MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 5966–5968

NEWS AND VIEWS

PERSPECTIVE

Genomics and museum specimens

MICHAEL W. NACHMAN

Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA

Nearly 25 years ago, Allan Wilson and colleagues isolated DNA sequences from museum specimens of kangaroo rats (*Dipodomys panamintinus*) and compared these sequences with those from freshly collected animals (Thomas *et al.* 1990). The museum specimens had been collected up to 78 years earlier, so the two samples provided a direct temporal comparison of patterns of genetic variation. This was not the first time DNA sequences had been isolated from preserved material, but it was the first time it had been carried out with a population sample. Population geneticists often try to make inferences about the influence of historical processes such as selection, drift, mutation and migration on patterns of genetic variation in the present. The work of Wilson and colleagues was important in part because it suggested a way in which population geneticists could actually study genetic change in natural populations through time, much the same way that experimentalists can do with artificial populations in the laboratory. Indeed, the work of Thomas *et al.* (1990) spawned dozens of studies in which museum specimens were used to compare historical and present-day genetic diversity (reviewed in Wandeler *et al.* 2007). All of these studies, however, were limited by the same fundamental problem: old DNA is degraded into short fragments. As a consequence, these studies mostly involved PCR amplification of short templates, usually short stretches of mitochondrial DNA or microsatellites. In this issue, Bi *et al.* (2013) report a breakthrough that should open the door to studies of genomic variation in museum specimens. They used target enrichment (exon capture) and next-generation (Illumina) sequencing to compare patterns of genetic variation in historic and present-day population samples of alpine chipmunks (*Tamias alpinus*) (Fig. 1). The historic samples came from specimens collected in 1915, so the temporal span of this comparison is nearly 100 years.

Keywords: adaptation, conservation genetics, contemporary evolution, ecological genetics, genomics/proteomics

Received 2 October 2013; revised 10 October 2013; accepted 12 October 2013

Correspondence: Michael W. Nachman, Fax: (510) 643-8238; E-mail: mnachman@berkeley.edu

Dealing with degraded DNA

The first step in Illumina sequencing of any sample of DNA is to break the DNA into short fragments. Thus, the short fragments of old DNA, which are problematic for PCR amplification of genes, are not a problem for Illumina sequencing. It is as if nature has performed the first step in the Illumina-library preparation protocol for us. This opens the door to surveys of large portions of the genome through comparisons of many small overlapping fragments.

In this case, Bi *et al.* (2013) first generated a transcriptome using deep sequencing of fresh material from one contemporary specimen. They combined transcripts from a number of different tissues (liver, kidney, spleen and heart) to increase the likelihood that many genes would be surveyed. Importantly, the alpine chipmunk is not a model species, and no whole-genome sequence is yet available for this species. Nonetheless, the authors were able to assemble and annotate many transcripts de novo, showing that this approach can be used for species that have not previously been characterized genetically (Bi *et al.* 2012). From this transcriptome, they then designed 60-bp probes for a subset of exons and used these to build a capture array, against which barcoded DNA from either museum specimens or modern samples was hybridized. This process captured a large subset of genes in the genome, and these were then subject to sequencing.

In addition to being degraded into short fragments, DNA from museum specimens may be damaged in other ways. In particular, post-mortem nucleotide changes may occur, introducing errors in the DNA sequence. The most common of these involve deamination of cytosine to uracil, which is then converted to thymine by DNA polymerase, creating an apparent C to T nucleotide change (or G to A on the complementary strand). Bi *et al.* (2013) looked for evidence of such damage by comparing the frequency of all 12 kinds of nucleotide changes among the four bases in historic and in modern samples. They found that the frequencies were similar except for C to T and G to A changes, which were much higher in historic compared to modern samples, and therefore probably due to DNA damage. The authors used a conservative approach to this problem and simply eliminated all of these sites from further analyses. In the end, after applying this and various other quality filters, the authors were able to compare 3 Mb of sequence with 1578 variable sites in ~11 000 exons across the genome.

Low-medium coverage sequence data, such as the authors generated here, also present challenges in correctly calling genotypes, as some individuals may be represented by few sequencing reads for a particular site. The authors circumvented this problem using a Bayesian probabilistic

Fig. 1 Yosemite National Park and alpine chipmunks (*Tamias alpinus*). Left panels show habitat near Vogelsang Lake in Yosemite National Park in 1915 (above) and 2004 (below). Right panels show museum specimens of alpine chipmunks from the Museum of Vertebrate Zoology at U.C. Berkeley (above) and an animal in the field (below). Photograph credits: Joseph Grinnell, Chris Conroy, Ke Bi, Mark A Chappell.

framework to call genotypes and estimate allele frequencies. This method is based on genotype likelihoods that use information from all individuals at all sites at the same time.

Climate change and range contractions

The authors chose the alpine chipmunk as their subject (Fig. 1). This species lives at high elevations in the Sierra Nevada of California and has experienced a range contraction to higher elevations during the last century as the climate has warmed (Moritz *et al.* 2008). The authors were interested in asking how this range contraction affected patterns of genetic variation. They found that the total amount of genetic variation, as well as the distribution of allele frequencies, had not changed in the nearly 100 years between sampling dates. This is perhaps not surprising as substantial reductions in genetic variation are expected only when the duration of a population bottleneck, measured in generations, is large relative to the effective population size of the reduced population. Nonetheless, the authors did find substantial genetic structure in the present-day population and no population structure in the ancestral population, suggesting that there may be reduced gene flow among more isolated present-day populations. This raises the possibility that climate change is reducing genetic connectivity among populations of alpine chipmunks today.

Moving forward with genomic studies

Museum collections have traditionally been used to study such things as phylogenetic relationships and geographic patterns of variation. The study by Bi *et al.* (2013) opens up many avenues for new kinds of research with museum specimens. We can now anticipate tracking changes in allele frequencies through time at many sites throughout the genome. In the not-too-distant future, we can even imagine whole-genome resequencing of population samples. This should enable us to track the effects of severe demographic processes on patterns of genetic variation. As the oldest museum specimens are only on the order of one hundred or more years old, we will not be able to observe processes that occur over thousands of generations. But some of the most interesting aspects of evolution occur rapidly. For example, selection that acts on standing genetic variation can lead to large and very rapid changes in allele frequencies, over relatively few generations. Genome-wide surveys of genetic variation in museum and present-day samples should thus enable us to identify loci that have responded to selection, including particular environmental changes.

The work conducted by Wilson and colleagues nearly 25 years ago and the study by Bi *et al.* (2013) were both based on specimens from the Museum of Vertebrate Zoology (MVZ) at the University of California, Berkeley. In the interest of full disclosure, I should point out that I am the new director of this institution. The MVZ and many other natural history museums collectively offer a rich source of preserved material that is now amenable to genomic studies. We should remember that these samples are finite, that extraction of DNA from them is destructive and that they were collected in the field often at great effort and expense. Moreover, comparisons such as the one described here require fieldwork today. Nonetheless, the effort required is well spent: studies such as this one should allow us not only to study population genetic patterns of variation but also to make explicit links between genotype and phenotype for traits of ecological importance, thus helping to explain the rich diversity of life on earth.

Acknowledgements

I thank K. Bi, J. Good, A. Geraldes, and one anonymous reviewer for comments on the manuscript.

References

Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, **13**, 403.

Bi K, Linderoth T, Vanderpool D *et al.* (2013) Unlocking the vault: next generation museum population genomics. *Molecular Ecology*, **22**, 6018–6032.

Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, Beissinger SR (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science*, **322**, 261–264.

Thomas WK, Paabo S, Villablanca FX, Wilson AC (1990) Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. *Journal of Molecular Evolution*, **31**, 101–112.

Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, **22**, 634–642.

This paper was written exclusively by M.W.N.

---

doi: 10.1111/mec.12563