Preserving microsatellites? Conservation genetics of the giant Galápagos tortoise.

Etienne Loire[1], Nicolas Galtier[2]

[1] Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UMR CMAEE and C2B, Montpellier (France)

[2] Institute of Evolutionary Sciences, University Montpellier, CNRS, IRD, EPHE (France)

Abstract

Conservation policy in the giant Galápagos tortoise, an iconic endangered animal, has been assisted by genetic markers for ~15 years: mitochondrial DNA and microsatellite loci have been used to delineate thirteen (sub)species, between which hybridization is prevented. Here, comparative reanalysis of a previously published NGS data set reveals a conflict with traditional markers. Genetic diversity and population substructure in the giant Galápagos tortoise are found to be particularly low, questioning the genetic relevance of current conservation practices.

Biological conservation implies collaboration between scientists, biodiversity managers and policy makers. We can only applaud when scientific progress influences policies and practices in a way that is beneficial to nature, and ultimately to humans. A recent Comment in Nature (Nichols 2015) highlights a spectacular conservation initiative geared towards an iconic endangered species – the giant Galápagos tortoise.

The Galápagos Islands host a number of fascinating endemic animals, such as Darwin's finches, Galápagos iguanas and Galápagos penguins and their preservation is clearly a major issue. These Islands are a paramount conservation priority. Remarkable efforts have been made by the Charles Darwin Research Station and the Galápagos National Park Service to protect indigenous species while eradicating invasive ones (e.g. Marquez et al. 2013). The Galapagos tortoise, *Chelonoidis nigra*, is probably the most emblematic of all species of the archipelago. These large reptiles have been heavily impacted by human activities over the centuries. Since the mid-1960s they have been the focus of a specific long-term conservation program, which has been fully successful in terms of population density recovery (e.g. Gibbs et al. 2014, Aguilera et al. 2015).

Population genetics research has been instrumental in defining Galápagos tortoise conservation targets. Thirteen giant Galápagos tortoise species (Poulakakis et al. 2015) have been described, or confirmed, using a combination of geographic, morphological and genetic data – specifically, mitochondrial DNA and 9-12 microsatellite loci (Caccone et al. 2002, Ciofi et al. 2002). Based on this information, population densities have been monitored in a way that aims at preserving, or even restoring, the integrity of the distinct gene pools, i.e. by avoiding inter-specific crosses (e.g. Milinkovitch et al. 2004, Garrick et al. 2012, Edwards et al. 2014). The most recent implementation of this strategy is an attempt to recreate the genome of the famous Lonesome George tortoise, the last member of the *C.* (*nigra*) *abingdoni* species, which died in 2012. The plan is to cross individuals from another species that carry *abingdoni* alleles – hence the recent transfer by helicopter and boat of 32 tortoises from their current habitat to the island of Santa Cruz (Nichols 2015).

Some problems, however, still hamper the task of defining *C. nigra* conservation units. For instance, mitochondrial DNA and microsatellite data do not fully agree (Poulakakis et al. 2012). Published levels of mitochondrial coding sequence divergence between named species of giant Galápagos tortoises (<1%, Caccone et al. 2004) are not higher than typical amounts of within-

species mitochondrial polymorphism in reptiles (Bazin et al. 2006). Importantly, individuals carrying so-called ‘hybrid’ genotypes or unexpected mitochondrial haplotypes, given their geographic origin, are often considered as being the result of non-natural causes (Russello et al. 2007, Poulakakis et al. 2008) and have sometimes been removed from analyses (Garrick et al. 2014), thus increasing measured levels of differentiation between entities. The link between morphological and genetic variation, finally, is anything but clear, the uncommon "saddleback" morphotype being observed in several unrelated species (e.g. Poulakakis et al 2012).

Three years ago, we published the blood transcriptomes of five giant Galápagos tortoises which, based on mtDNA and microsatellite data, were assigned to three distinct species – *becki*, *vandenburghi*, and *porteri* (Loire et al. 2013). It should be noted that these three species are evenly distributed across the mitochondrial genealogy of all *C. nigra* (sub)-species (Loire et al. 2013, supplementary fig 1). To our surprise, our analysis of ~1000 coding loci did not reveal any evidence of genetic structure in this sample, i.e. the homozygosity of the five individuals was not higher than expected under the hypothesis that they belonged to a single panmictic population. We calculated the genetic distance between individuals, defined as the proportion of sites at which the two considered individuals carry distinct genotypes, and found no pattern related to current taxonomy, with for instance individual GA05H from species *becki* being closer to GA05E (*vandenburghi*, distance=$4.2\ 10^{-4}$) than to GA05A (*becki*, distance=$8.9\ 10^{-4}$). The total genetic variation detected in the sample was also quite low – heterozygosity was around 0.1%, consistent with earlier findings (Caccone et al. 2004).

Here we looked at the situation from a comparative standpoint and plotted F_{IT} as a function of log-transformed synonymous diversity, π_S, in a dataset of 54 species of animals for which the transcriptome of at least four individuals was available (Perry et al. 2014, Romiguier et al. 2014, Table S1). F_{IT} measures the departure from Hardy-Weinberg equilibrium. It is expected to equal zero in absence of population structure and inbreeding, and to take positive values when individuals are more homozygous, on average, than expected under the assumption of panmixy. Data on all species in this dataset were analysed via the same procedure (see Gayral et al. 2013 for a complete description of the methodology), thus ensuring optimal comparability between data points. Figure 1 indicates that *C. nigra* harbours lower genetic polymorphism and population structuring than most vertebrates or other animal species, which casts serious doubts on the existence of several differentiated gene pools in this taxon. Interestingly, the data point closest to *C. nigra* in this figure is *Homo sapiens* – a species in which very few people would argue the need to define (sub)species.

The above analysis is based on relatively small samples – from 4 to 11 individuals per species – which in most cases precluded definition of sub-populations and calculation of the F_{ST} statistics. But note that a significantly positive F_{ST}, if any, should imply a significantly positive F_{IT}. The limited sample size, together with a potential impact of sequencing and genotyping errors, might also explain the unexpected observation of negative values of F_{IT} in the least polymorphic species. To control for these possible sources of noise and confirm that our approach has the power to identify genetic structure when it exists, we similarly analysed samples made of individuals from two congeneric species (Galtier 2016, Table S1). Distinctively high values of F_{IT} were obtained with the two-species samples compared to the single-species ones (Figure 1, open circles), demonstrating the power of our approach to detect population genetic substructure when it exists.

Our ~1000-locus analysis demonstrates that microsatellites and mtDNA data, on which the whole conservation strategy in *C. nigra* is based, are not representative of the whole genome. The five individuals we sampled carry divergent mitochondrial haplotypes (Loire et al. 2013, Figure S1) yet they clearly belong to the same species/population. This casts serious doubts on the reliability of current taxonomy and on the existence of truly differentiated gene pools in *C. nigra*. As a consequence, we suggest that the current practice whereby so-called 'purebred' individuals are crossed is pointless and potentially harmful to the species due to the risk of inbreeding depression. The risk is particularly high when large numbers of closely related individuals are simultaneously released in a closed area (Milinkovitch et al. 2004). In the same vein, physical removal from the field of naturally occurring 'hybrids' (Milinkovitch et al. 2007) and sterilization of individuals to avoid undesired crosses (Hunter et al. 2013) might well actively participate in the decline in the mean fitness of local populations (Frankham et al. 2014).

Regretfully, even the most recent publications on this issue, including the description of a new species (Poulakakis et al. 2015), have overlooked the Loire et al. (2013) data and focused on the same limited set of markers developed some 15 years ago. It is odd that, nowadays in the genomic era, the conservation policy regarding Galápagos tortoises is based entirely on just a dozen loci, given the demonstrated potential of genome-wide data to enhance population genetics research in other groups of threatened animals (e.g. Liu et al. 2014, Moore et al. 2014, Der Sarkissian et al. 2015, Dobrynin et al. 2015, Lamichaney et al. 2015). Given the importance of inbreeding depression as a threat to conservation of endangered species (e.g. Allendorf et al. 2010, Walling et al. 2011, Frankham 2015) and the existing doubts regarding the representativeness of microsatellites (Väli et al. 2008, DeFaveri et al. 2013), we suggest that forced inbreeding and sterilization interventions in Galápagos tortoises be delayed until the genetic architecture of the species is

clarified via the analysis of genome-wide data in a sufficiently large sample.

There is some controversy and subjectivity in the definition of relevant conservation units and strategies (Edmands 2007, Funk et al. 2012, Shaffer et al. 2015) but mitochondrial haplotypes and microsatellite alleles are probably not what scientists, and the public, would actually like to preserve.

Acknowledgments

We thank David Manley for helpful suggestions and corrections.

Data Accessibility

Raw dataset and R script used to generate Figure 1 are available at the following address:
https://github.com/loire/GalapagosData_And_Script/
Position in the mitochondrial genealogy of *Chelonoidis nigra* of individuals sampled in this paper is available here:
https://static-content.springer.com/esm/art%3A10.1186%2Fgb-2013-14-12-r136/MediaObjects/13059_2013_3192_MOESM1_ESM.pdf

References

Aguilera WT, Málaga J, Gibbs JP. 2015. Conservation: Giant tortoises hatch on Galapagos island. Nature 517:271.

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. Nat Rev Genet 11:697-709.

Bazin E, Glémin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570-572.

Caccone A, Gentile G, Gibbs JP, Frirts TH, Snell HL, Betts J, Powell JR. 2002. Phylogeography and history of giant Galápagos tortoises. Evolution 56:2052-2066.

Caccone A, Gentile G, Burns CE, Sezzi E, Bergman W, Ruelle M, Saltonstall K, Powell JR. 2004. Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Galápagos tortoises. Mol Phylogenet Evol 31:794-798.

Ciofi C, Milinkovitch MC, Gibbs JP, Caccone A, Powell JR. 2002. Microsatellite analysis of genetic divergence among populations of giant Galápagos tortoises. Mol Ecol 11:2265-2283.

DeFaveri J, Viitaniemi H, Leder E, Merilä J. 2013. Characterizing genic and nongenic molecular

markers: comparison of microsatellites and SNPs. Mol Ecol Resour 13:377-392.

Der Sarkissian C, Ermini L, Schubert M, Yang MA, Librado P, Fumagalli M, Jónsson H, Bar-Gal GK, Albrechtsen A, Vieira FG, Petersen B, Ginolhac A, Seguin-Orlando A, Magnussen K, Fages A, Gamba C, Lorente-Galdos B, Polani S, Steiner C, Neuditschko M, Jagannathan V, Feh C, Greenblatt CL, Ludwig A, Abramson NI, Zimmermann W, Schafberg R, Tikhonov A, Sicheritz-Ponten T, Willerslev E, Marques-Bonet T, Ryder OA, McCue M, Rieder S, Leeb T, Slatkin M, Orlando L. 2015. Evolutionary genomics and conservation of the endangered Przewalski's horse. Curr Biol 25:2577-2583.

Dobrynin P, Liu S, Tamazian G, Xiong Z, Yurchenko AA, Krasheninnikova K, Kliver S, Schmidt-Küntzel A, Koepfli KP, Johnson W, Kuderna LF, García-Pérez R, Manuel Md, Godinez R, Komissarov A, Makunin A, Brukhin V, Qiu W, Zhou L, Li F, Yi J, Driscoll C, Antunes A, Oleksyk TK, Eizirik E, Perelman P, Roelke M, Wildt D, Diekhans M, Marques-Bonet T, Marker L, Bhak J, Wang J, Zhang G, O'Brien SJ. 2015. Genomic legacy of the African cheetah, *Acinonyx jubatus*. Genome Biol 16:277.

Edmands S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol Ecol 16:463-475.

Edwards DL, Garrick RC, Tapia W, Caccone A. 2014. Cryptic structure and niche divergence within threatened Galápagos giant tortoises from southern Isabela Island. Conserv. Genet. 15:1357–1369.

Frankham R, Bradshaw CJA, Brook BW. 2014. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biol Conserv 170:56-63.

Frankham R. 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Mol Ecol 24:2610-2618.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. 2012. Harnessing genomics for delineating conservation units. Trends Ecol Evol 27:489-496.

Galtier N. 2016. Adaptive protein evolution in animals and the effective population size hypothesis. PLoS Genet 12:e1005774.

Gayral P, Melo-Ferreira J, Glémin S, Bierne N, Carneiro M, Nabholz B, Lourenco JM, Alves PC, Ballenghien M, Faivre N, Belkhir K, Cahais V, Loire E, Bernard A, Galtier N. 2013. Reference-free population genomics from next-generation transcriptome data and the vertebrate-invertebrate gap. PLoS Genet 9:e1003457.

Garrick RC, Benavides E, Russello MA, Hyseni C, Edwards DL, Gibbs JP, Tapia W, Ciofi C, Caccone A. 2014. Lineage fusion in Galápagos giant tortoises. Mol Ecol 23:5276-5290.

Gibbs JP, Hunter EA, Shoemaker KT, Tapia WH, Cayot LJ. 2014. Demographic outcomes and

ecosystem implications of giant tortoise reintroduction to Española Island, Galapagos. PLoS One 9:e110742.

Hunter EA, Gibbs JP, Cayot LJ, Tapia W. 2013. Equivalency of Galápagos giant tortoises used as ecological replacement species to restore ecosystem functions. Conserv Biol 27:701-709.

Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin CJ, Wang C, Zamani N, Grant BR, Grant PR, Webster MT, Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371-375.

Liu S, Lorenzen ED, Fumagalli M, Li B, Harris K, Xiong Z, Zhou L, Korneliussen TS, Somel M, Babbitt C, Wray G, Li J, He W, Wang Z, Fu W, Xiang X, Morgan CC, Doherty A, O'Connell MJ, McInerney JO, Born EW, Dalén L, Dietz R, Orlando L, Sonne C, Zhang G, Nielsen R, Willerslev E, Wang J. 2014. Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. Cell 157:785-794.

Loire E, Chiari Y, Bernard A, Cahais V, Romiguier J, Nabholz B, Lourenço JM, Galtier N. 2013. Population genomics of the endangered giant Galápagos tortoise. Genome Biol 14:R136.

Marquez C, Gibbs JP, Carrion V, Naranjo S, Llerena A. 2013. Population response of Giant Galápagos Tortoises to feral goat removal. Restoration Ecol 2:181-185.

Milinkovitch MC, Monteyne D, Gibbs JP, Fritts TH, Tapia W, Snell HL, Tiedemann R, Caccone A, Powell JR. 2004. Genetic analysis of a successful repatriation programme: giant Galápagos tortoises. Proc Biol Sci 271:341-345.

Milinkovitch MC, Monteyne D, Russello M, Gibbs JP, Snell HL, Tapia W, Marquez C, Caccone A, Powell JR. 2007. Giant Galápagos tortoises; molecular genetic analyses identify a trans-island hybrid in a repatriation program of an endangered taxon. BMC Ecol 7:2.

Moore JS, Bourret V, Dionne M, Bradbury I, O'Reilly P, Kent M, Chaput G, Bernatchez L. 2014. Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. Mol Ecol 23:5680-5697.

Nichols H. 2015. Lonesome George 'look-alike' captured for tortoise-breeding effort. Nature News doi:10.1038/nature.2015.18955.

Perry GH, Melsted P, Marioni JC, Wang Y, Bainer R, Pickrell JK, Michelini K, Zehr S, Yoder AD, Stephens M, Pritchard JK, Gilad Y. 2012. Comparative RNA sequencing reveals substantial genetic variation in endangered primates. Genome Res 22:602-610.

Poulakakis N, Glaberman S, Russello M, Beheregaray LB, Ciofi C, Powell JR, Caccone A. 2008. Historical DNA analysis reveals living descendants of an extinct species of Galápagos tortoise. Proc Natl Acad Sci U S A 105:15464-15469.

Poulakakis N, Russello M, Geist D, Caccone A. 2012. Unravelling the peculiarities of island life:

vicariance, dispersal and the diversification of the extinct and extant giant Galápagos tortoises. Mol Ecol 21:160-173.

Poulakakis N, Edwards DL, Chiari Y, Garrick RC, Russello MA, Benavides E, Watkins-Colwell GJ, Glaberman S, Tapia W, Gibbs JP, Cayot LJ, Caccone A. 2015. Description of a new Galapagos giant tortoise species (Chelonoidis; Testudines: Testudinidae) from Cerro Fatal on Santa Cruz Island. PLoS One 10:e0138779.

Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, Faivre N, Loire E, Lourenco JM, Nabholz B, Roux C, Tsagkogeorga G, Weber AA, Weinert LA, Belkhir K, Bierne N, Glémin S, Galtier N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261-263.

Russello MA, Beheregaray LB, Gibbs JP, Fritts T, Havill N, Powell JR, Caccone A. 2007. Lonesome George is not alone among Galápagos tortoises. Curr Biol 17:R317-318.

Shafer AB, Wolf JB, Alves PC, Bergström L, Bruford MW, Brännström I, Colling G, Dalén L, De Meester L, Ekblom R, Fawcett KD, Fior S, Hajibabaei M, Hill JA, Hoezel AR, Höglund J, Jensen EL, Krause J, Kristensen TN, Krützen M, McKay JK, Norman AJ, Ogden R, Österling EM, Ouborg NJ, Piccolo J, Popović D, Primmer CR, Reed FA, Roumet M, Salmona J, Schenekar T, Schwartz MK, Segelbacher G, Senn H, Thaulow J, Valtonen M, Veale A, Vergeer P, Vijay N, Vilà C, Weissensteiner M, Wennerström L, Wheat CW, Zieliński P. Genomics and the challenging translation into conservation practice. Trends Ecol Evol 30:78-87.

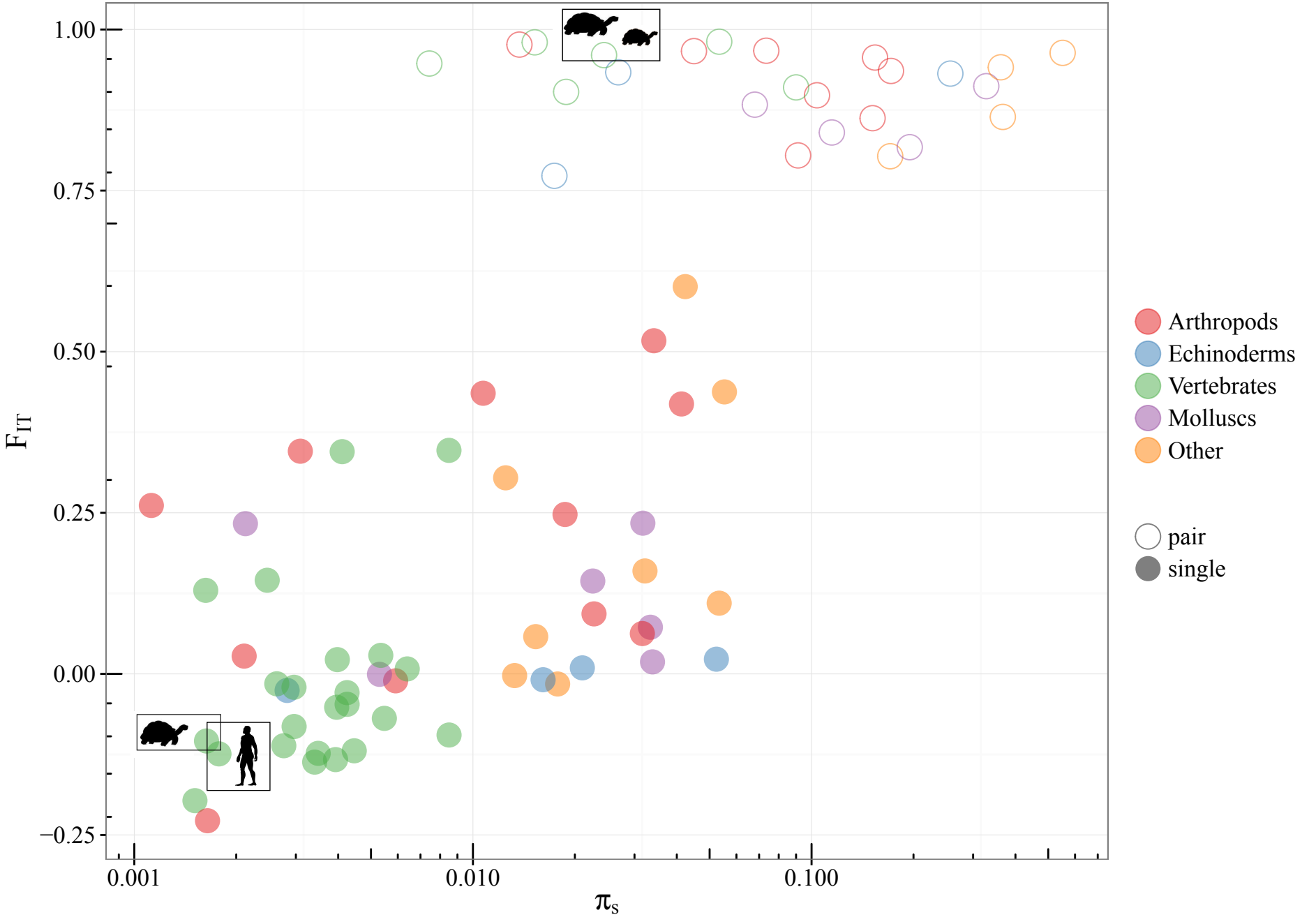
Väli U, Einarsson A, Waits L, Ellegren H. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Mol Ecol 17:3808-3817.

Walling CA, Nussey DH, Morris A, Clutton-Brock TH, Kruuk LE, Pemberton JM. 2011. Inbreeding depression in red deer calves. BMC Evol Biol 11:318.

Figure legend.

Figure 1. Genetic diversity and population substructure in 54 species and 25 pairs of congeneric species of animals.

Each filled circle corresponds to a single species of animals in which the transcriptome of at least four individuals was sequenced (Perry et al. 2012, Romiguier et al. 2014). Each open circle corresponds to a pair of congeneric species of animals in which the transcriptome of at least five individuals was sequenced (Galtier 2016). X-axis (log scale) reflects the amount of neutral genetic variation within the sample. This is measured by the synonymous-site heterozygosity, π_S, for single species, and by the proportion of fixed synonymous differences for species pairs. Y-axis reflects the detected amount of population substructure, measured by the F_{IT} statistics. The positions in the graph of *C. nigra*, *Homo sapiens* (single species) and of the *C. nigra*/*Chelonoidis carbonaria* pair are indicated by thumbnails.


1.00
0.75
0.50
0.25
0.00
−0.25
F_{IT}
0.001
0.010
0.100
π_s
Arthropods
Echinoderms
Vertebrates
Molluscs
Other
pair
single