trees. Mathematically speaking, the problem is NP-hard. Nevertheless, provided the two input trees share a number of common features (which is likely for many biological examples), the algorithm runs efficiently and gives the exact solution. We illustrate our algorithm on a grass data set.

## Counting coalescent histories

Rosenberg, N.A.

*Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA*

Given a species tree and a gene tree, a coalescent history is a list of the branches of the species tree on which coalescences in the gene tree take place. I develop a recursion for the number of valid coalescent histories that exist for an arbitrary gene tree/species tree pair, when one gene lineage is studied per species. The result is obtained by defining a concept of m-extended coalescent histories, enumerating and counting these histories, and taking the special case of m=1. As a sum over coalescent histories appears in the Degnan & Salter (2005) formula for the probability that a random gene tree evolving along the branches of a fixed species tree has a specified labeled topology, the enumeration of coalescent histories can reduce the effort required for evaluating this formula.

## Non-monophyletic species: testing for evidence of hybridization

Joly, S.[*] & P.J. Lockhart

*Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, New Zealand*

Examples of non-monophyletic species are common, both for plants and animals. In cases where we can exclude horizontal transfer and gene duplication as a source of phylogenetic incongruence, incomplete lineage sorting and hybridization remain possible explanations for non-monophyly. Until now, attempts to differentiate these evolutionary processes have not been successful. However, it should be possible to distinguish these processes, in at least some cases. An expectation of lineage sorting, but not of hybridization, is that divergence times of the sequences are greater or equal to the divergence times of the species. This observation can be used to identify events of hybridization, but how can this be tested statistically? Here, we use a Bayesian approach to estimate of divergence times and ancestral population sizes for a species tree. Then, gene trees are generated using the coalescent (which models only lineage sorting) on this tree in order to generate a null distribution that can be used to test potential hybridization events. This method is evaluated on NZ alpine buttercups (*Ranunculus*) and on North American roses (*Rosa*).

## Detecting hybridisation in collections of gene trees

Holland, B.,[*§] P.J. Lockhart,[§] K. Huber[†] & V. Moulton[†]

[§] *Allan Wilson Centre, Massey University, New Zealand;* [†] *University of East Anglia, UK*

New phylogenetic network methods have recently been developed that take gene trees as input and produce splits graphs capable of displaying conflicting splits. We discuss the application of these methods for studying hybridisation. We show that in some circumstances these methods are capable of distinguishing incongruence amongst gene trees resulting from hybridisation events from incongruence resulting from phylogenetic error and lineage sorting.

## New results on the SPR distance

Grünewald, S.

The subtree prune and regraft (SPR) distance has been used by phylogenetic tree search methods and to estimate the number of reticulate evolution events that must have occured to explain two different trees on the same taxa. However,

not much is known about the distribution of the SPR distance in the space of all trees with given taxa set and even the diameter, i.e. the maximal possible SPR distance between two trees with n taxa, is unknown. The best published lower and upper bounds are n/2 and n-2, respectively. I will present an improvement of both bounds and show that the diameter is n-o(n).

## The Complexity of Unrooted SPR Distance

Hickey, G.,[*§] F. Dehne,[§] A. Rau-Chaplin[†] & C. Blouin[†]

[§] *School of Computer Science, Carleton University, Canada;* [†] *Faculty of Computer Science, Dalhousie University, Canada*

The Subtree Prune and Regraft (SPR) operation induces a distance metric on phylogenetic trees. This metric is of interest because each SPR operation can be used to model a Lateral Gene Transfer (LGT) event, equating the SPR distance between two trees with the size of the most parsimonious LGT scenario that reconciles them. The complexity of the SPR distance problem has a storied history: the original proof of NP-Hardness has been shown false but the idea behind it has since been used to show that TBR and, more recently, rooted SPR distance are NP-Hard. As many inference algorithms generate unrooted trees, the complexity of unrooted SPR has remained of interest. We show, using techniques similar to those found in the previous work, that the problem is indeed NP-Hard as well. There has been little previous applied work on computing exact unrooted SPR distances, most existing algorithms being heuristic and/or designed to compute variants of this metric. We designed and implemented a novel, exact algorithm and tested it on a large dataset of trees inferred from real proteins as well as randomly generated data. The results of these experiments lead us to conclude that exact SPR distance computation is achievable in many practical cases. We additionally show that our experimental results support the conjecture that the problem is fixed parameter tractable with respect to the distance, as well as that the popular heuristic of decomposing input trees according to shared splits is not always valid.

## Searches through subtree prune-and-regraft space (Don't try this in Fangorn)

Beiko, R.G.[*§†] & N. Hamilton[†]

[§] *Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada, and Genome Atlantic;* [†] *Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia*

One of the myriad causes of disagreements among gene trees is the phenomenon of gene sharing or lateral genetic transfer (LGT). A 'reference' scenario of vertical descent (e.g., a tree describing the organismal history) can be used to infer LGT events from discordant gene trees: this is because an LGT event induces a subtree prune-and-regraft (SPR) operation on the reference tree. Reconciliation of a reference tree with a gene tree can be achieved by 'unwinding' the differences using a series of SPR operations. Each SPR event in such a series represents a potential transfer of the gene's evolutionary history. Unfortunately, this procedure is limited in certain fundamental ways, and increasing numbers of taxa in the gene tree yield a rapid increase in the number of SPR operations that must be considered. We have developed EEEP (Efficient Evaluation of Edit Paths), a method that applies reasonable biological constraints on the SPR search space. EEEP also has a number of heuristics that return optimal solutions in most cases. EEEP has been compared in principle and in practice to two other SPR search methods, HorizStory and LatTrans, and has been applied to the problem of LGT recovery to a data set of over 22,000 trees. In this presentation, I will describe the workings of EEEP, our practical insights into the problem, and challenges in interpreting the results of the analysis.

## Bayesian recombination detection based on distance between trees

Martins, L.O.,[*§] E.S. Leal[†] & H. Kishino[§]

[§] *Graduate School of Agriculture and Life Sciences, University of Tokyo, Japan;* [†] *DIPA, Federal University of Sao Paulo, Brasil*

Recombinant DNA sequences can not be represented by a single topology since regions participating recombination will support distinct evolutionary histories. Nonetheless, existing methods for recombination detection are reliable only for a limited number of taxa, constraining the recombination analysis to cases where the phylogeny can be assumed known for most sequences. If the number of putative recombinants increases the analysis needs to be conducted independently for each query sequence, neglecting potential recombinations between them. We newly developed a distance measure between tree topology pairs that closely resembles the number of recombinations. By introducing a prior distribution on these recombination distances between neighboring DNA segments, a Bayesian hierarchical model

was devised to detect inconsistencies due to recombinations. The method simultaneously estimates the distribution of recombination break-points along the genomes and the pattern of recombination events. We applied the procedure on simulated datasets and on a genomic HIV-1 dataset, where we found evidence for recombination hot-spots and intra-subtype recombination. The inference of recombination hot-spots follows directly from the approximate recombination distance between segments.

# Dates and Rates

## Phylogenies reveal constant rates of speciation

Pagel, M.,* A. Meade & C. Venditti

*School of Biological Sciences, University of Reading, Reading, England*

The lengths of the branches of phylogenetic trees record the amount of evolution between inferred speciation events. Frequency distributions of phylogenetic branch lengths typically decline precipitously from a large number of short branches to a small number of longer branches. We derive four models to characterize these frequency distributions: a constant rate of speciation model, speciation allowing for incomplete sampling, a variable rate of speciation model, and a model allowing for variable rates of speciation and incomplete sampling. We apply the models to 130 phylogenetic trees comprising narrow taxonomic ranges of well- sampled species. We find that the constant rate of speciation model provides a remarkably good fit to a large number of these data sets (63%) with the variable rate of speciation model fitting a further 12%. The two models incorporating incomplete sampling fit the remaining 25% of the data sets. These results are surprising for their uniformity, and support the Red Queen view of speciation – that species do not get better adapted the longer they live. They are also interesting in the light of the common observation that phylogenetic trees are less balanced in the number of species in daughter clades than expected from null models of speciation. Constant rates of speciation should, other things equal produce balanced trees. We discuss the possibility that speciation rates may be heritable. This can produce exponentially or approximately exponential distributions of branch lengths and yield unbalanced trees.

## Out of Australia: Dating the origin and diversification of the worldwide cicada tribe Cicadettini and its connection to the aridification of the Southern Hemisphere (and some interesting things about among- site rate variation)

Vanderpool, D., D.C. Marshall, K.B.R. Hill, and C. Simon*

*Dept. of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269 USA*

The worldwide cicada tribe Cicadettini is found on all temperate continents except South America and is most diverse in temperate and subtropical semiarid habitats, especially grass, scrub, and open woodland. Cicadettine cicadas appear to have diversified most extensively in Australia, where perhaps two-thirds of world species are found. Although one of just 12 Australian cicada tribes, the tribe contains approximately two-thirds of Australia's extensive cicada fauna, with the majority of those species undescribed. We have constructed Bayesian and maximum-likelihood molecular phylogenies of the tribe using nuclear (EF1-a) and mitochondrial (COI, COII) genetic datasets from up to 150 taxa. The results suggest that the main Cicadettini lineages originated in Australia and possibly New Caledonia and spread to their current worldwide distribution during a rapid radiation. We then used Bayesian and penalized-likelihood relaxed-clock dating techniques together with independent geological calibrations to explore the timing of the dramatic cicadettine radiation and its possible connections to the aridification of Australia. The results suggest that the cicadettine radiation began only 15-20 mya, coincident with the late Tertiary and Quaternary intensification of aridification in Australia and worldwide. Cicadettine lineages appear to have spread at least twice from Australia into Asia and then Europe. One of these lineages also colonized Africa and the other North America. The post-Gondwanan origin of the tribe and the lack of affinity of the group for tropical habitats (separating North and South America) might explain the absence of Cicadettini from South America. Cicadettine cicadas have twice colonized New Zealand from Australia+New Caledonia, with one of these lineages diversifying extensively in the late Pliocene and Pleistocene into approximately 55 species, in concert with extensive habitat diversification.

## One Rate to Rule Them All and at the Ancestor Bind Them

Drummond. A.J.*§ & M.A. Suchard†

[§] *Department of Computer Science, University of Auckland, New Zealand;* [†] *Department of Biomathematics, UCLA, USA*

We present a new method for relaxing the assumption of a strict molecular clock using Markov chain Monte Carlo called the Random Local Molecular clock model. The new method approaches the problem by proposing a series of local molecular clocks, each extending over a subregion of the full phylogeny. Each branch in a phylogeny (representing a split) is a possible location for a change of rate from one local clock to a new one. So there are 2n-3 models that have exactly two local molecular clocks. Thus including both the global clock and the unconstrained model results in 2^{2n-3) possible rate models covering models with 1, 2, …, 2n-3 different rate categories. We demonstrate an efficient method to sample this space while simultaneously estimating the phylogeny. The new method allows a direct test of the strict molecular clock with all possible partitions of the phylogeny into local clocks. We demonstrate its utility on a number of example data sets, demonstrating that many large datasets only require a small number of local molecular clocks to reconcile their branch lengths with a time scale. Finally we explore methods to visualize the complex posterior distribution that results from inference under such models.

## Glimpses into the strange world of phylogenetic mixtures

Matsen, F.A.* & M. Steel

*Biomathematics Research Centre, Canterbury University, Christchurch, New Zealand*

Rates of evolution and evolutionary history may vary across the genome of a given organism. One way of modelling the resulting data is a "mixture model" where a given phylogenetic pattern comes from one of a collection of trees with fixed probabilities. In this talk I will focus on recent results (joint with Mike Steel) which investigate what pattern probabilities are and are not possible when rates of evolution vary on a single topology. In particular, we find some surprising results concerning when the pattern probabilities for a tree of a given topology can be exactly mimicked by mixtures of trees with the same or a star topology. Some highlights include the creation of "phantom" internal edges (edges not appearing in any of the mixed trees) and pendant edges longer than the maximum of the corresponding pendant edges on the mixed trees.

## How useful are lineages through time plots for examining the tempo of diversification?

Wong, D.H.J.* & S.B. Heard

*Biology Department, University of New Brunswick*

Phylogenies are an important tool in our enterprise to understand the evolution of biological diversity. To extract the signal of macroevolutionary processes from a phylogeny, a variety of approaches have been developed that can be compared to expectations generated from simulation. Statistics such as imbalance or gamma provide a single number to summarize macroevolutionary process. Lineage accumulation plots are translated directly from a phylogeny and provide a visual interpretation of diversification—features in the shape of the plot presumably provide insight into macroevolutionary processes. Without physical evidence to corroborate, what can an examination of the shape of an LTT plot tell us about diversification? We use a simulation approach to show that LTT plots generally have low power to distinguish different models of diversification from each other and from the simplest null model (equal rates Markov). In addition to the mechanisms of macroevolution, we also explore the parameters that influence power such as the strength of extinction, the size of the phylogenies and the presence or absence of extinct lineages.

## Estimating speciation times under neutral models

Gernhard, T.

*Mathematics Department, TU Munich, Germany*

Recently, another neutral model for speciation, the critical branching process model, was introduced by D. Aldous and L. Popovic. We provide exact results for the distribution of the time of the $k$-th speciation event and obtain easy formulas for its expectation and its variance. With these formulas, we can calculate the expected time and the variance for any speciation event in a given phylogenetic tree, i.e. we can date all the inner nodes. The obtained values are compared to the expectation and the variance for the time of a speciation event in a phylogenetic tree under the popular Yule model, a neutral model without extinction. Methods for dating a speciation event become valuable, if for

constructed phylogenetic trees, no time scale is available. A missing time scale could be due to supertree methods, morphological data or molecular data which violate the molecular clock.

### The biology of the molecular clock

Bromham, L.

*Centre for Macroevolution and Macroecology, School of Botany and Zoology, Australian national University, Canberra, ACT 0200 Australia*

Rates of molecular evolution can vary substantially between lineages. This observation has important implications for the use of DNA sequence data to investigate evolutionary and ecological patterns, through molecular phylogenies and molecular dating. Molecular dating methods have been developed that allow for rate change over phylogenies, but how do we know if the assumptions that these "relaxed clock" methods make reflect real patterns of genomic evolution? I use a comparative approach, comparing a range of DNA sequences from a wide array of taxa, to ask which species characteristics, or modes of evolution, influence the rate of molecular evolution, and to examine what effect these factors may have on our ability to infer ecological and evolutionary patterns and processes from DNA sequences.

### Lineage Specific Sequence Evolution

Shavit, L.,[*] B. Holland, M. Hendy and D. Penny

*Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, New Zealand*

It has become increasingly evident that as lineages diverge they acquire their own properties, the constraints on the sequences change and the proportion of variable sites may not be constant among different lineages. For long timescales, the historical signal that is left in the sequences is expected to be relatively small. Accordingly, we expect lineage specific effects to become more prominent. The models used in phylogenetic estimation assume a constant process over the whole tree and do not allow for lineage-specific evolution. However, methods for inferring phylogenies construct trees for any input sequence data regardless of whether it is meaningful. An interesting question is, whether we are able to detect cases in which methods are confounded by lineage-specific effects, or will we naively trust a wrong tree. We discuss our strategy in understanding the significance of ignoring these lineage-specific processes when estimating phylogenies over progressively longer time scales. A wrap-around for Seqgen that allows the generation of sequence data which contains lineage-specific effects will be presented.

## Rooting Difficult Trees

### Rooting the eutherian tree-The power and pitfalls of phylogenomics

Hasegawa, M.[*§] H. Nishihara,[§†] & N. Okada[†]

[§] *Institute of Statistical Mathematics, Tokyo, Japan;* [†] *Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan*

In the post-genomic era, genome-scale approaches to phylogenetic inference (phylogenomics) are being applied extensively to overcome the large sampling errors inherent in commonly used approaches based on a single or a small number of genes. Sampling error vanishes as the number of genes provided for the analysis increases, but the fully resolved tree can still be wrong if the phylogenetic inference is biased (systematic error), and several such cases have been reported. Thus, to root the eutherian tree using genome-scale data and the maximum likelihood method, we show a case in which a concatenate analysis strongly supports a putatively wrong tree, whereas the total evaluation of separate analyses of different genes grossly reduced the bias of the phylogenetic inference. Although a conventional method of concatenate analysis of nucleotide sequences from our dataset of more than 1 Mbp alignment of 2,789 nuclear genes suggests a misled monophyly of Afrotheria (e.g., elephant) and Xenarthra (e.g., armadillo) with 100% BP, this tree is not supported by our "Separate Method" that takes into account the different tempos and modes of evolution among genes, and instead the basal Afrotheria tree is supported with 86-95% BP depending on the assumed model. Our analysis demonstrates that, in cases of very large variation in the evolutionary features among different genes, the separate model, rather than the concatenate model, should be used for phylogenetic inference, especially for genome-scale data.

### Highly Conserved Indels Support Atlantogenata at the Root of Placental Mammals

Waddell, P.J.,[*§†] S. Umehara,[§] K.-C. Griche,[‡] & H. Kishino[§]

[§] *Laboratory of Biometry and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657;* [†] *SCCC, USC, Columbia, SC 29208, USA;* [‡] *C/- Hiro Kishino, Rm 206, Building 1, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657*

Recently sequenced genomes of mammals and other vertebrates potentially contain many rarely occurring indels. These should provide an abundant source of phylogenetic characters with which to test difficult to resolve parts of the tree of placental (eutherian) mammals. One of the hardest and most important parts of the tree is the location of the root. While the hypothesis Boreotheria (also called Boreoeutheria) is consistent with LINE insertion data, it is unclear which of the three hypotheses, Epitheria (Xenarthra diverging first), Atlantogenata (Xenarthra with Afrotheria) or Exafricoplacentalia (Afrotheria first) is correct. Here we test these hypotheses using a log likelihood ratio test and highly conserved well-aligned indels that are five base pairs or larger in size. Finding a data configuration of 51 for Boreotheria with just 2 and 1 supporting alternatives, and finding 4 15 3 for Epitheria, Atlantogenata and Exafricoplacentalia, respectively, our tests reject the alternatives to Boreotheria ($p << 0.05$) and Atlantogenata ($p < 0.05$), so the root of placentals appears to fall between these two main lineages of placental mammals. Exploratory analyses of large collections of indels reveal some important possibilities. One of these is that the duration of the two main lineages of placentals may have been very brief, possibly ~ 1 million years or so. Another is that indels, which are located some distance from other indels, indeed do seem to show a much lower homoplasy (C.I. e 0.8) than sequence data (C.I. ~ 0.5-0.6). On a more cautionary note, there are also indications of unexpected biases in the homology and alignment of the genomic data. Given these caveats, it is important that maximally independent types of low-homoplasy molecular data corroborate all major clades of mammals. Estimating the ancestral population sizes of mammals at such an ancient age is shown to be particularly difficult, but potentially soluble. Taking account of biases in estimating the proportion of characters involved in ancestral sorting and in the lengths of deep internal edges, are essential. The resolved root of placentals is important to reconstructions of ancestral characters and genomes, and should yield important insights in these areas.

### Quantitative Analysis and Modelling of Indels from a 2mb Region in Ten Primates

Umehara, S.,[*§] P.J. Waddell[§†] & H. Kishino[§]

[§] *Laboratory of Biometry and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657;* [†] *SCCC, USC, Columbia, SC 29208, USA*

Indels are supplanting sequence substitutions for deep phylogenetic inference, so it is essential that we can accurately model them. To this end, we analyse a ~2mega base pair (bp) region of ten primates in detail. Sequenced BAC clones from ENCODE were retrieved from GenBank and assembled into contigs. A satisfactory alignment required three steps; bulk alignment with MLAGAN, refinement with Muscle, and finally hand editing using BioEdit. Many errors had to be corrected by eye, which is essential for the quality of such studies. "Strict" indels , or those that have identical end points, do not overlap with any other indel, and are e x based pairs from the next indel, worked best for our purpose. Of the ~30,000 indels detected in total, ~19,000 were "strict" (distance to the next indel e 5), providing ample data. For each indel length, distance to the next indel, identical local alignments (called "movability"), location ("coding", intron or intergenic) and number of changes per character on the known species tree were measured. To analyse these we developed an ML regression model based on conditional Poisson distributions. When applied, the only significant variable explaining homoplasy was size, indeed by a size of ~10 bp, homoplasy was almost completely gone. However, when applied to partially overlapping data sets of six species, that removed probable ancestral polymorphism, "movability" also emerged as a significant variable. The residuals from the model for small indels remain large and may follow a wave-like pattern suggesting there is more to understand.

## Mitochondria: Heteroplasmy and Recombination

### Mathematically Modelling of Mitochondrial Mixing

Hendy, M.D.[*] & M.D. Woodhams

*Allan Wilson Centre, Massey University, Palmerston North, New Zealand*

The observation that maternal mitochondrial heteroplasmy is generally conserved across a generation in a number of species indicates that multiple independent copies (segregating units) of the mitochondrial genomes can be inherited. If the number *N* of these segregating units is known, then using a "bean-bag" genetic model among the genomes inherited within a maternal descent line, the mutation rate can be estimated from the proportion of heteroplasmic sites. However the threshold θ of accurately detecting heteroplasmy is an important parameter. We find *N* can be estimated by the variation in level of heteroplasmy change across a generation.

### Blood of the Matriarchs: The descent of mitochondrial heteroplasmy.

M. Woodhams

*Allan Wilson Centre, Massey University, Palmerston North, New Zealand.*

We have estimated the persistence of mitochondrial heteroplasmies (and hence the pedigree evolution rate in Adelie penguins) by measuring the variation in allele frequencies between mother and chick. However, we are measuring the allele frequencies from blood samples, whereas evolution only cares about the germline. I present a model for attempting to account for the variation between germline and blood. Reanalysing the data under this model has a large effect on the estimated evolution rate.

### Good molecules go bad, but does it matter?

Gemmell, N.J.

*School of Biological Sciences, University of Canterbury, PB4800, New Zealand*

Mitochondrial DNA (mtDNA) is the linchpin of modern population and evolutionary genetics. It is widely used to examine the evolutionary history of numerous species and has been employed to determine, for example, the origins and the global expansion of modern humans. The power of mtDNA analyses derives from the apparent simplicity of mitochondrial inheritance (maternal, without recombination), which has enabled models of population history to be much simpler than those needed for the analysis of nuclear DNA. However, in biology things are seldom simple and paternal inheritance of mtDNA, heteroplasmy, and recombination have recently been documented in humans and other species. Here I will present an update of our ongoing experiments to evaluate the extent and frequency of paternal inheritance, heteroplasmy and recombination. New work planned for 2007 to investigate the influence of these phenomena on population and evolutionary interpretations of mtDNA data will also be discussed.

## Graphs, Trees and Distances

### Quality of maximum likelihood estimates

Schliep, K.

*Massey University & Allan Wilson Centre, Palmerston North, New Zealand*

Maximum likelihood is one of the most popular methods to estimate phylogenies since it was introduced by Felsenstein (1981). In statistics exist a big repertoire of tests and methods to evaluate the fit of a maximum likelihood estimators (MLE), many of them are also used in phylogenetic like LR-test, AIC, BIC, bootstrap, etc. I will present some strategies to detect poor fit and multiple optima as described in Steel (1994) or Chor et al. (2000), based on the covariance matrix of the MLE. I will give examples for different biological and simulated datasets. I will also present a fast alternative of fitting (Gamma + I)-like models.

### Realizing Phylogenetic Trees with Weighted Quartets

Grunewald, S.,[§] K.T. Huber,[†] V. Moulton, [†] C. Semple[*‡]

[§] *CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Shanghai, P. R. China;* [†] *School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, UK;* [‡] *Biomathematics Research Centre, Department of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand*

Results that say local information is enough to guarantee global information provide powerful tools in the study of constructing phylogenetic trees and networks. Such results include Buneman's Splits-Equivalence Theorem and the Tree-Metric Theorem. The first of these results says that, for a set Σ of splits, pairwise compatibility is enough to guarantee that Σ is compatible, that is, there is a phylogenetic tree that realizes Σ. The second result says that, for a distance matrix I, if every 4 × 4 distance submatrix of *D* can be realized by an edge-weighted phylogenetic tree, then *D* itself can be realized by an edge-weight phylogenetic tree. In this talk, we give an analogous result for when a set of weighted quartets can be realized by an edge- weighted phylogenetic tree.

## Computing planar split graphs

Bryant, D.,[§] V. Moulton[†] & A. Spillner[*†]

[§] *Dept. Mathematics, University of Auckland, New Zealand;* [†] *School of Computing Sciences, University of East Anglia, United Kingdom*

When genetic data from a set of species has been collected and the geographic locations where the individual samples have been taken are known it is natural to ask about the relationship between the genetic and the geographic landscape in the sampling area. For example, in many situations one might expect a strong correlation between genetic distance and geographic distance between species. In ongoing research we are investigating ways in which to understand the correlation between genetic and geographic distance. As part of this investigation we compute a certain decomposition of the geographic distance in terms of planar splits. This decomposition can be visualized by a planar splits graph which might help indicate regions of poor correlation. In the talk we will give a brief description of our method for computing planar split graphs and discuss some of the algorithmic issues that arise from it.

## Is there a star paradox?

Steel, M.[*] & E. Matsen

*Allan Wilson Centre for Molecular Ecology and Evolution, University of Canterbury, New Zealand*

The 'star paradox' in phylogenetics is the tendency for a particular resolved tree to be sometimes strongly supported even when the data is generated by an unresolved ('star') tree. There have been contrary claims as to whether this phenomenon persists when very long sequences are considered. We settle one aspect of this debate by proving mathematically that there is always a chance that a resolved tree could be strongly supported, even as the length of the sequences becomes very large.

## Perfectly Misleading Distances from Ternary Characters

Fischer, M.[*§] & H.-J. Bandelt[†]

[§] *Biomathematics Research Centre, University of Canterbury, Private Bag 4800, Christchurch, New Zealand;* [†] *Department of Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany*

For the estimation of phylogenetic trees, large DNA data sets are often transformed into distance matrices, to which distance methods are then applied. This transformation, however, inevitably leads to some loss of information, but more seriously, distances can be positively misleading in extreme cases. Huson and Steel (2004) have constructed sequences yielding a unique most parsimonious tree that is totally different from the tree univocally supported by the corresponding distances. Unfortunately, their construction required n-1character states for n taxa and therefore cannot be realized with DNA sequences whenever n > 5. In my talk, I will show that no more than three states are actually needed, that is, binary and ternary characters suffice for generating the extreme contrast between character- and distance-based trees – even if we insist that the distances conform to an ultrametric (i.e. fit a molecular clock).

## Some developments on the consistency of Balanced Minimum Evolution Algorithms

Bordewich, M.,[*§] V. Berry,[†] O. Gascuel,[†] K. Huber,[‡] V. Moulton[‡]

[§] *Durham University, Durham, U.K.;* [†] *LIRMM, Montpellier, France;* [‡] *University of East Anglia, Norwich, U.K.*

We are concerned with distance-based phylogeny reconstruction. Balanced Minimum Evolution (BME) is a method of estimating edge and tree lengths of a given tree topology from a distance matrix. If the distance matrix corresponds to a true tree then the BME principle is consistent, in the sense that the tree topology that minimizes total tree length is the true topology. FASTBME is an algorithmic method for reconstructing a phylogeny given a distance matrix, consisting of two phases. A Neighbour Joining phase, to construct a starting topology, followed by a local topology search, which iteratively checks nearby topologies and moves to the one which minimizes tree length. Typically topologies one Nearest Neighbour Interchange (NNI) away are checked. The topic of the talk is developments in determining whether the algorithm is consistent, that is it is guaranteed to converge to the globally optimal topology.

### The effect of reduced taxon sampling on phylogenetic tree imbalance

Heath, T.A.,[*§] D.J. Zwickl,[§†] J. Kim,[‡] & D.M. Hillis[§]

[§] *Section of Integrative Biology and Center for Computational Biology and Bioinformatics, University of Texas at Austin, Austin, Texas,USA;* [‡] *Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, USA;* [†] *National Evolutionary Synthesis Center, Durham, NorthCarolina, USA*

Patterns of tree shape such as phylogenetic tree imbalance (the degree of asymmetry among lineages) are often used to make inferences about macroevolutionary processes. These patterns may be obscured by non-biological factors that can bias tree shape, such as incomplete taxon sampling, phylogenetic noise, improper rooting, or the methods used to reconstruct the tree. Using simulated trees and published phylogenies inferred from empirical data, we investigated the effect of taxon sampling on tree imbalance. Simulated trees were generated under a branching model, where the speciation and extinction rates were variable and autocorrelated over the tree, and compared to trees simulated under the equal-rates Markov (ERM) model and the proportional-to-distinguishable arrangements (PDA) model. We found that trees simulated under variable rates of speciation and extinction show a pattern of imbalance similar to that of empirical trees, with imbalance increasing as node size increases. Our simulations also show that incomplete taxon sampling causes an increase in tree imbalance in the presence of variable and autocorrelated speciation/extinction rates. In contrast, reducing the number of taxa sampled does not alter the functional relationship between imbalance and node size for trees simulated under the ERM or PDA models. Our results indicate that incomplete taxon sampling in the presence of variable rates of speciation and extinction may be sufficient to explain much of the tree imbalance observed in empirical phylogenies.

## Pacific

### Bayesian coalescent analysis of mtDNA diversity reveals major Southern Asian phase in human prehistory

Atkinson, Q.D., [§] A.J. Drummond[†] & R.D. Gray[*§]

[§] *Psychology Department, University of Auckland, Private Bag 92019, Auckland 1001, New Zealand;* [†] *Computer Science Department, University of Auckland, Private Bag 92019, Auckland 1001, New Zealand*

Human genetic diversity carries a legacy of our population history. Here we apply Bayesian coalescent methods (Drummond et al, 2005) to a global dataset of 361 mtDNA sequences to generate plots of effective population size through time without assuming an a priori growth model and without the problems of information loss associated with distance-based methods. Relative regional effective population sizes for the “present” were found to be good indicators of relative census population sizes across Africa and Eurasia. Plots of effective human population size through time show slow growth in Africa from an MRCA 142-194kya, with a signal of rapid expansion out of Africa into Eurasia 50-70kya. Early growth outside Africa appears to have been centred in South Asia, such that the majority of the world’s population is estimated to have lived in South Asia by ~45kya. A more recent growth phase inferred in Europe beginning ~10-15kya may be associated with the spread of agriculture from ~10kya or an earlier expansion from Southern Europe at the end of the last ice age ~15kya.

### Pacific settlement and Austronesian languages.

Greenhill, S.J.[*] & R.D. Gray

*Department of Psychology, University of Auckland, New Zealand*

The settlement of the Pacific is one of the greatest population movements in the last 10,000 years, and lead to the settlement of the region bounded by Taiwan, Hawaii, Easter Island (Rapanui), New Zealand, and Madagascar. This Austronesian expansion brought with it (and developed along the way) a distinctive Lapita cultural complex and what has become the largest language family in the world, with over 1,000 languages. There are a number of scenarios describing this Austronesian expansion as either a rapid tree-like spread from Taiwan beginning around 6000 BP, or expansion from a deeper Island South-East Asia origin around 13,000 BP. Over the last few years we have built a large comparative database of linguistic information from these languages (http://language.psy.auckland.ac.nz/austronesian) and have begun using phylogenetic methods on it. We will present results from some large analyses of lexical data from over 300 languages, and demonstrate the power of this data at resolving these questions.

### Spatial and temporal distribution of *Rattus exulans* in Near Oceania

Hingston, M.[*§], H.A. Ross,[§] E. Matisoo-Smith[†] & J.H. Robins[†]

[§] *School of Biological Sciences and* [†] *Department of Anthropology, University of Auckland, Auckland, New Zealand*

People of the Lapita cultural complex colonised Remote Oceania via Near Oceania around 3500bp, bringing with them Kiore, the commensal Pacific rat (*Rattus exulans*). By using these rats as a proxy for humans, we aim to observe the phylogenetic relationship among Near and Remote Oceanic populations of *R. exulans* in order to distinguish between the hypotheses of single or multiple introductions and thus untangle the sets of models of Lapita origins. We are extending previous work with more focused sampling of rats in the vicinity of Papua New Guinea, and by applying Bayesian inference to the estimation of phylogenetic relationships and migration rates among populations of *R. exulans*. We will report on our recent findings.

### The Evolution of Social Stratification in the Pacific: A Phylogenetic Approach

Currie, T.

*Department of Anthropology, University College London, London, UK*

Phylogenetic comparative methods developed in the Biological Sciences can be applied to answer questions about human cultural evolution. Austronesian-speaking societies of the Pacific exhibited a great deal of variation in the extent to which some individuals were able to exercise power over others. Here I present an investigation into the evolution of this social stratification using phylogenetic methods. Linguistic data from the Austronesian Basic Vocabulary Database was used to construct a Maximum Parsimony phylogenetic tree of the societies under investigation. Social stratification is shown to be co-evolving with the presence irrigated crops. However the direction of evolution is unable to be assessed with the current data. Limitations of the current investigation are discussed and future directions for research are outlined.

## Organisms

### Marsupial “herbivore” phylogeny and the failure of non-evolutionary morphological analysis

Phillips, M. J.

*Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand*

The Australasian marsupial order Diprotodontia includes ten extant families, which together comprise the most ecologically diverse assemblage of terrestrial species of any mammalian order. In the study presented, separately modeled process partitions for a mitochondrial/nuclear sequence supermatrix provide near-complete resolution of diprotodontian family-level phylogeny. In order to gauge the similarity of the resultant phylogeny with past studies that used different data-types or methodological approaches, MRP supertrees were constructed to summarize diprotodontian phylogenetic inference from (a) DNA-DNA hybridization, (b) albumin micro-complement fixation, (c) mitochondrial and (d) nuclear sequences, as well as morphological data under (e) evolutionary character analysis and (f) cladistic analysis. On the supermatrix data, the favoured tree and MRP trees (a) to (e) are similar to each other relative to the MRP tree for morphological caldistic analyses, which is an extreme outlier. Relationships in the morphological cladistic

analysis MRP tree appear to be correlated with size (allometric scaling) and ecology. Bayesian inference reanalysis of the data matrices from which these trees derive fail to correct the putatively erroneous relationships.

### Meme analysis of bird songs

Ranjard, L.* & H. Ross

*Bioinformatics Institute and School of Biological Sciences, University of Auckland, Auckland, New Zealand*

The songs of passeriformes bird species are learned rather than innate. The learning occurs when the bird is immature and in some cases learning continues throughout a bird's life. This learning process is divided into two steps; firstly a bird hears and memorizes a song from another bird, and secondly it tries to imitate the song as accurately as possible. This imitation is not perfect, therefore songs evolve through generations as small changes occur. Thus, it appears that in certain cases, bird song is a culturally transmitted trait and therefore these songs can be considered as a collection of memes. Using the framework of meme evolution, an analogy of molecular evolution, song evolution involves deletion, substitution and repetition of syllables. In studying birdsong evolution, the first critical step is to convert a continuous song into a sequence of discrete syllables or notes, which can then be classified as memes. First, we'll present a bio-acoustic method using neural networks which have been developed for encoding bird songs as sequences of discrete syllables. A pairwise syllable distance measure has been used in order to train a self-organising map from which syllable types, i.e. memes, can be defined. Then, songs are binary encoded, each specific meme can feature in a song or not, in order to build phylogenetic trees. It is therefore possible to analyse how those memes are involved in song evolution and if their distribution reflects population history.

### The Jekyll and Hyde lifestyle of bacteria revealed by genome phylogenies

Darling, A.E.,* & M.A. Ragan

*Institute for Molecular Bioscience, Brisbane, Australia*

Whole genome sequence data offers the tantalizing possibility of using phylogenetic inference to discover historic changes in microbial lifestyles. At the genome scale, mutations such as rearrangement, gene gain, gene loss, and the proliferation of mobile DNA can be combined with the traditional repertoire of molecular phylogenetic characters to provide a comprehensive view of genome evolution. Development of a comprehensive population-genetic model of genome evolution stands to resolve several ongoing controversies such as the minimum rate of LGT and the nature of speciation. Using the finished genome sequences of nine *E. coli* and six *Shigella*, I will describe how genome phylogenies can be used to infer ecological niche adaptation in microbes. In general, *E. coli* is a non-pathogenic, commensal organism that forms a normal part of the human gut microflora. *Shigella*, on the other hand, is a diarrhea-causing pathogen that has recently evolved as a subpopulation within the population of *E. coli*. During its transition to a pathogenic lifestyle, *Shigella* has undergone major rearrangement and reformulation of its genome that I propose as characteristic of niche-ification. Using genome rearrangement phylogenies it may be possible to detect the signature of historic transitions to/from pathogenic lifestyles as bursts of rearrangement activity associated with niche change.

## Molecular and Genome Evolution

### Accounting for Exposition and Secondary Structure in Protein Evolution: Models and Gain

Le, S.Q. & O. Gascuel*

*LIRMM-CNRS, Montpellier, France*

It has been recognized for a long time that substitution processes vary depending on structural configurations. However, this information is not (or rarely) used in phylogenetic studies, even though the structure of dozen thousands of proteins has been elucidated. Here we reinvestigate the question in order to fill this gap. We used a very large dataset comprising 4,389 protein alignments with structural annotations to estimate new amino-acid substitution matrices for various structural configurations. Moreover, we used an independent sample of 500 alignments to evaluate the gain in tree likelihood brought by these new matrices. Various ways to combine these models (matrices) were envisaged, namely, separate analysis based on available annotations, mixtures (assuming no structural information), and a combination of both based on an estimated parameter that reflects the reliability of structural annotations. Our results show that separate analysis and mixtures are nearly equivalent in average, while our confidence-based approach is best thanks to its ability

to detect poorly annotated proteins. Highest likelihood values are obtained with six structural categories combining exposed/buried and alpha/beta/other status of the sites; the average gain is as high as 1.16 AIC points per site, compared to standard WAG model. This six-category model is closely followed by the two-category exposed/buried model, while the secondary structure-based three-category model is worse, but still better than WAG. All these likelihood gains induce significant topological changes in the trees being inferred, indicating that our models should be used routinely by phylogeneticists. Datasets, substitution matrices and a PHYML-like implementation will be soon downloadable from our URL (http://www.lirmm.fr/mab).

## Disentangling selection and mutation in comparative genomics: the greater-than-equal and opposite 'rule'

Roy, S.W.

*Allan Wilson Centre for Molecular Ecology and Evolution, Palmerton North, New Zealand*

A central question of evolutionary biology concerns the relative influence of mutation and selection in driving genetic differences between species. However, in the absence of relevant data on standing polymorphism, or in the absence of significant standing polymorphism, it is difficult to attribute observed genetic differences between lineages to mutational or selective differences. Here I consider the general case of a class of reversible mutations in which one of two states is selectively favoured over the other. I derive a relationship between the expected rates of fixation of favoured and disfavoured mutations. In general for any change in Nes, the rate of fixation of disfavoured mutations is expected to change by a factor greater than and in an opposite direction to) the factor by which the rate of fixation of favoured mutations is expected to change. I call this the 'greater-than-equal and opposite rule': for every response of the rate of fixation of favoured mutations to selection, we expect a more-than-equal and opposite response in the rate of fixation of disfavoured mutations. I discuss this insight in light of the example of codon bias and intron loss/gain in *Drosophila* species.

## Detecting Asymmetric Markov Processes in Aligned Sequence Data

Jermiin, L.S.,[*§†] F. Ababneh,[‡] C. Ma[£] & J. Robinson[‡]

[§] *School of Biological Sciences, University of Sydney, NSW 2006, Australia;* [†] *Sydney Bioinformatics, University of Sydney, NSW 2006, Australia;* [‡] *School of Mathematics & Statistics, University of Sydney, NSW 2006, Australia;* [£] *Department of Mathematics and Statistics, Wichita State University, Kansas 67260-0033, USA*

Most phylogenetic methods assume that the sequences of nucleotides or amino acids have evolved under stationary, reversible, and homogeneous conditions. When these assumptions are violated by the data, as would be the case if there is compositional heterogeneity across the sequences, the phylogenetic estimates are obtained under an incorrect model and thus subject to error. Methods to examine aligned sequences for violation of these assumptions have been available for years, but they are rarely used, presumably because they are not widely known or because they are poorly understood. Here we describe and compare matched-pairs tests for symmetry of two-dimensional contingency tables from homologous sequences and show that the tests of symmetry, marginal symmetry and internal symmetry can be used not only to detect violation of the assumption of stationarity, reversibility and homogeneity, but also to identify what may have caused this violation. The tests are unaffected by invariant sites and divergence between the pairs of sequences, implying that they may be used to identify suitable substitution models for estimation of evolutionary relationships under a Markovian model.

## A maximum-likelihood framework for gene expression evolution models with mutational and non-mutational effects

Rosskopf, M.,[*§] & A. von Haeseler[†]

[§] *Dept. of Computer Science, University of Düsseldorf, Germany;* [†] *Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, Dr. Bohr-Gasse 9, A-1030 Vienna, Austria, University of Vienna, Medical University of Vienna, and University of Veterinary Medicine Vienna*

Several studies describe that the level of gene expression between species is positively correlated with the time that has passed since the species split from a common ancestor (Ranz 2006 TREE). Moreover, Khaitovich et al. (2004 PLoS

Biology) found a linear relationship between divergence time and expression differences. This linearity can be explained by the neutral theory (Kimura 1983). To model the process of gene expression according, the mutation rate is described by a Poisson process. The strength of changes in expression level is described by a continuous distribution, the so-called mutation effect distribution. That is, whenever a mutation occurs, the gene expression level changes according to the mutation effect distribution. However, the strength of gene expression is also affected by metabolism and epistatic effects which overlay mutational changes of gene expression. To take these effects into account we introduce an additional parameter that summarizes all non-mutational changes. In order to estimate the model parameters, we suggest a maximum-likelihood framework, which is applicable to any type of mutation effect distribution. However, arbitrary mutation effects can increase the computational complexity rapidly. Thus, we suggest a model with normal distributed mutational and non-mutational effects and present an optimisation method for parameter estimation. In addition, we developed a method to determine if a single gene changed its expression by mutations or by non-mutational effects. Both methods were evaluated with synthetic data. Furthermore, biological data from different human and chimpanzee tissues were studied. Our results indicate a larger number of mutations in testis than in the remaining tissues brain, heart, kidney, and liver. The largest component of non-mutational effects was estimated in heart.

## Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution

Dagan, T. & W. Martin[*]

*Institut fuer Botanik III, Heinrich-Heine Universitaet Duesseldorf*

The amount of lateral gene transfer (LGT) that has occurred in microbial evolution is heavily debated. Efforts to quantify LGT through gene tree comparisons have delivered estimates that anywhere between 2% and 60% of all prokaryotic genes have been affected by LGT, the 30-fold discrepancy reflecting differences among gene samples studied and uncertainties inherent to phylogenetic reconstruction. Here we present a simple, but novel, method that is independent of gene tree comparisons to estimate the LGT rate among sequenced prokaryotic genomes. If little or no LGT has occurred during evolution, ancestral genome sizes would become unrealistically large, while too much LGT would render them far too small; we determine the amount of LGT that is necessary and sufficient to bring the distribution of inferred ancestral genome sizes into agreement with that observed among modern microbes. Rather than testing for phylogenetic congruence or lack thereof across genes, we assume that all gene trees are compatible, hence our method delivers very conservative lower-bound estimates of the average LGT rate. The results indicate that among 57,670 gene families distributed across 190 sequenced genomes, at least two-thirds, and probably all, have been affected by LGT at some time in their evolutionary past. A component of common ancestry nonetheless remains detectable in gene distribution patterns. We estimate the minimum lower bound for the average LGT rate across all genes as 1.1 LGT events per gene family and gene family lifespan and this minimum rate increases sharply when genes present in only a few genomes are excluded from the analysis.

## Modeling context-dependent evolution: the influence of the immediate flanking bases

Baele G.,[*§†] Y. Van de Peer[†] & S. Vansteelandt[§]

[§] *Department of Applied Mathematics and Computer Science, Ghent University, Belgium;* [†] *Department of Plant Systems Biology, Ghent University / VIB, Belgium*

A number of approaches in the past have already relaxed the assumption that sites evolve independently by modeling for example base-pair evolution. In our work we assume that the evolution of a site is dependent on its two immediate neighboring sites, also known as the site's "context of evolution". To efficiently evaluate the corresponding models, we employ a "data augmentation" approach in an MCMC framework. Depending on the identities of its neighbors, a site is assigned a branch-specific evolutionary model as the identities of the neighbors are not necessarily identical for all the branches of the tree, i.e. not only the site itself but also its neighbors may have a different state in different internal nodes of the tree. Taking into account every possible context of evolution will lead to a drastic increase in parameters, which not necessarily leads to an important increase in the goodness-of-fit of the model. Using thermodynamic integration to calculate Bayes factors, we illustrate ways to handle the trade-off between added number of parameters and increased model fit.

## Is Microevolution sufficient for Macroevolution; birds and mammals.

David Penny (and gang)

*Allan Wilson Centre, Massey University*

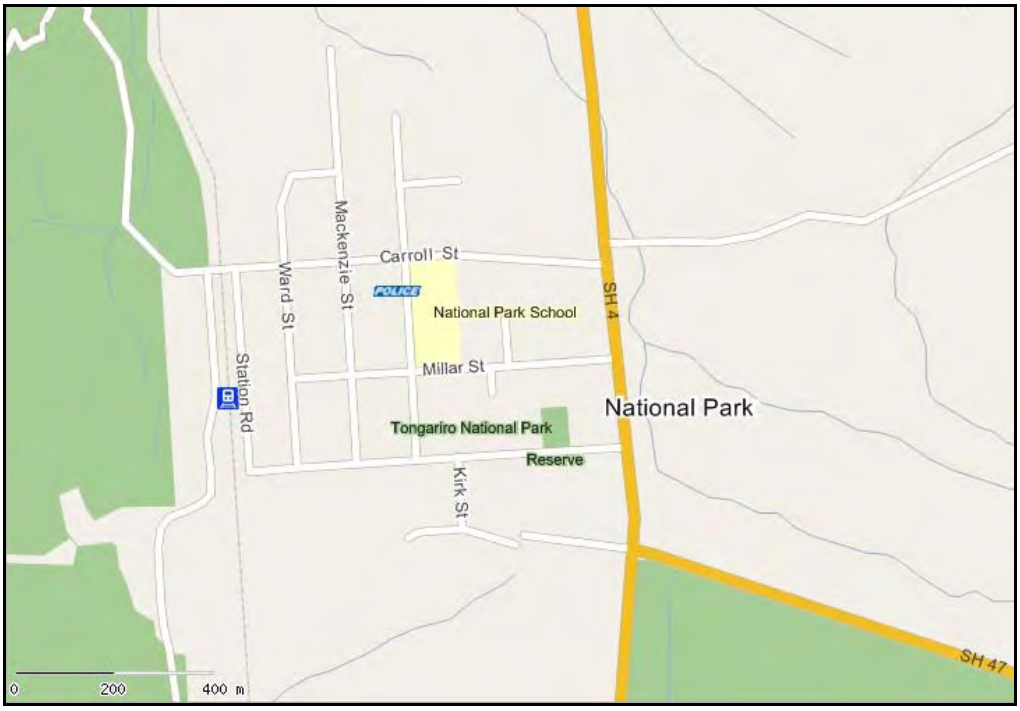
The question whether the processes of microevolution are sufficient for macroevolution is conceptually difficult to test. Much of it is inferring processes that may, or may not have, occurred tens of millions of years ago. For birds and mammals, a popular viewpoint has been that it was necessary for some external agent, in this case an asteroid on a collision course with earth, had to wipe out dinosaurs and pterosaurs before the early birds and mammals could radiate into the vacant niches. This implies both a "sudden and unexpected extinction, at the Cretaceous/Tertiary boundary". All three aspects ('sudden', 'unexpected' and 'at the K/T boundary') are amenable to testing. By combining molecular phylogenies and fossil calibration points there is some quite rapid progress that will be discussed. For both birds and mammals, the latest results imply modern bids and mammals are diversifying from around 100Ma (million years ago). These will be discussed; maybe the good guys are winning!

# Maps

## Regional map

## Whakapapa map


## National Park map



## Activities in and around Tongariro National Park

*Hiking - Please see reception regarding walks – min 20, max 7 hrs*

*Mt Ruapehu Guided Crater Hike - Tel : 07 8923 738*

*Mountain Biking - Tongaririo Mountain Bike, Tours - Tel 0800 10 1024 Also*

*Bike hire is available from Howard's Lodge - Tel : 07 8922 827*

*National Park Climbing Wall, Indoor climbing hall - Tel : 08 8922 870*

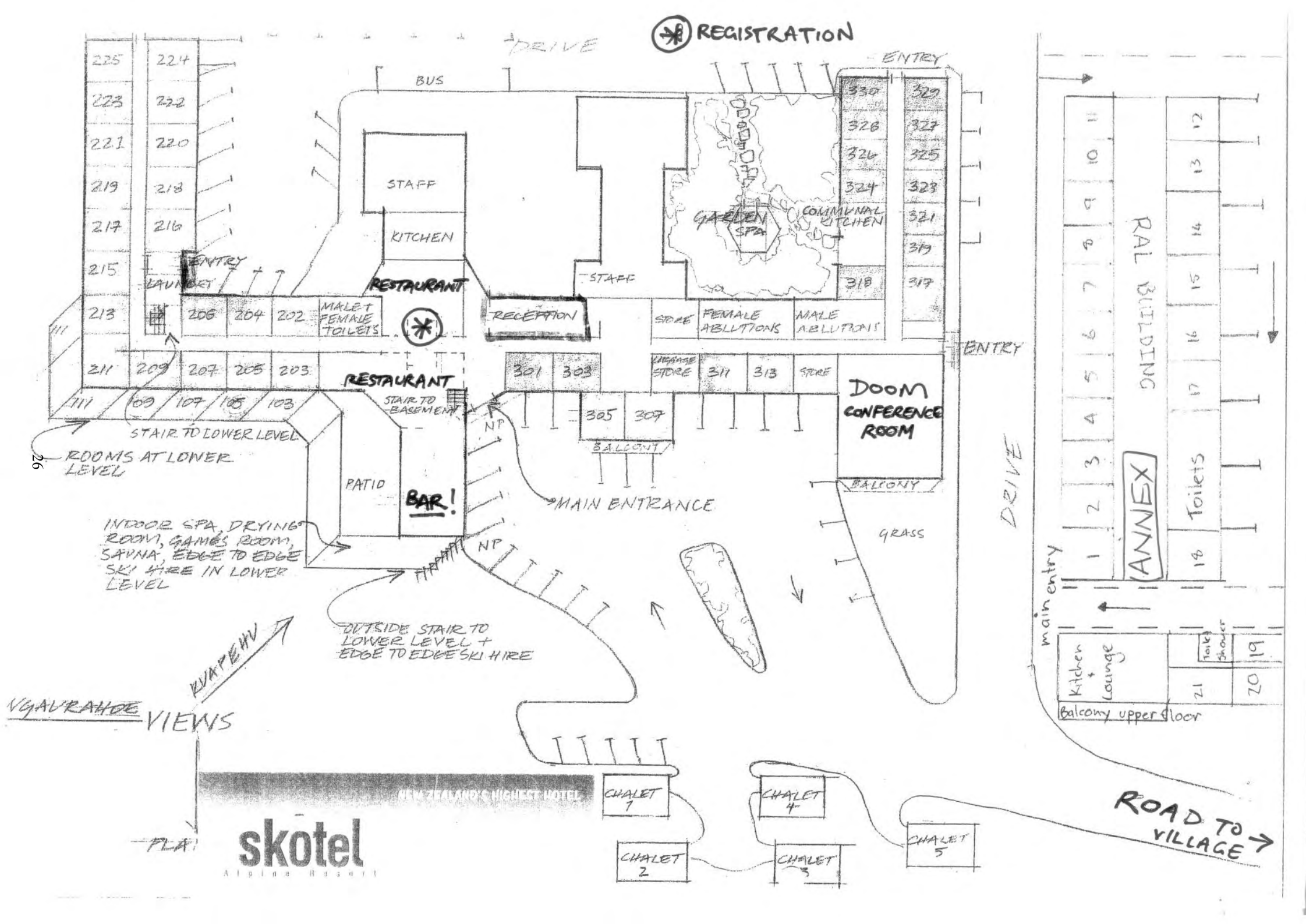
*Horse Back Riding / Quad Biking - Tel : 0800 628 642*

*Scenic Flight, Mountain Air - Tel : 0800922 812*

*Adventure Kayak Tours - Tel : 07 8922 778 / 027 242 7209*

*Whanganui River Adventures & Jet Boat Tours to Bridge to Nowhere - Tel : 0800 862 743*

*Golf - Clubs available for hire at Grand Chateau*


REGISTRATION
DRIVE
ENTRY
BUS
STAFF
KITCHEN
RESTAURANT
ENTRY
LAUNDRET
STAFF
GARDEN SPA
COMMUNAL KITCHEN
RECEPTION
STORE
FEMALE ABLUTIONS
MALE ABLUTIONS
MALET FEMALE TOILETS
RESTAURANT
STAIR TO BASEMENT
STAIR TO LOWER LEVEL
ROOMS AT LOWER LEVEL
PATIO
BAR!
INDOOR SPA, DRYING ROOM, GAMES ROOM, SAUNA, EDGE TO EDGE SKI HIRE IN LOWER LEVEL
MAIN ENTRANCE
BALCONY
DOOM CONFERENCE ROOM
BALCONY
GRASS
DRIVE
ENTRY
OUTSIDE STAIR TO LOWER LEVEL + EDGE TO EDGE SKI HIRE
KVAPEHV
VGALVRAHDE VIEWS
FLAT
CHALET 1
CHALET 2
CHALET 3
CHALET 4
CHALET 5
ROAD TO VILLAGE
RAL BUILDING
ANNEX
Toilets
main entry
Kitchen + Lounge
Balcony upper floor
NEW ZEALAND'S HIGHEST HOTEL
skotel
Alpine Resort

Nestled amongst

heather and Mountain

Beech you will find a

beautiful alpine retreat

**mind**

far from the madding crowd.

The Skotel lies in the heart of

the Tongariro National Park,

in Whakapapa village at the

**body**

base of Mt Ruapehu - 1142m

above sea level. Up here the

air is clear, the water clean

and the views spectacular.

**spirit**

In summer we are the best

base for the Tongariro Crossing

and other walks.

For skiiing and snowboarding

we are your winter ideal.

skotel
Alpine Resort

menu being updated
(bigger + better)

## ENTRÉE

**Soup of the Day - $7.50**
Home-made with warm bread

**Bread Selection - $7.00**
Selection of Breads with three dipping sauces
Olive oil, green pesto and red pesto

**Crumbed Mussels - $13.50**
On a bed of salad

**Caesar Salad - $10.50 entrée - $14.00 main**
Classic salad served with bacon and croutons

**Fettuccine Cabonara - $9.00 entrée $15.50 main**
Bacon and mushroom cream sauce with Parmesan

## MAINS

**Lamb Curry - $19.50**
With basmati rice, poppadoms, cucumber raita and fruit chutney

**Chicken Breast - $22.50**
On tomato rice with mushrooms in a Pernod cream

**Kiwi Steak – $22.00**
Steak served with fries and salad garnish with choice of three sauces.
Mushroom, pepper or garlic butter

**Fish of the Day - $25.50**
Served on a warm potato salad, Mediterranean leaves with chilli lime hollandaise

**Spanish Vegetable Pasta - $18.50**
Sundried tomatoes, capsicum & black olives in provencal sauce with parmesan cheese

**Venison Denver Leg - $26.50**
Oven baked medium rare served atop garlic tossed vegetables finished with blueberries and red wine demi

**Roasted Lamb Rump - $26.50**
Served atop a pumpkin and herb rosti with grilled bell pepper and red wine jus

## Desserts – All $9.50

***Ice Cream Sundae***
Chocolate, Berry or Caramel sauce
with Ice cream and cream

***Vodka & Lime Cheesecake***
Served with fresh cream and Ice cream

***Black forest Gateau***
Served with fresh cream and Ice cream

***Sticky Date Pudding***
Drizzled with caramel sauce
served with Ice cream and cream



Our hotel has

a wide variety of

rooms to suit all

tastes and budgets.

**best**

Shoulder season prices

and great package deals

give you the chance

to take a holiday which

**park**

is economical and yet

brings you the comfort

of a hotel with all the

usual facilities.

**value**

Above all, our staff

are determined to

make you really welcome.

Skotel

# BREAKFAST MENU

*(All prices quoted include G.S.T.)*

**With any of our breakfasts please help yourself to tea or coffee from the buffet.**

## LIGHT BREAKFAST

| | Adult | Child |
|---|---|---|
| **Toast**<br>With tea & coffee. | **$7.00** | **$5.00** |
| **Continental**<br>Select from the buffet: cereals, preserved and fresh fruit, yoghurt, meats and cheese. | **$16.50** | **$9.50** |

## COOKED BREAKFAST

| | | |
|---|---|---|
| **Toasted Muffin with Poached Egg and Bacon** | **$8.50** | |
| **Hot Porridge**<br>With cream & brown sugar. | **$8.50** | **$5.50** |
| **Baked Beans or Spaghetti on Toast** | **$9.50** | **$5.50** |
| **Eggs on Toast**<br>2 eggs any style on toast. | **$9.50** | **$5.50** |
| **Blueberry Pancakes** | | |
| With maple syrup & whipped cream. | **$13.50** | **$7.50** |
| With bacon, banana & maple syrup. | **$15.50** | **$9.50** |
| **Eggs Benedict**<br>Toasted muffin with smoked salmon, poached eggs & hollandaise sauce. | **$15.50** | |
| **Kiwi Breakfast**<br>Bacon, hash browns & 2 eggs (any style). | **$15.50** | **$9.50** |
| **Ngauruhoe Breakfast**<br>Grilled bacon, sausage, mushrooms, tomatoes, hash brown & two eggs (any style). | **$18.50** | |

**Extras to any Cooked Breakfast.**

| | | | |
|---|---|---|---|
| Bacon(2 slices) | **$5.00** | Mushrooms | **$3.50** |
| Sausages | **$3.00** | Hash Browns | **$3.00** |
| Baked Beans & Spaghetti | **$3.00** | | |

**Special Coffees** *(please order from the waitress)* **$3.50**

*The children's menu is available for up to and including 12 year olds.*
*Children under 5 years of age may have a continental breakfast free of charge.*

***Please inform your waitperson if you have any allergies or dietary requirements.***

11.30am → 8.45pm (last orders for food)

BAYVIEW
CHATEAU TONGARIRO
MOUNT RUAPEHU • NEW ZEALAND

W. Village

# Pihanga Menu

| | |
|---|---|
| Soup of the day | $8 |
| Breads, hummus & kalamata olives | $9 |
| Garlic bread | $4 |
| Seafood chowder with toasted baguette | $14 |
| BLT on Turkish Pide with basil mayonnaise & fries | $15 |
| Smoked Chicken & Bacon Salad, mango salsa & crispy noodles | $15 |
| Classic Shrimp and Avocado Salad | $16 |
| The Chateau's Creamy Beef & Mushroom pie | $14 |
| Premium Beef Burger with gruyere cheese, pickle, lettuce & spiced mayo served with fries | $15 |
| Quattro Formaggio with Penne | $17 |
| Smoked Venison "Sausage Dog", onion jam & Dijon mustard | $17 |
| Salmon Fillet on Nicoise Salad with basil dressing | $21 |
| Sirloin Steak with fries, béarnaise sauce and rocket salad | $22 |
| Butter Chicken on steamed rice with poppadom & chutney | $18 |
| Chermoula spiced Chicken Breast on warm Mediterranean Salad | $21 |

Sides

| | |
|---|---|
| Steamed vegetables | $5 |
| Fries with garlic mayo | $5 |
| Side salad | $5 |

# Dessert

| | |
|---|---|
| Ice Cream Sundae with Chocolate, Caramel or Strawberry Sauce | $8 |
| Cheesecake of the Day | $9 |
| Plum and Port Pudding with Marscapone | $9 |
| Fresh Fruit Salad with choice of Yoghurt or Vanilla Ice Cream | $9 |

# Hot Beverage's

| | |
|---|---|
| Cappucinno, Latte, Flat White, Long Black, Espresso and Mocca | $4 |
| English Breakfast or Earl Grey Tea | $3 |
| Hot Chocolate | $4 |
| Chamomile, Peppermint, Apple and Blackcurrant, Lemon Twist or Green Tea | $3 |
| Chai Latte | $5 |

NATIONAL HOTEL PARK
THE 'ORIGINAL' INN AT THE PARK

Nat. Park Pub

YOUR HOSTS LYNDA & GLENN

## ✳ STRAY BUS OFFER OF SUSTINANCE.....

Homemade *soup of the day*, served with buttered toast $5.50

*Potato wedges*, topped with crispy bacon, cheese, sour cream & sweet chili sauce $7.90

N.B. these prices are for ✳ a special deal; food is a little more expensive than indicated on individual basis

*Cheeseburger* with fries $8.50

*Vegetarian* Burger and fries $9.50

*Fish* burger and fries $10.50

*Steak* burger with fries $11.50

*Crispy crumbed schnitzel* served with fries and garden salad $13.50

*Fish and chip* meal with fresh garden salad $11.50

*Beef Lasagne* served with salad and fries $12.50

*Chicken kiev:* chicken breast filled with garlic butter, served with golden potato croquettes. $14.50

*Rump steak* served with chips and eggs or salad and chips $16.50

*Hot sticky date pudding* with ice cream $9.50

*Chocolate overload*: iced chocolate cake and chocolate ice cream, smothered in chocolate sauce $10.50

*Ice cream Sunday*, with your choice of topping $7.50

Open 5pm – 8.30pm for meals

Gibson Enterprises Limited operating as NATIONAL PARK HOTEL & BACKPACKERS
Phone (07) 892 2805 • Fax (07) 892 2746 • Email: nat.park.hotel@xtra.co.nz
61 Carroll Street, PO Box 60, National Park Village 2653

# Eivins Off Piste Bistro Bar Jan 2007

Menu is
current at time of reporting and can change

## SUMMER MENU REPORT

Nat. Park

### ALPINE AMBIENCE

| FOR LUNCH | | EVENING | |
|---|---|---|---|
| **Weather Permitting** | Balcony dining | **Temperature** | Warm and Cosy |
| **Wind** | Just a flicker of the tussocks | **Wind** | Just a flicker of the candles |
| **Visibility** | Panoramic Mountain Views | **Visibility** | Variable (Don't forget your camera) Wicked sunsets! |

### FACILITIES OPERATING

| | |
|---|---|
| **Eivins Bistro & Bar** | ***OPEN from 12pm until late*** * |
| **Ski and Board Rental** | ***CLOSED inquiries welcome*** |

### Evening Apres Tracks Menu

| | | |
|---|---|---|
| Soup | Chilled Watermelon Gazpacho $ 6 Mushroom and port | $ 8 |
| Bread | Garlic Ciabatta Bread $ 6 Bread and dips | |
| Cheese | Puhoi Brie Cheese Filo wrapped oven baked served w/ spicy plum sauce | $7 |
| | Crumbed camembert, garnish salad and plum sauce | |
| Seafood | Calamari sweet chili sautéed, tender and tasty served w/ ginger rice | $12 |
| | Warm Malibu prawns cutlets and a garnish salad | $ 8 |
| Potatoes | | |
| | Bowl of potato fries choice of garlic mayonnaise or tomato sauce | $ 5 |
| | Bowl of kumera (sweet potato) fries served w/ orange mayonnaise | $ 6 |
| Salads | | |
| | Side salad, fresh salad greens topped w/ honey-mustard vinaigrette | $ 6 |

Beef

*Scotch fillet, Salad and potato Fries, topped w/ Dijon mustard $26.50

*Drunken Scotch, stuffed with whisky marinated banana cajun prawns, served w/ potato mash and a baked tomato $26.50

Pork

Spare Ribs baked in our BBQ Sauce served w/ potato fries and a side salad $25

Prawns

Laksa, asian slaw, cashew nuts nestled on noodles topped w/ thai curry sauce $20

Chicken

*Breast studded w/ cream cheese, sun dried tomato and coriander pesto, bound in filo pastry, oven baked, served w/ baked potatoes and Julienne vegetables, mango chutney flavored $26

*Manuka smoked, tossed in a putanesca sauce w/ fusilli pasta $18

*Vegitarian option with out chicken extra SEASONAL vegetables added $16

Lamb

Rack, basil pesto encrusted served on baked potato, and a savory tomato mint sauce. $26

**Junior Menu**

Creamy bacon & tomato Pasta $8 Steak & Fries $10 Chicken Apricot Skewers $ 8

Cheese Toasty $3

Vegetarian and special dietary needs happily catered for. Please discuss with attendant

**Reservations ph 07 8922844 0272507808**

* normally till 8.30pm, but will take people later if booked

# SCHNAPPS BAR

Just off SH4 Nat. Park
12pm – 9pm (or later)

## Main Meals

**Lamb Shanks** 19.50
Braised in Mint jus served on Potato & Kumara Mash

**Vegetable Stack Grilled** 19.50
Layered Red & Yellow Peppers, Eggplant, kumara with
Crispy Black Fettuccini & Hollandaise Sauce

**Scotch Fillet** 23.50
Served with Rosemary & Garlic Baby Potatoes with a Creamy Garlic Sauce

Why Not Surf & Turf Add 4.50

**400g T – Bone Steak** 25.00
With chips and a choice of Pepper or Mushroom Sauce

Top it up With Homemade Herb Crumbed Onion Rings Add 2.50

**Razor Back Ribs** 18.50
Razor Back Ribs glazed in a Smokey BBQ Sauce & Fries

**Kiwi Favourite Hapuka** 19.50
Pan Fried with Basil & Pesto Potatoes topped with Spinach and
Lemon Pepper Hollandaise

**Venison Pie** 22.00
A Rich Juniper Berry Hot Pot, topped with a Flaky Pastry Top

**Horopito Chicken Breast** 23.50
Oven baked in Horopito Spices, Char Grilled Peppers,
Tomato & Garlic with Gourmet Baby Potatoes

## Desserts All 9.00

**Apple Crumble**
Homemade Apple & Berry Crumble served with Ice Cream & Fresh Fruit

**Ice- Cream Sunda**
R18 Ice-cream Sundae with a Toffee & Baileys Sauce

**Cheesecake**
Rich White & Dark Marbled Chocolate Baked Cheesecake

**Brulee**
Traditional Brulee served with Ice cream & Cream

# SCHNAPPS BAR

## Light Meals

**Garlic Bread** 6.50
Oven Baked Crusty Loaf with Garlic Butter

**Today's Soup** 9.00
Made Fresh with Local Produce

**Tandoori Lamb Cutlets** 14.50
New Zealand Quality Lamb Cutlets Marinated in Tandoori Spices served on Rice drizzled with Fresh Mint Yoghurt

**Buffalo Wings** 12.50
Spicy Chicken Wings accompanied by Blue Cheese Dip

**Caesar Salad** (Vego Optional) 13.50
Crispy Cos Leaves with Bacon, Croutons, tossed in a Traditional Dressing

With Grilled Chicken Add 15.50
With BBQ King Prawn Add 18.50

**Thai Beef Salad** 15.50
Marinated Strips of Prime Beef, Thai Style with Crispy Noodles & Summer Salad

**New Zealand Mussels** 13.50
Steamed in either a Light Chilli Tomato or Creamy Garlic Sauce

**Grilled Scallops** 12.50
On a Bed of Sweet Crispy Spinach

- Closed Mon/Tue nights
- bookings essential

- avoid 1-2pm! (when train arrives)

# THE STATION

## • CAFE • BAR • RESTAURANT

Nat. Park

The Station is a fully licensed cafe/bar/restaurant situated within the historic National Park Railway Station. An ideal venue for travellers passing through or those staying over. The Station offers a superb atmosphere for daytime and evening dining. Carleen and Warren extend a warm welcome inviting you to come on in to The Station and experience our fine food and friendly service.

Reservations recommended for evening dining.

# Summer Menu

## Entrees

**Soup of the day** $8.00

**Antipasti Platter** $16.50
A selection of delicious tastes to share.

**Focaccia** $6.50

**Prawn Kebabs** $14.00
Cooked in sweet chilli and lime.

**Scallops** $15.50
In a pastry case with pink peppercorn sauce.

**Homestyle Pâté** $12.50
Served with melba toast and port wine jelly.

## Childrens Menu

**Selection of children's dishes available.**

## Mains

**Caramelised Orange Duck** $29.00
Served with wild rice.

**Lamb Rack** $28.50
Filled with pistachio, apricot and fresh mint.

**Salmon Buerre Blanc** $29.50

**Cajun Rubbed Beef Fillet** $29.00
Drapped with pesto hollandaise.

**Red Risotto** $22.50
Creamy rice, beetroot, feta, olive and walnut.

**Pork and Watercress** $28.00
Served with a warm blue cheese sauce.

**Sirloin Steak** $23.50
Cooked to your liking with garlic or mushroom sauce.

**Chicken Breast** $25.00
Coated with chilli, lime and cashew glaze.

## Dessert

**Passionfruit Brûlée** $9.50
Served with sweet tamarillo.

**Sweet Velvet Chocolate Parfait** $9.50
With espresso jelly.

**Pear Tart** $9.50
With blue cheese and hazelnuts.

**Sticky Date Pudding** $9.50
With caramel and cream sauce.

**Sorbet Duet** $9.50

**Cheeseboard Selection** $16.00

## Beverages

**A fine selection of beer, wine, spirits and non-alcoholic beverages.**

**SUNDAY SPECIAL** - Traditional Roast Dinner $16.50 (kids half price)