Conservation genetics

The Institute's researchers develop methods and tools in conservation genetics, and apply these to inform the conservation of threatened species.

Genetic diversity influences the health and long-term survival of populations, with decreased genetic diversity associated with reduced fitness and adaptability to changing environments. Conservation genetics provides the theory and methods to estimate historical and current genetic diversity and to predict the future distribution of genetic diversity in populations and species. This crucial information informs conservation strategy in order to minimise risk of extinction in threatened species.

Estimating effective population size

Threatened species usually have small and declining population sizes, and genetic diversity is lost due to genetic stochasticity. The rate of loss of genetic diversity is determined by the effective population size (N_e) of the species, which is affected by factors including the total population size (N), variance in reproduction and sex ratio. For wild populations, N_e can be dramatically smaller than N, especially for species with extremely unbalanced sex ratio and high reproductive dominance.

Therefore, it is essential to measure the N_e of a population for conservation management. We have developed methods to estimate N_e from genetic marker data (most recently Wang 2016a), and produced user-friendly computer software called Colony, which is available on the Institute website for use by conservation genetics practitioners. Thousands of downloads of the Colony software have been recorded over the past 12 months and new Institute research used these methods to investigate the historical and current effective population size of Hainan gibbons (*Nomascus hainanus*) from microsatellite data (Bryant et al. 2016).

Inferring population structure

Individuals of a species may not constitute a genetically homogeneous population. Instead, they are usually clustered into different populations due to patchy geographic distribution and a limited ability to disperse. Inferring population structure by delineating the number and distribution of populations and by measuring the genetic differentiation among populations helps us to understand the current distribution of genetic diversity.

Institute research has shown that the most appropriate statistic for quantifying genetic differentiation is Fst (a measure of genetic differentiation between populations). However, Fst can underestimate differentiation when calculated from highly polymorphic microsatellite data. We have shown for the first time the circumstances when microsatellites give locus-dependent and underestimates of Fst, such as high mutation rate and low migration rate. We have since developed a correlation analysis to detect such circumstances from genotype data (Wang 2015).

We have also used a Bayesian clustering method (Structure) to make individual assignments to populations from multi-locus genotype data, most recently to investigate the population structures of grey squirrels (*Sciurus carolinensis*) (Signorile et al. 2016). A new study showed that Structure is easily misused when sampling is unbalanced,
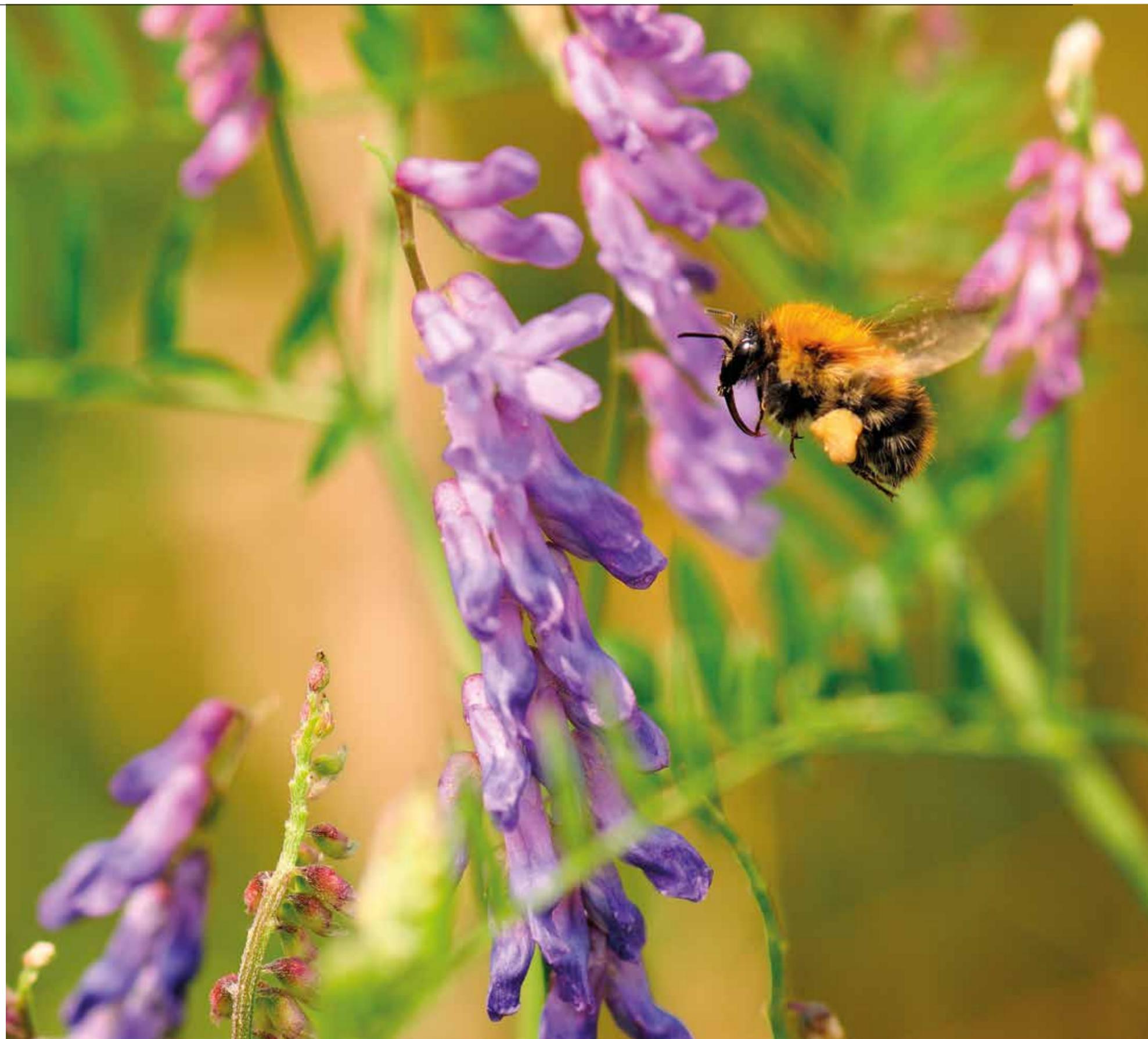
and could yield biased results when the default parameter values of the software were adopted (Wang 2017). Our research has provided alternative parameter values, which yield more accurate population structure inferences in both simulated and real datasets.

Relatedness and relationship ascertainment

Knowing the genetic relatedness among individuals is essential for investigating breeding behaviour, inbreeding and inbreeding depression, social structure, and population demographics and connectivity. Traditionally, relatedness is calculated from pedigrees (eg Brekke et al. 2015). However, pedigree data are rarely available in wild populations.

To facilitate studies of wild populations, we developed methods to estimate relationship and relatedness from genetic marker data, producing a computer program called Coancestry. Our studies, among many others, have used these methods to assess inbreeding and inbreeding depression in diverse species, and were most recently used to investigate the survival and reproduction of bumblebee (*Bombus spp.*) queens (Carvell et al. 2017).

Conservation genetics by non-invasive genotyping

Threatened species are often difficult to sample for genetics studies. Some animals are dangerous, such as large carnivores, or difficult to access, such as marine mammals, while some others are elusive, such as okapis (*Okapia johnstoni*).

Non-invasive genotyping alleviates some of these difficulties by obtaining DNA by indirect sampling of a specimen. This does not require handling (or even the presence of) the animal, as DNA can be extracted from small quantities of hair, feathers, scales or scats, which can be genotyped by polymerase chain reaction. We have applied this technique to a range of species, enabling ecological, evolutionary and conservation genetics studies of species that could not have been studied otherwise.

However, data quality can be problematic because the quality and quantity of DNA from non-invasive samples is often low and genotyping errors are common. An important application of non-invasive genotyping, estimating population census size by the genetic capture-recapture model, is particularly vulnerable to genotyping errors.

We have proposed advanced likelihood methods for identifying individuals from multi-locus genotypes obtained from non-invasive samples, accounting for genotyping errors and missing data (Wang 2016b). Our methods have recently been used to estimate the density and effective population size of cheetahs using scats detected by trained dogs (Becker et al. 2017).

Disease emergence

Infectious diseases can threaten populations, species and entire communities of wildlife, and the consequences of disease can be exacerbated when genetic diversity of wildlife populations is low. Wildlife hosts of pathogens can respond to the emergence of disease through selection, and we have used population genetics to detect both natural and sexual selection in diseased populations.

We have shown how the genetic interactions between host and pathogen are strongly regulated by environmental variation caused by climate change (Clare et al. 2016), and we now need to examine how genes are expressed under these different conditions. Hosts cannot fend off pathogens with a single molecular toolkit, so we will continue to unpick the complexities of host immune gene expression.

> To download Colony, visit zsl.org/science/software/colony – and for Coancestry, visit zsl.org/science/software/coancestry

References

Becker, MS, Durant, SM, Watson, FGR, Parker, M, Gottelli, D, M'soka, J, Droge, E, Nyirenda, M, Schuette, P, Dunkley, S and Brummer, R (2017) Using dogs to find cats: detection dogs as a survey method for wide-ranging cheetah. *Journal of Zoology* **302**: 184-192

Brekke, P, Ewen, JG, Clucas, G and Santure, AW (2015) Determinants of male floating behaviour and floater reproduction in a threatened population of the hihi (*Notiomystis cincta*). *Evolutionary Applications* **8**: 796-806

Bryant, JV, Gottelli, D, Zeng, X, Hong, X, Chan, BPL, Fellowes, JR, Zhang, Y, Luo, J, Durrant, C, Geissmann, T and Chatterjee, HJ (2016) Assessing current genetic status of the Hainan gibbon using historical and demographic baselines: implications for conservation management of species of extreme rarity. *Molecular Ecology* **25**: 3540-3556

Carvell, C, Bourke, AF, Dreier, S, Freeman, SN, Hulmes, S, Jordan, WC, Redhead, JW, Sumner, S, Wang, J and Heard, MS (2017) Bumblebee family lineage survival is enhanced in high-quality landscapes. *Nature* **543**: 547-549

Clare, FC, Halder, JB, Daniel, O, Bielby, J, Semenov, MA, Jombart, T, Loyau, A, Schmeller, DS, Cunningham, AA, Rowcliffe, JM,

Garner, TWJ, Bosch, J and Fisher, MC (2016) Climate forcing of an emerging pathogenic fungus across a montane multihost community. *Philosophical Transactions of the Royal Society* **371**: 201505454

Signorile, AL, Lurz, PWW, Wang, J, Reuman, DC and Carbone, C (2016) Mixture or mosaic? Genetic patterns in UK grey squirrels support a human-mediated 'long-jump' invasion mechanism. *Diversity and Distributions* **22**: 566-577

Wang, J (2015) Does Gst underestimate genetic differentiation from marker data? *Molecular Ecology* **24**: 3546-3558

Wang, J (2016a) A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology* **25**: 4692-4711

Wang, J (2016b) Individual identification from genetic marker data: developments and accuracy comparisons of methods. *Molecular Ecology Resources* **16**: 163-175

Wang, J (2017) The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12650