

Research article

Open Access

Progressive colonization and restricted gene flow shape island-dependent population structure in Galápagos marine iguanas (*Amblyrhynchus cristatus*)

Sebastian Steinfartz*^{†1,2}, Scott Glaberman^{†1}, Deborah Lanterbecq^{1,3}, Michael A Russello⁴, Sabrina Rosa¹, Torrance C Hanley¹, Cruz Marquez⁵, Howard L Snell⁶, Heidi M Snell⁶, Gabriele Gentile⁷, Giacomo Dell'Olmo⁸, Alessandro M Powell⁹ and Adalgisa Caccone¹

Address: ¹Department of Ecology and Evolutionary Biology and Yale Institute for Biospheric Studies - Molecular Systematics and Conservation Genetics Laboratory, New Haven, Connecticut 06511, USA, ²Current address: Department of Animal Behaviour, University of Bielefeld, D-33501 Bielefeld, Germany, ³Current address: Marine Biology Laboratory, University of Mons-Hainaut, 7000 Mons, Belgium, ⁴Department of Biology, University of British Columbia Okanagan, Kelowna, British Columbia V1V 1V7, Canada, ⁵Galápagos National Park, Puerto Ayora, Galápagos, Ecuador, ⁶Museum of Southwestern Biology, Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA, ⁷Department of Biology, Tor Vergata University, 00133 Rome, Italy, ⁸Ornis italica, Piazza Crati 15, 00199 Rome, Italy and ⁹36 Carew Rd, Hamden, CT 06517, USA

Email: Sebastian Steinfartz* - sebastian.steinfartz@uni-bielefeld.de; Scott Glaberman - scott.glaberman@gmail.com; Deborah Lanterbecq - deborah.lanterbecq@gmail.com; Michael A Russello - michael.russello@ubc.ca; Sabrina Rosa - sabrina.rosa@yale.edu; Torrance C Hanley - torrance.hanley@yale.edu; Cruz Marquez - cruz.marquez02@gmail.com; Howard L Snell - snell@unm.edu; Heidi M Snell - hsnell@unm.edu; Gabriele Gentile - gabriele.gentile@uniroma2.it; Giacomo Dell'Olmo - giacomo.dellomo@iss.it; Alessandro M Powell - apowell@hamdenhall.org; Adalgisa Caccone - adalgisa.caccone@yale.edu

* Corresponding author †Equal contributors

Published: 22 December 2009

Received: 5 August 2009

BMC Evolutionary Biology 2009, 9:297 doi:10.1186/1471-2148-9-297

Accepted: 22 December 2009

This article is available from: <http://www.biomedcentral.com/1471-2148/9/297>

© 2009 Steinfartz et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Marine iguanas (*Amblyrhynchus cristatus*) inhabit the coastlines of large and small islands throughout the Galápagos archipelago, providing a rich system to study the spatial and temporal factors influencing the phylogeographic distribution and population structure of a species. Here, we analyze the microevolution of marine iguanas using the complete mitochondrial control region (CR) as well as 13 microsatellite loci representing more than 1200 individuals from 13 islands.

Results: CR data show that marine iguanas occupy three general clades: one that is widely distributed across the northern archipelago, and likely spread from east to west by way of the South Equatorial current, a second that is found mostly on the older eastern and central islands, and a third that is limited to the younger northern and western islands. Generally, the CR haplotype distribution pattern supports the colonization of the archipelago from the older, eastern islands to the younger, western islands. However, there are also signatures of recurrent, historical gene flow between islands after population establishment. Bayesian cluster analysis of microsatellite genotypes indicates the existence of twenty distinct genetic clusters generally following a one-cluster-per-island pattern. However, two well-differentiated clusters were found on the easternmost island of San Cristóbal, while nine distinct and highly intermixed clusters were found on youngest, westernmost islands of Isabela and Fernandina. High mtDNA and microsatellite

genetic diversity were observed for populations on Isabela and Fernandina that may be the result of a recent population expansion and founder events from multiple sources.

Conclusions: While a past genetic study based on pure F_{ST} analysis suggested that marine iguana populations display high levels of nuclear (but not mitochondrial) gene flow due to male-biased dispersal, the results of our sex-biased dispersal tests and the finding of strong genetic differentiation between islands do not support this view. Therefore, our study is a nice example of how recently developed analytical tools such as Bayesian clustering analysis and DNA sequence-based demographic analyses can overcome potential biases introduced by simply relying on F_{ST} estimates from markers with different inheritance patterns.

Background

The legacy of the Galápagos archipelago may be forever attached to the development of Darwin's theory. Yet, the unique character of these islands have continued to make them an ideal model for evolutionary study [1,2]. The Galápagos archipelago is geographically isolated - approximately 1,000 km west of South America - and has never been attached to any continental land mass [3]. Consequently, it is home to many endemic taxa that have colonized the islands either once [4-10] or very few times [9,11,12]. Such a system offers an opportunity to study the radiation of species from a limited ancestral stock without the confounding signals of recurrent colonization. In addition, the islands vary significantly in size and degree of isolation, providing a range of conditions under which to examine the interplay between evolutionary diversification and different demographic processes [13,14]. Finally, the geologic history of the Galápagos is well known, supplying a temporal framework upon which to reconstruct the biogeographic history of various species. The islands were produced by a hotspot that lies beneath the Nazca plate, which is traveling in an eastward direction. Consequently, island ages generally decrease from east to west [3,15-17], and the colonization sequences of many organisms show a progression from older to younger islands (i.e. progression rule or "island progression hypothesis" sensu Wagner and Funk [18]; reviewed in [2]).

Among the organisms that inhabit the Galápagos, one of the most amenable to evolutionary study is the marine iguana (*Amblyrhynchus cristatus*). This species is endemic to the archipelago and is frequently found along the coasts of all the major islands as well as many smaller ones. Marine iguanas exhibit a unique natural history among lizards, feeding almost exclusively on specific algae species in the intertidal or subtidal zones, while breeding and nesting completely on land [19,20]. They possess physical attributes that enable them to negotiate the marine environment, including a flattened tail for swimming and long, sharp nails for clinging to rocks in the surf. The presence of marine iguana populations on islands throughout the archipelago provides a rich system

for examining the roles of island population size, age and isolation, as well as current flow, on patterns of migration, distribution, and long-term population history.

Morphological and genetic data have revealed that the closest relative of *Amblyrhynchus* is the genus of terrestrial iguanas, *Conolophus*, which is also endemic to the Galápagos [8,21-23]. Molecular dating based on mitochondrial DNA (mtDNA) and immunological comparisons yielded a divergence time estimate of 10-20 million years (myr) for the two genera [8,24], which is significantly more ancient than the 3-5 myr age estimate of the oldest islands [3,15,16]. A proposed explanation for these results is that marine and land iguanas diverged from each other on now-sunken islands lying to the east of the present-day archipelago [8,24-26].

A comprehensive genetic study based on both mitochondrial cytochrome b (*cytb*) data as well as nuclear-coded loci (three microsatellite and three minisatellite loci), traced back the microevolution of 22 population/subpopulations of marine iguanas from 15 islands in the archipelago [27]. Despite at least ten million years of independent evolutionary history on the Galápagos, this study suggested that only one or a few related mtDNA haplotypes were involved in the colonization of the present-day archipelago from now-submerged islands [27]. While the overall levels of genetic divergence at both mtDNA and nuclear markers were low, these two markers revealed different patterns of genetic structuring and migration between island populations. Based on *cytb* data, marine iguana populations were grouped into three major lineages: one occupying the older eastern and central islands, another found mostly in the geographically distant northern islands, and a third distributed across the northern and younger western islands. However, this genetic structuring was not supported by nuclear markers, which did not show any clear sign of population differentiation among islands. Based on these results, nuclear gene flow was supposed to be high across the archipelago and is mainly the result of male-biased dispersal as males have been observed to swim to different islands during the breeding season [27]. This was one of the first exam-

ples of male-biased dispersal based on molecular evidence and is still used in textbooks (e.g. p. 229 in [28]).

In this study, we revisit the population genetic structure of marine iguanas in order to lend insight into the patterns of gene flow, genetic diversity, and demographic history of this species. Data is presented from over 1200 individuals from 23 populations sampled at two different time points using an increased number of nuclear markers (thirteen microsatellite loci) as well as the typically fast-evolving mtDNA control region (CR). We use this information to unravel patterns of past and current dispersal in marine iguanas, and discuss results within the context of the island progression hypothesis and sex-biased dispersal.

Methods

Sampling and genetic data collection

Marine iguanas were sampled during two different time periods, 1991/1993 and 2004. Samples from 1991/93 were a subset of those collected and analyzed from the Rassmann et al. study [27]. The total sampling effort resulted in more than 1200 marine iguana specimens spanning 13 islands and 23 populations (Figure 1b; Table 1). Eleven of these locations were sampled during both time periods to meet the objectives of a parallel study examining changes in genetic diversity due to an intense El Niño event [29].

Total genomic DNA was extracted from blood samples in a 96-well format using the QIAamp 96 DNA Blood Kit (QIAGEN Inc.) following the manufacturer's instructions. Complete mtDNA CR sequences (1183 bp) were generated for 1203 marine iguana specimens using PCR protocols that are described elsewhere [29,30]. We genotyped thirteen microsatellite loci for 1225 individuals: locus Am(GT)4 from [27] and twelve loci from [31] following the same procedures as described before [29].

Phylogenetic analysis of CR sequences

CR sequences were edited in the program SEQUENCHER v4.2.2 [32] and aligned using the program MUSCLE v3.6 [33]. Unique haplotypes were defined and numbered using the program DNASP v4.20.2 [34]. The program MRMODELTEST v2 [35], which is based on code from the MODELTEST software [36], was used to evaluate the fit of different nucleotide substitution models to the data. The Hasegawa-Kishino-Yano model [37] with additional parameters for gamma distribution and fraction of invariable sites (HKY+G+I) provided the best fit to the haplotype data according to both the hierarchical likelihood ratio test and the Akaike information criterion. This substitution model was implemented in a Bayesian framework using the program MRBAYES v3.1.2 [38] with a search of 2.5×10^6 generations where the first 10% of parameter samples were discarded as burn-in.

Since the point of branch connection between marine iguana mtDNA phylogroups and Galápagos land iguanas has previously been shown to be unclear due to the large divergence between the two species [27], the root of the marine iguana phylogeny was inferred for CR sequences using a relaxed clock model implemented in the program BEAST v.1.4.7 [39]. As in the phylogenetic analysis, the HKY+G+I model of nucleotide substitution was applied. An evolutionary model was chosen where the substitution rates among branches were uncorrelated, and an expansion growth prior was assumed since marine iguanas exhibit shallow divergence and are likely still in the process of reaching population genetic equilibrium in parts of their range. The analysis was run twice with 2×10^6 generations each, and the first 10% of parameter values were discarded as burn-in. Adequate mixing was determined by examining the effective sample size and parameter trace values as visualized in the program TRACER v1.4 [40] and the two runs were combined to obtain an overall estimate of the posterior distribution of parameters.

CR analysis of genetic diversity and population structure

The number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), and the average number of nucleotide differences between sequences (K) were calculated for each population using DNASP. Data were generated separately for the two temporal samples (1993 and 2004) from Marchena island, as a previous study showed that a bottleneck occurred in between samplings [29]. For all other populations, samples from the two time-points were combined. F_{ST} calculations based on Wright's F_{ST} [41] and an analysis of molecular variance (AMOVA) were performed in the program ARLEQUIN v3.11 [42] in order to determine the level of genetic differentiation within and between marine iguana populations and the following major island groups: north (Pinta, Marchena and Genovesa), west (Fernandina and Isabela), central (Santiago, Rábida, Pinzon, Santa Cruz, Floreana and Santa Fé) and east (Española and San Cristóbal).

We additionally tested for correlation of island age and molecular distance by applying a Mantel test as implemented in the program ARLEQUIN v3.11 [42]. For this F_{ST} differentiation was estimated by haplotype frequencies between island populations and the corresponding age difference between islands was estimated as the difference of maximum island age as provided by Table 1. The significance of test results were determined by performing 10,000 randomizations.

Population structure based on microsatellite loci data

Using microsatellite data, we employed a Bayesian-based analysis of population structure in order to identify genetic clusters, patterns of migration, and gene flow within *Amblyrhynchus* without using any *a priori* sampling information. Starting from individual genotypes, the

Table 1: Information on Galápagos marine iguanas.

Island	Island Age (Myr)	Population size estimates	Population	Microsatellite loci				Mitochondrial Control Region (CR)						
				N	N _{alleles}	H _{obs}	k-test	N	H	Hd	K	F _s	D	
Española	3.31-3.54	1,700-21,000	EPC (Punta Cevallos) ^{a,b}	99	9.07	0.79	7 (p = 0.46)	98	6	0.656	0.00193	2.271	2.65	-1.20
Fernandina	0.04	15,000-120,000	FCH (Capo Hammond) ^a	28	9.5	0.72	9 (p = 0.11)	20	10	0.900	0.00281	3.311	-2.35	-0.80
			FPE (Punta Espinosa) ^{a,b}	99	11.6	0.78	8 (p = 0.25)	86	27	0.946	0.00332	3.911	-11.40**	-0.55
			FPM (Punta Mangle) ^{a,b}	78	10.7	0.79	10* (p = 0.037)	73	32	0.950	0.00328	3.872	-20.84**	-1.48*
Floreana	1.52-2.34	2,000-16,000	FMO (Punta Montura) ^{a,b}	51	8.7	0.78	6 (p = 0.67)	60	12	0.786	0.00423	4.983	0.98	0.18
Genovesa	0.35	900-15,000	GCA (Campamente) ^{a,b}	92	6.7	0.70	6 (p = 0.67)	81	5	0.574	0.00299	3.021	6.26	0.48
Isabela	0.313-535	5,000-40,000	IBU (Bahia Urvina) ^b	17	8.9	0.79	8 (p = 0.25)	15	10	0.924	0.00367	4.324	-2.65	-0.49
			IWE (Cabo West) ^b	-	-	-	-	13	5	0.628	0.00137	1.615	-0.58	-1.45
			ICB (Caleta Black) ^a	-	-	-	-	7	7	1.000	0.00364	4.286	-3.40*	-1.05
			ICW (Caleta Webb) ^{a,b}	40	10	0.74	8 (p = 0.25)	30	8	0.724	0.00333	3.922	1.02	0.64
Marchena	0.56	1,000-10,000	IBA (Cerro Ballena) ^b	22	9.1	0.76	8 (p = 0.25)	21	4	0.271	0.00078	0.914	-0.16	-1.72*
			IPA (Punta Albemarle) ^{a,b}	61	9.4	0.77	7 (p = 0.46)	54	7	0.681	0.00337	3.969	3.42	0.62
			MBN 1993 (Bahia Negra)	30	7.7	0.76	5 (p = 0.84)	29	4	0.685	0.00324	3.823	5.37	1.60

Table 1: Information on Galápagos marine iguanas. (Continued)

			MBN 2004 (Bahia Negra)	49	7.7	0.78	7 (p = 0.46)	49	2	0.250	0.00212	2.500	8.57	0.33
Pinta	0.70	800-6,000	PCI (Caleta Ibetson) ^{a,b}	94	6.1	0.64	6 (p = 0.67)	94	6	0.431	0.00118	1.388	0.63	0.42
Pinzón	1.40-1.73	200-900	PDL (Dumb Landing) ^a	9	5.9	0.76	7 (p = 0.46)	12	4	0.682	0.00084	0.985	-0.65	-0.90
Rábida	1.06-1.56	200-2,000	RAB (No name) ^b	10	5.1	0.71	7 (p = 0.46)	11	3	0.473	0.00308	3.756	4.00	-0.14
San Cristóbal	2.35-4.04	50-400	SRL (Loberia) ^{a,b}	82	6.3	0.70	4 (p = 0.94)	83	3	0.616	0.00220	2.589	7.53	2.67
			SRP (Punta Pitt) ^a	20	3.9	0.51	5 (p = 0.84)	20	1	0.000	0.00000	0.000	NA	NA
Santa Cruz	1.31-2.26	2,000-13,000	SCZ (Estacion/ Camaño) ^{a,b}	135	9.5	0.80	7 (p = 0.46)	116	2	0.017	0.00007	0.086	-0.56	-1.89
Santa Fé	2.85	3,000-16,000	SFN (North) ^a	52	8.4	0.80	8 (p = 0.25)	52	4	0.587	0.00079	0.932	0.72	0.12
			SFM (Miedo) ^{a,b}	82	8.5	0.75	9 (p = 0.11)	82	6	0.702	0.00283	3.341	4.53	2.18
			SFX (Bahia Paraiso) ^b	24	8.4	0.80	7 (p = 0.46)	25	5	0.643	0.00175	2.067	1.14	0.43
Santiago	0.77-1.42	450-4,000	SJB (James Bay) ^{a,b}	51	7.7	0.76	10* (p = 0.037)	72	3	0.547	0.00607	7.160	18.49	3.81
Overall		37,000-280,000		1225				1203	106					

Population localities, sample sizes, island ages, and summary data for 13 microsatellite loci and the mitochondrial control region (CR).

Island ages are in millions of years (myr) and were derived from plate velocity and potassium-argon (KAr) dating estimates from multiple sources [3,17,82,86,87]. Estimates of population sizes were compiled in [78]. **Microsatellite loci data:** *N*, sample size for each population; *N_{alleles}*, average number of alleles; *H_{obs}*, observed heterozygosity; intra-locus *k*-test statistic, number of loci out of 13 loci that showed negative *k* values and the corresponding *p* value (specific *k* values are provided by suppl. table 3). **Control Region data:** *N*, sample size for each population; *H*, number of haplotypes; *H_d*, haplotype diversity; *D*, nucleotide diversity; *K*, average number of nucleotide differences between sequences; *F_s*, Fu's neutrality statistic [51]; *D*, Tajima's neutrality statistic [52]. For the complete table: * and ** denote significance at the = 0.05 and = 0.01 levels respectively. Populations were sampled in 1991/1993 and/or 2004^b.

number of genetic clusters was inferred using the program STRUCTURE v2.1 [43]. K (the number of inferred genetic clusters) ranged from 1 to 25 with 15 iterations for each K . The run length was set to 100,000 MCMC replicates after a burn-in period of one million replicates. For the ancestry model, we chose the admixture model with ALPHA being inferred from the data in combination with correlated allele frequencies. Based on the log probability of these runs, the true value of K was estimated using the approach of Evanno et al. [44]. The program DISTRUCT [45] was then used to graphically display the number of genetic clusters as well as genetic intermixing of individuals based on the true value of K in STRUCTURE. The program BAPS v3.2 [46], which uses a slightly different Bayesian approach than STRUCTURE to infer population differentiation, was also used to identify the optimal number of genetic clusters. The upper bound for the number of populations was set to the number of sampling locations in our dataset. This corresponded to 36 different sampling locations/events when considering different time points from the same location as independent sampling events. The optimal number of clusters identified by BAPS3.2 was selected from a list of the ten best visited partitions according to their $\log(\text{ml})$ values (see [46]). Genotypic assignment of individuals to populations was tested by applying an individual-based self-assignment test using the program GENECLASS2 [47] and the implemented Bayesian approach of Rannala and Mountain [48].

In order to allow a direct comparison between our results and those of the Rassmann et al. study [27], we calculated overall differentiation as Reynolds $F_{ST} = -\ln(1-\Theta)$; [49]) using the program MICROSAT [50]. Analysis of molecular variance (AMOVA) was performed for the same sets of populations and major island groups as done for the equivalent mitochondrial analysis (see above) in the program ARLEQUIN v3.11 [42].

In order to determine the correlation between island age *vs* genetic distance (F_{ST}) we used the same approach as described for the D-loop sequences (see above). The correlation between possible dispersal distance of marine iguanas *vs* genetic distance (F_{ST}) was tested with a Mantel test as implemented in the program ARLEQUIN v3.11 [42]. The significance of test results were determined by performing 10,000 randomizations. Underlying dispersal distances in the matrix (see Table S1 [see Additional file 1]) were estimated under the assumption that iguanas from localities on different islands dispersed by drifting in the ocean, meaning that distances are the closest straight lines around intervening islands between localities. If localities were on the same island, the distances were the closest "coastal" distances - assuming that iguanas dispersed by either swimming or walking along the coast rather than walking across islands.

Analysis of population expansion

Tests of recent population expansion were conducted on the combined population samples (i.e. 1991/93 and 2004 samples; see Table 1) for the mitochondrial and microsatellite datasets separately. For CR sequences, we obtained frequency distributions of pairwise nucleotide differences between sequences (mismatch distributions), Fu's F_s test [51], and Tajima's D statistic [52] using ARLEQUIN. Particular focus was placed on marine iguana populations on the comparatively young islands in the western (Fernandina and western Isabela) and northern (Genovesa, Marchena, and Pinta) regions of the archipelago, which may have been host to recent expansions in both range and population size. Details on methods and parameters used for these analyses are provided in the additional material [see Additional file 2].

The distribution of alleles at microsatellite loci were examined for signs of population expansion using the intra-locus k test as implemented in the Excel Macro KGTESTS [53]. Assuming a simple stepwise mutation model, locus-specific allele length distributions from constant-sized populations should have several modes, whereas a single mode is expected for expanding populations [54]. The intra-locus k test examines differences in allele length distributions for specific loci, where a negative k value is indicative of recent population expansion, as well as whether the number of negative k values across all loci is significant [55].

Tests of sex-biased dispersal

Based on the results of the Rassmann et al. study [27] that indicated male-biased dispersal in marine iguanas, we performed four different tests of sex-biased dispersal for sets of adult individuals sampled in 1991/93 and 2004 [56]: i) The F_{IS} statistic describes how well genotype frequencies follow expectations under Hardy-Weinberg Equilibrium and assumes that the dispersing sex will be a mixture of residents and immigrants resulting in a heterozygote deficiency and a positive F_{IS} value; ii) F_{ST} values should be lower for the dispersing sex because allele frequencies of the dispersing sex should be more similar across populations; iii) Immigrants tend to have a lower Assignment Index (AI; see [57] and [58] for details) than residents. Therefore, the dispersing sex should display, on average, a lower value of AI than the resident sex; iv) Since members of the dispersing sex will include both residents (with common genotypes) and immigrants (with rare genotypes), the variance of AI should be larger for the dispersing sex than for the resident sex. All tests were run in the program FSTAT version 2.9.3.2 [59] and the significance of different test results were determined by performing 10,000 randomizations.

Results

CR-based phylogenetic analysis and clade distribution

Complete mtDNA CR sequences were obtained for 1203 marine iguanas and contained 1183 bp, or 1179 bp when positions with insertions/deletions (indels) were removed as they did not provide significant information and would have extremely limited important applications within the program DNASP (the dataset with gaps can be obtained from SS or SG upon request). D-loop haplotypes have been deposited in GeneBank under accession numbers [GQ293462-GQ293497](#) and [EU278255-EU278326](#) and detailed information on haplotypes (e.g. island and clade designation) is provided by Table S2 [see Additional file 3]. The resulting alignment contained 106 haplotypes with a pairwise uncorrected sequence divergence of 1.6% between the most distant haplotypes. Thirty-five of these haplotypes were singletons.

The topologies of trees generated by the MRBAYES (Figure 1a) and BEAST (data not shown) software were very similar and resulted in three general clades -designated A, B, and C - that were supported by high posterior probability values (PPV > 0.95). Under the relaxed clock model, Bayesian-based analysis in the software BEAST placed clade B as the basal group in the phylogeny. Many islands possessed haplotypes from multiple clades (Figure 1a and Figure 2). Clade A (PPV = 1) possessed the fewest haplotypes ($h = 7$) and was widely distributed across the archipelago; clade B (PPV = 1) contained 30 haplotypes that were found mostly on the central and eastern islands, but also on southern Isabela (IBA and IWE); clade C (PPV = 0.96) accounted for the majority of haplotypes ($h = 69$) and occurred on the northern (Pinta, Marchena, and Genovesa) and western (Fernandina and Isabela) islands as well as on the two central islands of Floreana and Rábida at low frequency. The average pairwise divergence within each of the three clades was low ($\sim 0.4\%$; Table S3a [see Additional file 4]) while the maximum divergence was apparently slightly higher in clade B (1.0%) than for the other two groups (0.8%). The average percent distance between clades ranged from 1.0 - 1.3% (Table S3b [see Additional file 4]).

CR haplotype relationships and distribution

Statistical parsimony networks were constructed separately for each of the three clades. In clade A, haplotype H5 was the only variant that was widespread and found on multiple islands, including San Cristóbal to the south-east, Marchena to the north, and Santiago in the central archipelago (Figure 2). One clade A haplotype (H4) was also found on northern Isabela (IPA), showing that this lineage extends to the far west. In clade B, some haplotypes (e.g. H8, H15) were widespread across several of the eastern and central islands while others (e.g. H9-H12 on Española) formed groups that clustered within specific islands. Members of clade B were generally restricted to

the eastern and central islands, except that the majority of southern Isabela haplotypes were also part of this clade. In addition, a clade B haplotype (H23) was identified from a single individual in the far western FPM population on Fernandina. Haplotypes from the island of Floreana appeared at several disjunct places in the network.

For clade C a high degree of reticulation was observed in the network. As a result, the complete network could not be represented in a clear fashion, and singleton individuals ($N = 26$) were removed from the statistical parsimony analysis to simplify the general phylogeographic patterns. The majority of clade C haplotypes were limited to the far western (Fernandina and western Isabela) and northern (Pinta, Marchena, and Genovesa) regions of the archipelago, although two clade C haplotypes, H60 and H106, were also found on two central islands, Floreana and Rábida.

Population genetic diversity and differentiation based on CR data

Many populations from the west (i.e. Fernandina and western Isabela) showed high levels of genetic diversity ($Hd > 0.900$; Table 1). Conversely, a number of populations from the north, east, and central regions had particularly low levels of variation, many containing three haplotypes or less. The variation was particularly low on Santa Cruz, where 115 of the 116 individuals sampled shared the same CR haplotype.

The analysis of genetic structure (Table 2) showed that populations from the older central/eastern islands were highly differentiated from those in the north ($F_{ST} = 0.74$) and west ($F_{ST} = 0.70$). There was also significant structure between northern and western islands, but to a lesser degree ($F_{ST} = 0.42$). Within regions, genetic distinctiveness among populations was much greater in the eastern/central archipelago ($F_{ST} = 0.55$) than in the north ($F_{ST} = 0.18$) or west ($F_{ST} = 0.09$). However, some eastern/central islands (e.g. Santa Fé, Española) contained many individuals that shared haplotypes with populations from other islands (see Figure 1b). The most striking separation of marine iguana populations in the east/central region was observed on opposite ends of San Cristóbal island ($F_{ST} = 0.87$); a result of the fact that haplotypes from these populations are from two different clades (Figure 1b; Figure 2). The results of the AMOVA, in which populations were grouped into northern, eastern/central, and western regions, showed that variation was similarly partitioned among groups (40.1%), among populations within groups (31.2%), and within populations (28.7%). These results were further supported by the fact that mitochondrial differentiation between island populations correlated with island age ($r = 0.45$; $p = 0.004$) and could explain 20% of the observed variation.

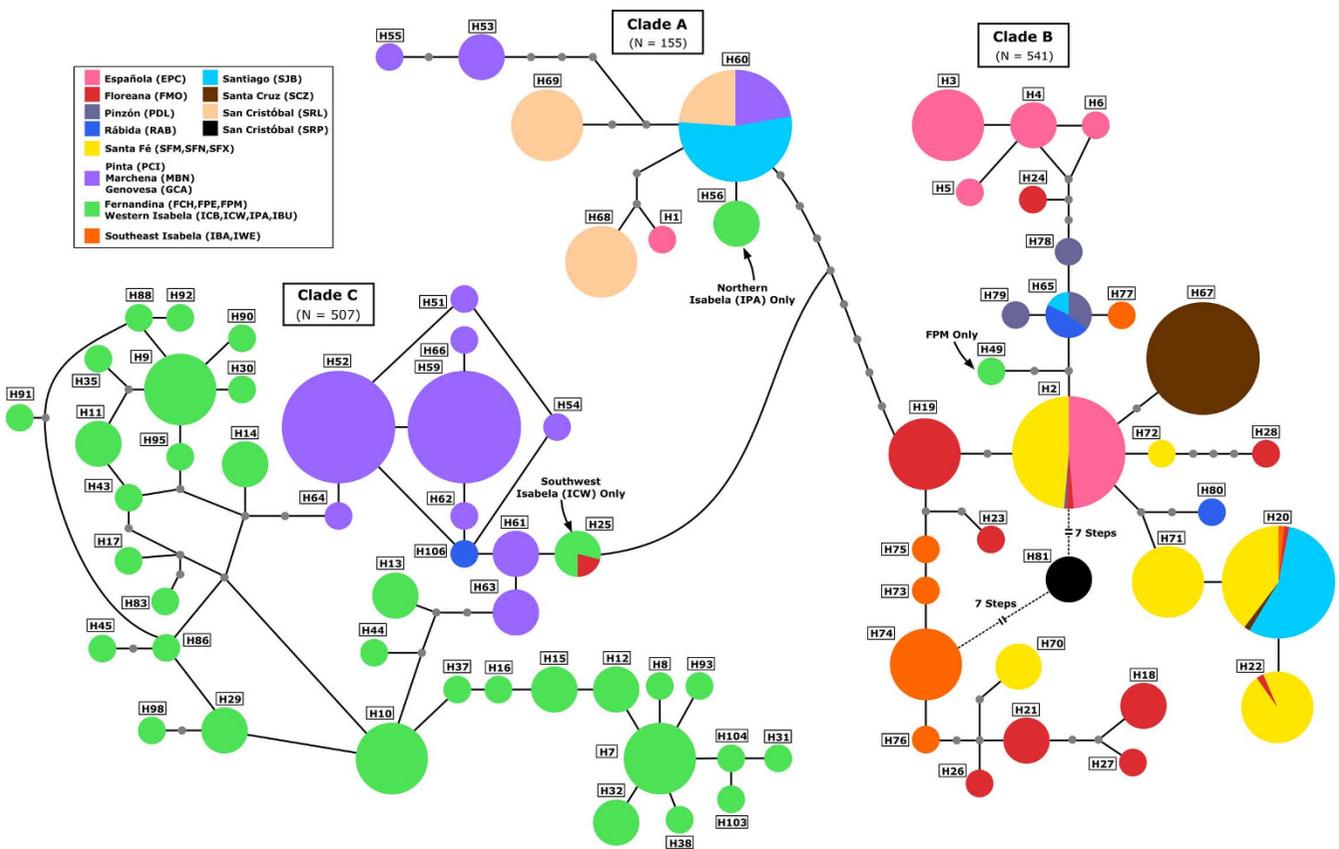


Figure 2
Haplotype network constructed under statistical parsimony for all Galápagos marine iguana mtDNA CR haplotypes. Haplotype numbers are preceded by the letter "H." The smallest circles denote missing haplotypes. The other four circle sizes reflect the number of individuals. In order of increasing area, this corresponds to 1-6, 7-20, 21-50, and 50 or more individuals. All singletons were removed from populations on Fernandina and western Isabela in order to more clearly display the overall phylogeographic patterns. Island distribution of each haplotype is indicated by color according to the embedded legend.

Microsatellite loci dependent population structure and gene flow

Table 1 shows the average number of alleles and observed heterozygosity of each population that was genotyped for thirteen microsatellite loci. Complete genotypic data was obtained for 1225 individuals from 23 sampling locations as well as from different time points for the same island (see Table 1 for details). Applying the approach of Evanno et al. [44], the most likely number of genetic clusters for the complete dataset based on the results of STRUCTURE was estimated at 20 (see Figure 3 and Figure S1 [see Additional file 5]). The individual-based Bayesian structure analysis using BAPS indicated the same number of genetic clusters (see Table S4 [see Additional file 6]).

We found different patterns of genetic differentiation across the archipelago. Individuals sampled at different time points (i.e. 1991/93 and 2004) from the northern (Pinta, Marchena, Genovesa), central (Santiago, Santa Fé,

Santa Cruz) and one of the eastern (Española) islands each form a single island-specific genetic cluster with a low degree of intermixing with genetic clusters from other islands. On the other easternmost island of San Cristóbal, two highly distinct genetic clusters were found corresponding to the western (SRL) and eastern (SRP) populations. The strong genetic differentiation among these islands (as well as among populations within islands in the case of San Cristóbal) was further supported by the results of the individual-based self-assignment test. Proportions of individuals correctly assigned to these island populations ranged from 96-100% indicating low levels of gene flow between islands (Table 3). Gene flow from Española to Floreana was indicated by 12% of individuals sampled on Floreana, but incorrectly assigning to Española. The smaller islands such as Pinzón and Rábida show signatures of genetic exchange with the larger islands of Fernandina and Isabela/Santiago, respectively (see Table 3).

In contrast, STRUCTURE analysis (Figure 3a) showed that individuals sampled on Fernandina and Isabela are genetically similar, with populations on both islands showing a high degree of genetic mixing. The population from Punta Albermarle (IPA) in northern Isabela was an exception since it was genetically differentiated from the remaining populations on Fernandina and Isabela. Also, the self-assignment analysis indicated high levels of gene flow between these two islands as all of the incorrectly assigned individuals on Fernandina (18%) were from Isabela and nearly all incorrectly assigned individuals on Isabela (30%) were from Fernandina (Table 3). Importantly, the nine distinct genetic clusters recovered by our analyses on Fernandina and Isabela are mostly absent from the remaining islands of the archipelago (Figure 3b).

The average F_{ST} values between the 16 populations sampled in 1991/93 and the 18 populations sampled in 2004 were quite similar (0.13 ± 0.057 and to 0.12 ± 0.051 , mean \pm SD, respectively). Within-island F_{ST} differentiation ranged from 0 (on Fernandina) to 0.11 (on San Cristóbal) and F_{ST} between islands from the same region ranged from 0.01 (west) to 0.13 (north). Differentiation between geographic regions was low, ranging from 0.1 (East/Central vs. West) to 0.14 (North vs. East/Central) (see Table 2). In line with these results, the AMOVA analysis revealed that only 3% of the variation was partitioned between geographic groups, 8% among populations within geographic groups, and 89% within populations. Also, no evidence for a relationship between island age and genetic

distance was found on the basis of microsatellite loci ($r = 0.14$; $p = 0.25$). A slight ($r = 0.26$), but highly significantly ($p = 0.001$) correlation between possible dispersal distances and genetic distance was found that could explain 7% of the observed variation between matrices (see also Figure S2 [see Additional file 7] for corresponding scatter plots).

Analysis of population expansion

For the mtDNA dataset, both the western and northern populations conformed to the model of sudden-expansion ($p > 0.05$; Figure S3a,b [see Additional file 8]), which was further supported by low raggedness values ($p > 0.05$). However, only the west showed a clear unimodal mismatch pattern which is suggestive of recent population expansion (Figure S3a-d [see Additional file 8]). Also, highly negative F_S values ($p < 0.05$) were obtained for two sites on Fernandina (FPE and FPM) and for one site on western Isabela (ICB), indicating an excess of recent mutations in these populations due to population growth (see Table 1). Of these three sites, only FPM was also significant for Tajima's D neutrality test ($p = 0.04$), but with a much weaker signal than for the F_S test. In addition, when all the western populations were combined, F_S values were negative and significant ($F_S = -25.729$; $p = 0.00$), but D values were not ($D = -1.144$; $p = 0.11$).

Intra-locus k -test statistics based on microsatellites showed evidence of significant recent population expansion for only two populations. Ten out of thirteen loci (Table 1 and Table S5 [see Additional file 9]) showed negative k values that are indicative of recent population expansion for the FPM site on Fernandina and the Saint James Bay (SJB) site on Santiago.

Analysis of sex-biased dispersal

Table 4 provides the results of the four different test statistics for sex-biased dispersal for marine iguana individuals sampled in 1991/93 and 2004. Although some of the test results are slightly different for females and males sampled in 1991/93 and 2004, none showed significant evidence of detectable genetic consequences from potentially sex-biased dispersal in marine iguanas.

Discussion

Do marine iguanas follow the progression hypothesis?

Genetic studies have revealed that many species in the Galápagos follow the progression hypothesis, where islands are colonized in order of their emergence (reviewed in [2]). This pattern has been attributed to the relative ease in which migrating individuals can successfully occupy younger, uninhabited islands in comparison to older ones, where populations have already been established [18]. The progression hypothesis appears most supported by poorly dispersing species such as giant tortoises

Table 2: Population structure of Galápagos marine iguanas.

	CR (F_{ST})	Microsatellite loci (F_{ST})
Within Islands		
Fernandina	0.00 \pm 0.00	0.00 \pm 0.00
Western Isabela	0.12 \pm 0.07	0.03 \pm 0.01
Santa Fé	0.16 \pm 0.09	0.00 \pm 0.00
San Cristóbal	0.87 \pm 0.00	0.11 \pm 0.09
Within/Between		
North	0.18 \pm 0.08	0.13 \pm 0.08
East/Central	0.55 \pm 0.27	0.11 \pm 0.05
West	0.09 \pm 0.08	0.01 \pm 0.01
North vs. East/Central	0.74 \pm 0.12	0.14 \pm 0.04
North vs. West	0.42 \pm 0.19	0.12 \pm 0.03
East/Central vs. West	0.70 \pm 0.12	0.10 \pm 0.02
Among all 23	0.58 \pm 0.26	0.12 \pm 0.05

Population structure of Galápagos marine iguanas within and between islands and major geographic regions based on CR data (F_{ST} values) and 13 microsatellite loci (F_{ST} values).

Only islands with multiple samples could be used to examine within-island differentiation. The two populations from southeastern Isabela (IBA and IWE) were included in the east/central since they show a higher genetic affinity to those islands. Standard deviations represent the mean over all pairwise comparisons.

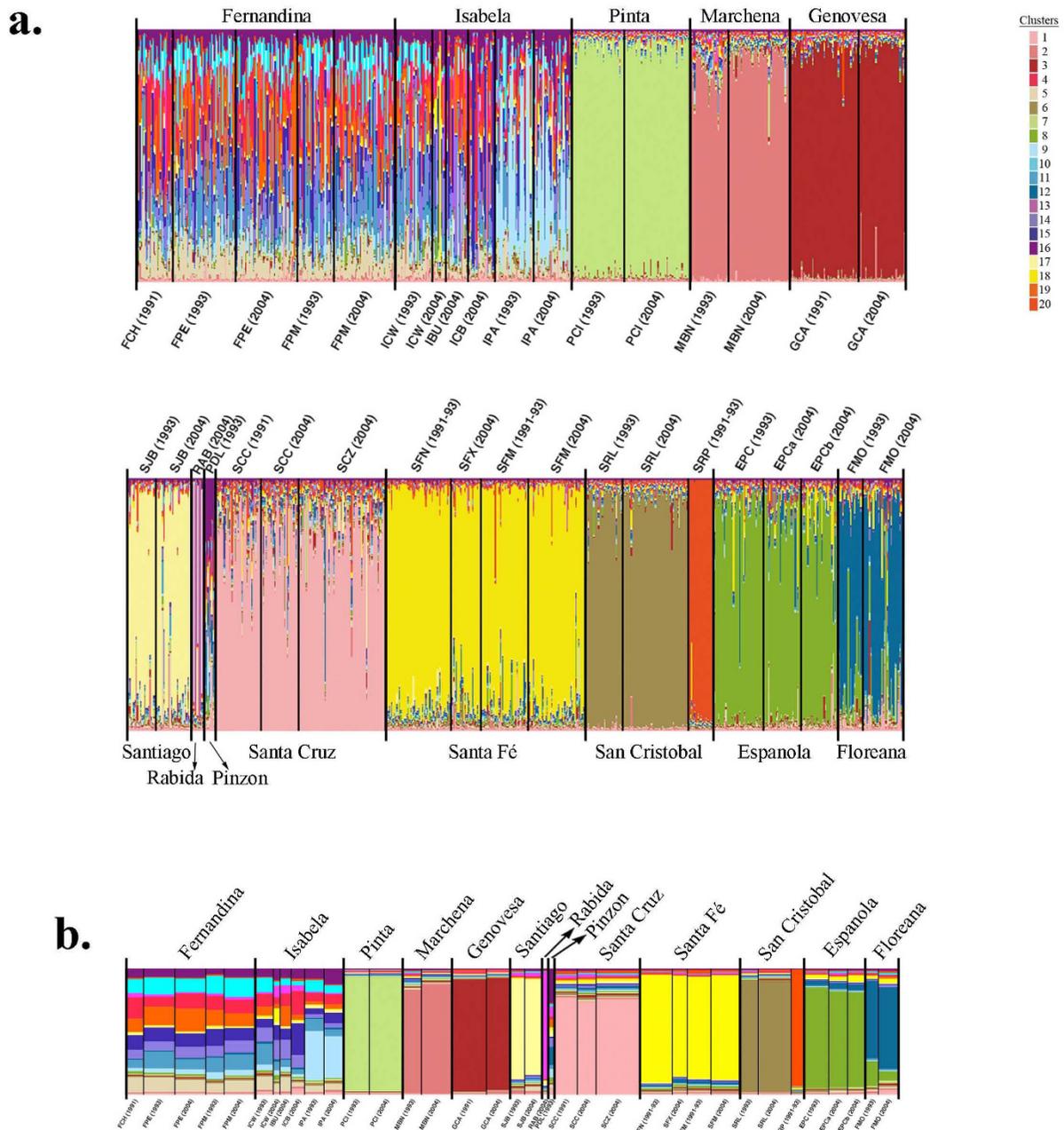


Figure 3
Bayesian mixture analysis. a) Individual-based mixture analysis of 1225 Galápagos marine iguanas to 20 genetic clusters as determined by the program STRUCTURE following the estimation of true number of genetic clusters (K) following [44] Evanno *et al.* (2005; see Figure S1 [see Additional file 5]). Note that individual samples taken at two different time points (i.e. in 1991/93 and 2004) on an island location represent the same genetic cluster. With the exception of Fernandina and Isabela all major islands each represent a distinct genetic cluster. **b)** Population-based mixture analysis for the same number of genetic clusters according to STRUCTURE. For (a) and (b) the program DISTRUCT [45] (Rosenberg 2004) was used to generate the display genetic clusters and degree of admixture of individuals.

Table 3: Results of the self-assignment test.

Island population	Proportion of correctly assigned individuals	Proportion of non-correctly assigned individuals
Fernandina	0.82	0.18 (Isabela)
San Christobal (SCR)	1.0	-
San Christobal (SRP)	1.0	-
Floreana	0.82	0.12 (Española) 0.06 (Santa Cruz)
Genovesa	1.0	-
Marchena	1.0	-
Pinta	1.0	-
Santiago	0.96	0.04 (Isabela)
Santa Cruz	0.98	0.02 (Isabela)
Española	0.96	0.04 (Floreana)
Isabela	0.66	0.3 (Fernandina) 0.007 (Floreana) 0.007 (Santa Cruz) 0.007 (Santa Fé)
Rábida	0.8	0.1 (Isabela) 0.1 (Santiago)
Pinzon	0.78	0.22 (Fernandina)
Santa Fé	0.994	0.006 (Isabela)

Results of the self-assignment test of 1225 Galápagos marine iguana samples using the approach of Rannala and Mountain [48] as implemented in the program GENECLASS2 [47]. Proportion of "correctly" assigned individuals includes individuals that were assigned to their island of origin irrespective of the sampling period (i.e. sampled in 1991/93 or 2004; see Material and Methods for details), whereas the second column gives the proportion of individuals assigned to islands other than the island of origin.

[60,61], lava lizards [62,63], land snails [7], and Galápagos land iguanas [64,65] since the probability of back migration to older islands is low. Conversely, the phylogeographic patterns of more mobile organisms such as Darwin finches ([1]; but see [4]), insects (e.g. weevils; [10]), sea lions [66] and hawks [67] do not always conform to the geologic history of the islands. In the case of Galápagos sea lions and hawks genetic data suggest that population structure was established and affected by quite different factors after a rapid population expansion across the whole archipelago: sea lions diverged according to ecological differences in a western and eastern clade [66], whereas an island-dependent population structure due to limited dispersal between islands was found for the Galápagos hawks [67].

Galápagos marine iguanas feed exclusively on aquatic algae and are closely tied to the coastal environment throughout their lives [20,68,69]. Large individuals are even known to forage offshore in the subtidal zone, swimming up to several hundred meters to reach feeding sites [20]. It is therefore reasonable to expect that these animals can disperse easily - either actively or passively - and that their evolutionary history should not conform to the progression hypothesis.

On a broad scale, the current study provides evidence that the progression hypothesis appears to hold true for marine iguanas. Our CR analyses, as well as the previously published *cytb* study [27], suggest that clade B may be basal to clades A and C (Figure 1a). Biogeographically,

this makes sense since clade B is found almost exclusively on the eastern and central islands (Figure 1b) that are considered to be the oldest in the archipelago. Meanwhile, the majority of clade C haplotypes was identified on younger islands in the north and west. The basal position of clade B would be made stronger if Galápagos land iguanas - the acknowledged sister species of marine iguanas - could be used as an outgroup for phylogenetic analysis. However, when this was attempted using Bayesian-based phylogenetic inference, independent runs failed to converge after 1×10^7 generations. This outcome is likely due to the large divergence between marine and land iguana CR sequences (approximately 10%) compared to the low divergence within marine iguanas (max. 1.6%), reflecting the long timeframe in which the two species have independently evolved on both existing and submerged islands.

Due to the difficulty in rooting the marine iguana phylogeny with CR data, frequency-based genetic information provides important additional support for the progression hypothesis at the large scale. The genetic division between populations from the eastern/central archipelago and those from the northern and western islands was evident in the CR haplotype frequencies, which showed high levels of differentiation as expressed by rather high F_{ST} -values between regions (see Table 2). F_{ST} differentiation based on microsatellite loci for corresponding comparisons is much lower (Table 2) and not in line with the mitochondrial data.

One consequence of the progression hypothesis is that levels of genetic structure are higher among populations on older islands since these are more likely to be at equilibrium and to have undergone significant lineage sorting. For example, in Galápagos tortoises, mitochondrial haplotypes are largely endemic to specific islands in the eastern archipelago while those in the west are often shared between populations and depict an intricate history of genetic exchange [60,61]. A qualitatively similar pattern is also seen for marine iguanas. Mitochondrial CR differentiation was much higher between populations from the older eastern/central archipelago than among the younger western and northern islands. Moreover, levels of genetic differentiation among populations on the youngest (~35,000-535,000 years) and westernmost islands of Fernandina and Isabela were not significantly different from 0 (see Table 2).

There are some exceptions to the general east/west division of marine iguanas in the archipelago that are evident from the CR data. While clade B haplotypes did predominate in the eastern/central region in our study, they were also in the majority on southeastern Isabela, suggesting that this part of the island is biogeographically more connected to the east/central archipelago than to the west. In addition, a single clade B haplotype (H23) was also found on the far-western island of Fernandina, demonstrating the potential for rare long-distance dispersal from the east. Similarly, haplotypes from clade C, which were previously shown to be restricted to the north and west, were also identified on two of the central islands, Floreana and Rábida. A unique clade A haplotype was also identified on northern Isabela, showing that this lineage extends across the entire northern archipelago. The ancestral haplotype in this clade (H5) was widely distributed, existing on San Cristóbal to the east, and Santiago and Marchena in the north-central archipelago. Since San Cristóbal is one of the two oldest islands in the Galápagos, clade A likely radiated from there to the north and west, all the way to the northern tip of Isabela. This fits well with the patterns of island age as well as the path of the prevailing Hum-

boldt current which approaches the archipelago from a southeasterly direction [70] and is thought to have produced similar east to west colonization patterns in other Galápagos organisms [2,63].

While the progression hypothesis generally explains the overall east/west division of marine iguanas, dispersal patterns have produced a more complicated history within regions. Unlike tortoises and other Galápagos organisms that exhibit complete genetic structuring among older populations, many marine iguana CR haplotypes are shared between different islands in the eastern/central region (Figure 1b and Figure 2). For example, more than half of the marine iguanas sampled on Santa Fé have CR haplotypes in common with individuals from other islands. In addition, the statistical parsimony analysis (Figure 2) shows that haplotypes restricted to several of the older islands in the eastern archipelago (e.g. Española and Santa Fé) are connected to other eastern/central island haplotypes via haplotypes from younger central islands (e.g. Pinzón, Isabela), suggesting that migration routes have not proceeded strictly according to the age of island emergence. The appearance of Floreana haplotypes throughout the network shows that this island has been subject to continual genetic exchange with other islands.

Microsatellite loci based population structure - evidence for sex-biased dispersal in marine iguanas?

Bayesian cluster analysis of microsatellite loci genotypes shows that most marine iguanas belong to highly distinct genetic island clusters (Figure 3). This result was independently supported by the high assignment rate of individuals to their island of origin for the majority of islands (Table 3). The Bayesian cluster analysis is especially convincing for two reasons. First, individuals sampled at different time points (i.e. in 1991/93 and 2004) but at the same sampling location were grouped together as one genetic unit. Second, the same number of genetic clusters was obtained by two different Bayesian clustering methods (STRUCTURE and BAPS). The microsatellite loci data therefore suggest that the structuring of marine iguana

Table 4: Test for sex-biased dispersal in Galápagos marine iguanas.

Sampling period	F_{IS} statistic	F_{ST} statistic	Mean AIC test	Variance AIC test
1991/93	F = 0.023 M = 0.040 p = 0.11	F = 0.117 M = 0.115 p = 0.35	F = 0.124 M = -0.106 p = 0.22	F = 11.24 M = 12.07 p = 0.35
2004	F = 0.033 M = 0.007 p = 0.98	F = 0.106 M = 0.129 p = 0.82	F = -0.145 M = 0.136 p = 0.78	F = 14.49 M = 13.39 p = 0.77

The results of four tests (F_{IS} test, F_{ST} test, corrected Mean Assignment Index test and variance of corrected assignment test; see material and methods and [56] for details) are shown for females (F) and males (M) as sampled in 1991/93 and 2004. Note that none of the tests for both sampling periods provide significant differences (p-value) concerning dispersal of sexes.

populations generally follows a pattern of one genetic unit per island. A similar pattern was also recently demonstrated for Galápagos lava lizards from the genus *Microlophus* [71].

Based on F_{ST} analyses of nuclear genetic data, high levels of gene flow between islands and populations were reported [27]. The overall F_{ST} was approximately 0.1 for the 22 marine iguana populations examined. Although the average F_{ST} differentiation found in our study was in the same range ($F_{ST}= 0.13$ and $F_{ST}= 0.12$, respectively, for 1991/93 and 2004 samplings), the high level of genetic differentiation between islands indicated by Bayesian cluster analyses suggests that recent gene flow has been limited, and consequently that F_{ST} values may provide little information about fine scale population structure (see reviews [72,73]). Another issue of microsatellite-based F_{ST} estimates is the underestimation of genetic structure due to allele size homoplasy. Such a pattern may result from the characteristically high mutation rates and allelic polymorphism of microsatellite loci and may diminish the observed genetic distance (reviewed in [74]). As an example, when microsatellite loci and sequences of the nuclear coded internal transcribed spacer (ITS-1) were analysed in the coral species *Corallium rubrum* across the western Mediterranean, ITS-1 sequences revealed significant genetic structure between different geographic regions while microsatellite loci failed to show any correlation between F_{ST} -based estimates for geographic distances ranging from 30-2650 km [75]. In another study in the same organism, however, very fine scale structuring based on microsatellites was observed for distances of ten meters to one km [76] indicating that spatial genetic structure does exist in this system but may not be detectable using microsatellite-based F_{ST} estimates over large distances and/or time frames.

Since the study of Rassmann et al. found significant differentiation between islands and regions based on maternally inherited mtDNA markers but not on bi-parentally inherited nuclear microsatellite and minisatellite loci, they concluded that male-mediated dispersal and gene flow was likely responsible for this discrepancy [27]. Support for this hypothesis came from frequent field observations in which male marine iguanas have been observed to swim to different islands during the breeding season - apparently to establish breeding territories (e.g. Santa Cruz and Caamaño; [77]). After the breeding season they return to the larger island where they "live". On the other hand females have been observed swimming to some islands (e.g. Plaza Sur, Plaza Norte and "Devine's Bay" in Academy Bay) during the breeding season and returned to their resident island after oviposition (unpublished data).

In our study none of the four tests conducted showed significant evidence for sex-biased dispersal in marine igua-

nas (Table 4). Since the set of individuals analyzed in the Rassman et al. study [27] for three microsatellite loci corresponds to the 1991/93 samples analyzed for thirteen microsatellite loci in our study, it was possible to directly compare of F_{ST} -based test statistics between the two studies. In the former study [27] males displayed a significantly lower F_{ST} -value ($F_{ST}= 0.09$) than females ($F_{ST}= 0.12$) suggesting that males disperse more than females. Conversely, F_{ST} estimates for the same set of males and females, which were genotyped at thirteen microsatellite loci, were not significantly different (males, $F_{ST} = 0.115$; females, $F_{ST}= 0.117$; $p = 0.35$; Table 4). For individuals sampled in 2004, we even found higher F_{ST} values for males ($F_{ST}= 0.129$) than for females ($F_{ST}= 0.106$), but also non-significant ($p = 0.82$). These findings suggest that the results of the Rassman et al. study [27], which were based only on three microsatellite loci, should be approached cautiously. Since sex bias has to be intense to be detected by any of the applied test statistics [56], we cannot rule out that such a bias might exist in the case of single dispersers. However, we can exclude that this is a major demographic characteristic of marine iguana populations, as previously suggested.

Beyond the results of the sex-biased dispersal tests, the observed genetic patterns make high dispersal rates in marine iguanas unlikely. Bayesian-based population structure analysis (Figure 3) indicates that, except in a few cases (e.g. Isabela and Fernandina), recent migration and gene flow between islands is low. The fact that we recovered separate genetic units in the west (location SRL) and east (location SRP) on San Cristóbal (also supported by the mitochondrial data; see Figure 1b; Figure 2) as well as detected population structure between populations on Isabela (IPA *versus* the rest of the Isabela; see Figure 3b) indicates that gene flow of marine iguanas can be limited even along the coasts of the same island. These findings are not in line with expectations of high male-mediated gene flow [27].

Evolution of high genetic diversity on Isabela and Fernandina

Marine iguana populations on the westernmost islands of Isabela and Fernandina are among the largest in the Galápagos, with Fernandina alone numbering in the hundreds of thousands [78]. Individuals from these populations are also the largest in the archipelago with a maximum body mass of 12 kg for individuals in southwestern Isabela [79]. Both large population size and individual body size have been attributed to the high marine productivity generated by upwelling of the nutrient-rich Cromwell current along the western archipelago [19,80]. The same region also harbors nearly 95% of the overall Galápagos penguin (*Spheniscus mendiculus*) population [81].

The western marine iguana populations also possess a high level of genetic diversity for both mitochondrial and microsatellite data, even though they occur on the two most recently emerged islands. The emergence age of the oldest volcano on Isabela is less than 535,000 years, while the emergence age of Fernandina is estimated to be only 35,000 years [82]. In contrast, central and southern islands of the archipelago are 1-4 million years old. Sixty-seven out of the total 106 mitochondrial control region haplotypes are found on Fernandina and Isabela with 65 haplotypes being distinct to these islands (Figure 1 and Figure 2) resulting in high levels of haplotype diversity. In contrast populations on the other islands showed significantly fewer haplotypes and lower levels of haplotype diversity. Also, at the level of microsatellite loci, up to nine different genetic clusters were detected for Fernandina and Isabela (Figure 3b), whereas a single predominant cluster is found in populations from most other islands. These genetic clusters are unique to the Isabela and Fernandina populations and therefore must have recently evolved on these islands.

On a broad scale across all islands mitochondrial as well as nuclear diversity indices did not correlate with island size (i.e. island perimeter; (see Figure S4 [see Additional file 10])). The most obvious explanation for the higher level of genetic diversity on the younger islands is that large population sizes along with recent population growth have resulted in many new mutations combined with decreased effects of genetic drift. Indeed, western populations on Isabela and Fernandina show clear signs of sudden-expansion based on mtDNA data as evidenced by a clear unimodal mismatch pattern (Figure S3 [see Additional file 8]) and highly negative F_s values for Punta Espinosa and Punta Mangle on Fernandina (see Table 1) indicating an excess of recent mutations. The microsatellite loci-based k -test statistic showed significant support for recent population expansion for the Punta Mangle population on Fernandina (Table 1 and Table S5 [see Additional file 9]), but not for any of the other western populations.

However, the analysis of CR haplotypes also shows that western marine iguana populations possess mtDNA haplotypes typical of some central, northern, and southern populations. This variation could have been introduced into the western populations through founder events from multiple sources. An increased variation within populations due to multiple recent colonizations has been shown for the invasive brown anolis (*Anolis sagrei*; [83]). In this respect, the finding that mitochondrial CR haplotypes of marine iguanas from southern Isabela are connected to

haplotypes typical of the central islands (Santa Cruz and Floreana) - a similar picture that is seen also in Galápagos tortoises [60,84,85] - underpins the colonization of Isabela from multiple sources. In the land iguana *C. subcristatus*, mitochondrial DNA sequence data indicate a single founder event of western islands from a single source located in the central islands [65]. Subsequently *C. subcristatus* seems to have dispersed in a south to north direction on Isabela. However, the occurrence of an old separate lineage on Isabela (now recognized as a separate species; [65]) indicates that land iguanas have colonized Isabela at least twice.

Conclusions

This comprehensive analysis of the population structure of Galápagos marine iguanas both supports and transforms our previous knowledge about the microevolution of this unique species. The detailed mtDNA analyses trace back migration routes in the evolutionary past and suggest that colonization of islands progressed from geologically older to younger islands in the archipelago. The existence of highly differentiated genetic clusters among islands as well as no specific support for sex-biased dispersal conflict with the previously held view of high male gene flow among islands. More generally, our study demonstrates how the development of recent analytical tools such as Bayesian clustering analysis and DNA sequence-based demographic analyses allow us to tease apart the role of past and present gene flow in shaping current patterns of population differentiation.

Authors' contributions

Project conception and design: SS SG AC HLS. Performed laboratory work: SS SG DL TCH. Data analysis: SS SG DL TCH. Wrote the paper: SS SG AC MR. Figure design: DL SG SS. Field sampling: SS SG AC MR CM SR HMS GG GO AMP. All authors read and approved the final manuscript.

Additional material

Additional file 1

Table S1: Possible dispersal distances of marine iguanas.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S1.XLS>]

Additional file 2

Supporting information on mitochondrial CR analyses [88-93].

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S2.DOC>]

Additional file 3

Table S2: Detailed information on D-loop haplotypes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S3.PDF>]

Additional file 4

Table S3: General characteristics of the three mitochondrial clades of Galápagos marine iguanas.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S4.DOC>]

Additional file 5

Figure S1: Estimation of genetic clusters according to Evanno et al.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S5.PDF>]

Additional file 6

Table S4: Best found partitions according to BAPS v3.2.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S6.DOC>]

Additional file 7

Figure S2: Scatter plot F_{ST} (based on microsatellite loci) versus possible dispersal distance.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S7.PDF>]

Additional file 8

Figure S3: Mismatch distribution of Galápagos marine iguana D-loop haplotypes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S8.PDF>]

Additional file 9

Table S5: Intra locus specific k -values of thirteen microsatellite loci of Galápagos marine iguana populations.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S9.DOC>]

Additional file 10

Figure S4: Scatter plot with regression analysis of island perimeter as measured by NOAA GIS resource (World Vector Coastline) vs genetic diversity values.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-297-S10.DOC>]

Acknowledgements

The authors thank the *Servicio Parque Nacional Galápagos* and the *Charles Darwin Research Station* for granting the sampling permits and for providing the necessary infrastructure, TAME airlines for providing reduced airfare from mainland Ecuador to the islands, Kornelia Rassmann, Fritz Trillmich,

Diethard Tautz, Martin Wikelski, Thomas Roedl and Maren Vitousek for providing blood samples of *A. cristatus*. Ryan Garrick provided valuable comments on an earlier version of the manuscript. The sample collection trip of the year 2004 was enabled through a grant from the National Geographic Society (NGS 7589-04) to Adalgisa Caccone and Sebastian Steinfartz. Sebastian Steinfartz has been supported by a fellowship (STE 1130/2-1 and 1130/2-2) of the German Research Community (DFG) and Deborah Lanterbecq by a postdoctoral fellowship of the Belgian American Educational Foundation. The Yale Institute for Biospheric studies supported this project through funds to Adalgisa Caccone.

References

- Grant PR, Grant BR: **How and why species multiply: the radiation of Darwin's finches.** Princeton: Princeton University Press; 2008.
- Parent CE, Caccone A, Petren K: **Colonisation and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis.** *Philos Trans Roy Soc B* 2008, **363**:3347-3361.
- Cox A: **Ages of the Galápagos Islands.** In *Patterns of Galápagos organisms* Edited by: Bowman RI, Berson M, Leviton AE. San Francisco: Pacific Division, AAAS; 1983:11-23.
- Arbogast BS, Drovetski SV, Curry RL, Boag PT, Seutin G, Grant PR, Grant BR, Anderson DJ: **The origin and diversification of Galápagos mockingbirds.** *Evolution* 2006, **60**:370-382.
- Bollmer JL, Kimball RT, Whiteman NK, Sarasola JH, Parker PG: **Phylogeography of the Galápagos hawk (*Buteo galapagoensis*): A recent arrival to the Galápagos Islands.** *Mol Phyl Evol* 2006, **39**:237-247.
- Caccone A, Gibbs JP, Ketmaier V, Suatoni E, Powell JR: **Origin and evolutionary relationships of giant Galápagos tortoises.** *Proc Nat Acad Sci USA* 1999, **96**:13223-13228.
- Parent CE, Crespi BJ: **Sequential colonization and diversification of Galápagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora).** *Evolution* 2006, **60**:2311-2328.
- Rassmann K: **Evolutionary age of the Galápagos iguanas pre-dates the age of the present Galápagos Islands.** *Mol Phyl Evol* 1997, **7**:158-172.
- Sato A, O'hUigin C, Figueroa F, Grant PR, Grant BR, Tichy H, Klein J: **Phylogeny of Darwin's finches as revealed by mtDNA sequences.** *Proc Nat Acad Sci USA* 1999, **96**:5101-5106.
- Sequeira AS, Lanteri AA, Albelo R, Bhattacharya S, Sijapati M: **Colonization history, ecological shifts and diversification in the evolution of endemic Galápagos weevils.** *Mol Ecol* 2008, **17**:1089-1107.
- Kizirian D, Trager A, Donnelly MA, Wright JW: **Evolution of Galápagos Island Lava Lizards (Iguania: Tropiduridae: Microlophus).** *Mol Phyl Evol* 2004, **32**:761-769.
- Wright JW: **The evolution and biogeography of the lizards of the Galápagos archipelago: evolutionary genetics of *Phyllodactylus* and *Tropidurus* populations.** In *Patterns of evolution in Galápagos organisms* Edited by: Bowman RI, Berson M, Leviton AE. San Francisco: Pacific Division, AAAS; 1983:123-155.
- Beheregaray LB, Caccone A: **Cryptic biodiversity in a changing world.** *J Biol* 2007, **6**:9.
- Beheregaray LB: **Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere.** *Mol Ecol* 2008, **17**:3754-3774.
- Hickman CS, Lipps JH: **Geologic Youth of Galápagos Islands Confirmed by Marine Stratigraphy and Paleontology.** *Science* 1985, **227**:1578-1580.
- Geist D: **On the emergence and submergence of the Galápagos islands.** *Not Galápagos* 1996, **5**:5-9.
- White WM, Mccirney AR, Duncan RA: **Petrology and Geochemistry of the Galápagos-Islands - Portrait of a Pathological Mantle Plume.** *J Geophys Res-Sol Ea* 1993, **98**:19533-19563.
- Wagner WL, Funk VA: *Hawaiian biogeography: evolution on a hot spot archipelago* Washington: Smithsonian Institution Press; 1995.
- Boersma PD: **An ecological study of the Galápagos marine iguana.** In *Patterns of evolution in Galápagos organisms* Edited by: Bowman RI, Berson M, Leviton AE. San Francisco: Pacific Division, AAAS; 1983:157-176.
- Trillmich KGK, Trillmich F: **Foraging Strategies of the marine iguana, *Amblyrhynchus cristatus*.** *Behav Ecol Sociobiol* 1986, **18**:259-266.

21. de Queiroz K: **Phylogenetic systematics of Iguanine lizards: a comparative osteological study.** *Univ Cal Publ Zool* 1987:118.
22. Sites JW, Davis SK, Guerra T, Iverson JB, Snell HL: **Character congruence and phylogenetic signal in molecular and morphological data sets: A case study in the living iguanas (Squamata, Iguanidae).** *Mol Biol Evol* 1996, **13**:1087-1105.
23. Wiens JJ, Hollingsworth BD: **War of the iguanas: Conflicting molecular and morphological phylogenies and long-branch attraction in iguanid lizards.** *Syst Biol* 2000, **49**:143-159.
24. Wyles JW, Sarich VM: **Are the Galápagos iguanas older than the Galápagos?** In *Patterns of evolution in Galápagos organisms* Edited by: Bowman RI, Berson M, Leviton AE. San Francisco: Pacific Division, AAAS; 1983:123-155.
25. Christie DM, Duncan RA, McBirney AR, Richards MA, White WM, Harpp KS, Fox CG: **Drowned Islands Downstream from the Galápagos Hotspot Imply Extended Speciation Times.** *Nature* 1992, **355**:246-248.
26. Werner R, Hoernle K, Bogaard P van den, Ranero C, von Huene R: **Drowned 14-m.y.-old Galápagos archipelago off the coast of Costa Rica: Implications for tectonic and evolutionary models.** *Geology* 1999, **27**:499-502.
27. Rassmann K, Tautz D, Trillmich F, Gliddon C: **The microevolution of the Galápagos marine iguana *Amblyrhynchus cristatus* assessed by nuclear and mitochondrial genetic analyses.** *Mol Ecol* 1997, **6**:437-452.
28. Freeland JR: *Molecular Ecology* Chichester: John Wiley & Sons; 2005.
29. Steinfartz S, Glaberman S, Lanterbecq D, Marquez C, Rassmann K, Caccone A: **Genetic Impact of a Severe El Niño Event on Galápagos Marine Iguanas (*Amblyrhynchus cristatus*).** *PLoS ONE* 2007, **2**:e1285.
30. Hanley TC, Caccone A: **Development of primers to characterize the mitochondrial control region of Galápagos land and marine iguanas (*Conolophus* and *Amblyrhynchus*).** *Mol Ecol Notes* 2005, **5**:599-601.
31. Steinfartz S, Caccone A: **A set of highly discriminating microsatellite loci for the Galápagos marine iguana *Amblyrhynchus cristatus*.** *Mol Ecol Notes* 2006, **6**:927-929.
32. Gene Codes Corporation, Ann Arbor, MI, USA: .
33. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**:1792-1797.
34. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R: **DnaSP, DNA polymorphism analyses by the coalescent and other methods.** *Bioinformatics* 2003, **19**:2496-2497.
35. Nylander JAA: **MrModeltest v2.** In *Program distributed by the author* Evolutionary Biology Centre, Uppsala University; 2004.
36. Posada D, Crandall KA: **MODELTEST: testing the model of DNA substitution.** *Bioinformatics* 1998, **14**:817-818.
37. Hasegawa M, Kishino H, Yano T: **Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.** *J Mol Evol* 1985, **22**:160-74.
38. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19**:1572-1574.
39. Drummond AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees.** *BMC Evol Biol* 2007, **7**:214.
40. Rambaut A, Drummond AJ: **Tracer v1.4.** 2007.
41. Wright S: **The Genetical Structure of Populations.** *Ann Eugenics* 1951, **15**:323-354.
42. Schneider S, Roessli D, Excoffier L: **Arlequin 2000 A software for population genetics data analysis. Version 2.000.** Version 2.000 2000.
43. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155**:945-959.
44. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.** *Mol Ecol* 2005, **14**:2611-2620.
45. Rosenberg NA: **DISTRUCT: a program for the graphical display of population structure.** *Mol Ecol Notes* 2004, **4**:137-138.
46. Corander J, Waldmann P, Sillanpää MJ: **Bayesian analysis of genetic differentiation between populations.** *Genetics* 2003, **163**:367-374.
47. Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A: **geneclass2: a software for genetic assignment and first-generation migrant detection.** *J Hered* 2004, **95**:536-539.
48. Rannala B, Mountain JL: **Detecting immigration by using multilocus genotypes.** *Proc Nat Acad Sci USA* 1997, **94**:9197-9221.
49. Reynolds J, Weir BS, Cockerham CC: **Estimation of the co-ancestry coefficient - basis for a short-term genetic distance.** *Genetics* 1983, **105**:767-779.
50. Minch E, Ruiz-Linares A, Goldstein DB: *Microsat (version 1.5d) a program for calculating statistics on microsatellite allele data* 1995.
51. Fu YX: **Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.** *Genetics* 1997, **147**:915-925.
52. Tajima F: **Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism.** *Genetics* 1989, **123**:585-595.
53. Bignon R: **Kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites.** *Mol Ecol Notes* 2007, **7**:416-417.
54. Reich DE, Goldstein DB: **Genetic evidence for a Paleolithic human population expansion in Africa.** *Proc Nat Acad Sci USA* 1998, **95**:8119-8123.
55. Reich DE, Feldman MW, Goldstein DB: **Statistical Properties of Two Tests that Use Multilocus Data Sets to Detect Population Expansions.** *Mol Biol Evol* 1999, **16**:453-466.
56. Goudet J, Perrin N, Waser P: **Tests for sex-biased dispersal using bi-parentally inherited genetic markers.** *Mol Ecol* 2002, **11**:1103-1114.
57. Paetkau D, Calvert W, Stirling I, Strobeck C: **Microsatellite analysis of population structure in Canadian polar bears.** *Mol Ecol* 1995, **4**:347-354.
58. Favre L, Balloux F, Goudet J, Perrin N: **Female-biased dispersal in the monogamous mammal *Crocodyrus russula*: evidence from field data and microsatellite patterns.** *P Roy Soc B-Biol Sci* 1997, **264**:127-132.
59. Goudet J: **FSTAT (version 1.2): A computer program to calculate F-statistics.** *J Hered* 1995, **86**:485-486.
60. Beheregaray LB, Gibbs JP, Havill N, Fritts TH, Powell JR, Caccone A: **Giant tortoises are not so slow: rapid diversification and biogeographic consensus in the Galápagos.** *Proc Nat Acad Sci USA* 2004, **101**:6514-6519.
61. Ciofi C, Wilson GA, Beheregaray LB, Marquez C, Gibbs JP, Tapia W, Snell HL, Caccone A, Powell JR: **Phylogeographic history and gene flow among giant Galápagos tortoises on southern Isabela Island.** *Genetics* 2006, **172**:1727-44.
62. Benavides E, Baum R, McClellan D, Sites JW Jr: **Molecular phylogenetics of the lizard genus *Microlophus* (Squamata: Tropiduridae): Aligning and retrieving indel signal from nuclear Introns.** *Syst Biol* 2007, **56**:776-797.
63. Kizirian D, Trager A, Donnelly MA, Wright JW: **Evolution of Galápagos Island Lava Lizards (Iguania: Tropiduridae: Microlophus).** *Mol Phy Evol* 2004, **32**:761-769.
64. Tzika AC, Rosa SFP, Fabiani A, Snell HL, Snell HM, Marquez C, Tapia W, Rassmann K, Gentile G, Milinkovitch MC: **Population genetics of Galápagos land iguana (genus *Conolophus*) remnant populations.** *Mol Ecol* 2008, **17**:4943-4952.
65. Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W, Sbordoni V: **An overlooked pink species of land iguana in the Galápagos.** *Proc Nat Acad Sci USA* 2009, **106**:507-511.
66. Wolf JBV, Harrod C, Brunner S, Salazar S, Trillmich F, Tautz D: **Tracing early stages of species differentiation: Ecological, morphological and genetic divergence of Galápagos sea lion populations.** *BMC Evol Biol* 2008, **8**:150.
67. Bollmer JL, Kimball RT, Whiteman NK, Sarasola JH, Parker PG: **Phylogeography of the Galápagos hawk (*Buteo galapagoensis*): A recent arrival to the Galápagos islands.** *Mol Phy Evol* 2007, **39**:237-247.
68. Carpenter CC: **The marine iguana of the Galápagos islands, its behavior and ecology.** *Proc Cal Acad Sci* 1966, **34**:329-376.
69. Wikelski M, Trillmich F: **Foraging Strategies of the Galápagos Marine Iguana (*Amblyrhynchus cristatus*): Adapting Behavioral Rules to Ontogenetic Size Change.** *Behaviour* 1994, **128**:255-279.
70. Wyrutki K: **Circulation and Water Masses in Eastern Equatorial Pacific Ocean.** *Int J Oceanol Limnol* 1967, **1**:1-17.
71. Benavides E, Baum R, Snell HM, Snell HL, Sites JW Jr: **Island biogeography of Galápagos lava lizards (Tropiduridae: Microlophus): species diversity and colonization of the archipelago.** *Evolution* 2009, **63**:1606-1626.

72. Pearse DE, Crandall KA: **Beyond F-ST: Analysis of population genetic data for conservation.** *Cons Gen* 2004, **5**:585-602.
73. Waples RS, Gaggiotti O: **What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity.** *Mol Ecol* 2006, **15**:1419-1439.
74. Estoup A, Jarne P, Cornuet JM: **Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis.** *Mol Ecol* 2002, **9**:1591-1604.
75. Constantini F, Fauvelot C, Abbiati M: **Genetic structuring of the temperate gorgonian coral (*Corallium rubrum*) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences.** *Mol Ecol* 2007, **16**:5168-5182.
76. Constantini F, Fauvelot C, Abbiati M: **Fine-scale genetic structuring in *Corallium rubrum* (L.): evidences of inbreeding and limited effective larval dispersal.** *Mar Ecol-Prog Ser* 2007, **340**:109-119.
77. Trillmich KGK: **The mating system of the marine iguana (*Amblyrhynchus cristatus*).** *Z Tierpsychol* 1983, **63**:141-172.
78. Wikelski M, Nelson K: **Conservation of Galápagos marine iguanas (*Amblyrhynchus cristatus*).** *Iguana* 2004, **11**:191-197.
79. Wikelski M, Trillmich F: **Body size and sexual size dimorphism in marine iguanas fluctuate as a result of opposing natural and sexual selection: An island comparison.** *Evolution* 1997, **51**:922-936.
80. Boersma PD: **An ecological and behavioral study of the Galápagos Penguin.** *Living Bird* 1977, **15**:43-93.
81. Vargas H, Lougheed C, Snell H: **Population size and trends of the Galápagos Penguin *Spheniscus mendiculus*.** *Ibis* 2005, **147**:367-374.
82. Kurz MD, Rowland S, Curtice J, Saal A, Naumann T: **Eruption rates at Fernandina volcano, Galápagos archipelago, from cosmogenic helium surficial laval flows.** *Fall Meeting Supplement* 2005.
83. Kolbe JJ, Glor RE, Schettino LRG, Lara AC, Larson A, Losos JB: **Genetic variation increases during biological invasion by a Cuban lizard.** *Nature* 2004, **431**:177-181.
84. Russello MA, Beheregaray LB, Gibbs JP, Fritts T, Havill N, Powell JR, Caccione A: **Lonesome George is not alone among Galápagos tortoises.** *Curr Biol* 2007, **17**:R317-R318.
85. Poulakakis N, Glaberman S, Russello M, Beheregaray LB, Ciofi C, Powell JR, Caccione A: **Historical DNA analysis reveals living descendants of an extinct species of Galápagos tortoise.** *Proc Nat Acad Sci USA* 2008, **105**:15464-15469.
86. Cullen A, Mcbirney AR: **The Volcanic Geology and Petrology of Isla Pinta, Galápagos Archipelago.** *Geol Soc Am Bull* 1987, **98**:294-301.
87. Naumann T, Geist D: **Physical volcanology and structural development of Cerro Azul Volcano, Isabela Island, Galápagos: implications for the development of Galápagos-type shield volcanoes.** *B Volcanol* 2000, **61**:497-514.
88. Richly E, Leister D: **NUMTs in sequenced eukaryotic genomes.** *Mol Biol Evol* 2004, **21**:1081-1084.
89. Triant DA, Dewoody JA: **The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography.** *J Mammal* 2007, **88**:908-920.
90. Slatkin M, Hudson RR: **Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations.** *Genetics* 1991, **129**:555-62.
91. Rogers AR, Harpending H: **Population growth makes waves in the distribution of pairwise genetic differences.** *Mol Biol Evol* 1992, **9**:552-569.
92. Schneider S, Excoffier L: **Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA.** *Genetics* 1999, **152**:1079-89.
93. Harpending HC: **Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution.** *Hum Biol* 1994, **66**:591-600.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

