

RESEARCH ARTICLE

Open Access



Inconsistent phylogeographic pattern between a sperm dependent fish and its host: *in situ* hybridization vs dispersal

Roland Vergilino[†], Christelle Leung[†] and Bernard Angers^{* ID}

Abstract

Background: Co-dispersal of sperm-dependent hybrids and their sexual relatives is expected to result in consistent spatial patterns between assemblages of hybrids and genetic structure of parental species. However, local hybridization events may blur this signal as assemblages could be organized under different connectivity constraints. This study aims at testing the hypothesis of local hybridization events by comparing the assemblage of hybrid fish *Chrosomus eos-neogaeus* to the genetic diversity of one of its parental species, *Chrosomus eos*.

Results: An extensive survey performed on a total of 132 sites located in two regions of Southern Quebec (West-Qc and East-Qc) revealed a distinct organization of hybrid lineages. One of the six hybrid lineages detected in West-Qc is widespread throughout this region resulting in a low α -diversity (1.38) and β -diversity (4.35). On the other hand, 36 hybrid lineages were detected in East-Qc and displayed narrow geographic distributions leading to a high α -diversity (2.30) and β -diversity (15.68). In addition, the *C. eos* multilocus haplotype of several of these hybrids is assigned to their respective sympatric *C. eos* population. Finally, contrasting with hybrids, the paternal species *C. eos* displayed a higher ρ_{ST} in West-Qc (0.2300) than in East-Qc (0.0734).

Conclusion: The unusually high diversity of hybrid lineages in East-Qc as well as the spatial organization and the close genetic relationship with *C. eos* sympatric populations support the hypothesis that multiple hybridization events occurred *in situ*. These findings coupled to the near absence of the maternal species *Chrosomus neogaeus* suggest that the decline of this species could be the trigger event at the origin of the high rates of spontaneous hybridization in this region.

Keywords: Local hybridization events, Gynogenesis, *Chrosomus eos-neogaeus*, Population genetics, Spatial autocorrelation, Sperm-dependence, Host-parasite

Background

If contemporary evolutionary forces and current landscape elements contribute to the spatial organization of genetic diversity [1–6], historical factors such as geological events, glacial cycles and post-glacial dispersal are also known to play a major role in shaping current species distributions [7–10]. Because all these factors act independently on species populations, two unrelated species are not expected to display a similar genetic structure except if they have been equally influenced by the action of important geological processes (e.g. vicariance

[11–13]) or are linked by strong ecological associations (e.g. host–parasite [14–17]).

In Metazoans, organisms resulting from interspecific hybridization are often associated with asexual reproduction and lead to the formation of all-female unisexual lineages [18–22]. Gynogenesis is the most frequent mode of reproduction observed in unisexual vertebrates [23–25] and is present in fishes and amphibians [25, 26]. Clonal reproduction could result in the competitive exclusion of the parental species population by the asexual hybrids [27, 28] benefiting from demographic advantages (the so-called "Two-fold cost of sex" [29]). In addition, asexual hybrids avoid outbreeding depression by perpetuating the F1 generation and the benefits of heterosis [30]. However,

* Correspondence: bernard.angers@umontreal.ca

[†]Equal contributors

Department of biological sciences, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, Québec H3C 3J7, Canada



gynogenesis, also called sperm-dependent parthenogenesis, requires the sperm of either parental species to trigger embryogenesis [31]. Sperm-dependent unisexual organisms are therefore closely tied to parental species for their persistence and are restricted to the geographic range of their sexual host [32].

Hybrids and their ecological relationships with parental species are of great importance to the study of evolution [18, 33–36] and conservation biology [28, 37, 38]. Hybridization also plays a substantial role in ecological differentiation [39, 40] and species diversification [18, 41–43]. Studying the evolution of gynogenetic organisms is key in order to assess how genomes without recombination deal with genetic decay and what ecological and behavioural processes may create and maintain stable coexistences between sexual and sperm-dependent unisexual organisms [31]. Elucidating when and how often gynogenetic hybrid clones originate is particularly important as a prerequisite to approach these topics. However, uncovering the origin of hybrids often represents a challenging task [44–47].

This study aims at assessing the origin of sperm-dependent hybrid lineages by comparing the distribution of hybrids with one of their parental species. We used fishes from the *Chrosomus eos-neogaeus* complex (Cyprinidae; Teleostei) that includes paternal species northern redbelly dace *Chrosomus eos*, maternal species finescale dace *Chrosomus neogaeus* and gynogenetic hybrids *Chrosomus eos-neogaeus* [48, 49]. While these hybrids are widely distributed in North-America, previous studies reported a low number of hybrid lineages and the presence of the same lineage in different hydrographic networks, suggesting few hybridization events predating the end of the Pleistocene [50, 51]. However, the presence of distinct hybrid lineages within a drainage basin in the Eastern region of Quebec (Canada) raises new questions on the origin and diversity of hybrid lineages [50].

In freshwater environments formerly covered by glaciers, the period during which hybridization events occurred relative to the Pleistocene-Holocene boundary strongly shaped current patterns of diversity at the regional scale [50]. Given the ecological dependence of gynogenetic hybrids to their parental species, different predictions could be formulated depending on whether hybridization events occurred in glacial refuges, during the Pleistocene, or *in situ*, during the Holocene.

Hybrids produced during the Pleistocene may have benefited from the pro-glacial lakes and temporary bridges among hydrographic networks [6, 8, 13, 52, 53] to spread throughout drainage basin(s) of one particular region [50]. This would have resulted in homogeneity in the initial assemblage of hybrids across drainage basins. Following colonization, isolation of hydrographic networks

and the process of lineage sorting may have resulted in different patterns according to the number of lineages in the founder group. Whereas the dispersal of a single lineage would result in a homogeneous distribution of the same genotype across drainage basins [50, 53–55], colonization by multiple lineages would provide a patchwork of assemblages with narrower geographic distribution and lineages present in multiple assemblages. On the other hand, hybridization events may have occurred during the Holocene in conditions similar to the current landscape where each hydrographic network is isolated from each other. These assemblages of hybrids are expected to have a narrow geographic distribution [56] but should harbor assemblage-specific lineages even in the absence of corresponding genetic structure in parental species. Both colonization by multiple lineages and local hybridization events are expected to produce a pattern of spatial autocorrelation limited by very short distances for the distribution of hybrid lineages. However, these scenarios can be discriminated, as only the multilocus genotypes from *in situ* hybridization events are expected to match the genetic diversity in the sympatric parental populations.

In this study, we report the results of an extensive survey of 132 sites from two regions (West-Qc and East-Qc) of Southern Quebec (Canada). We determined the diversity of hybrid lineages and the genetic diversity of parental species. Spatial autocorrelation analysis and assignment tests were performed to assess the hypothesis that hybrids were produced *in situ*.

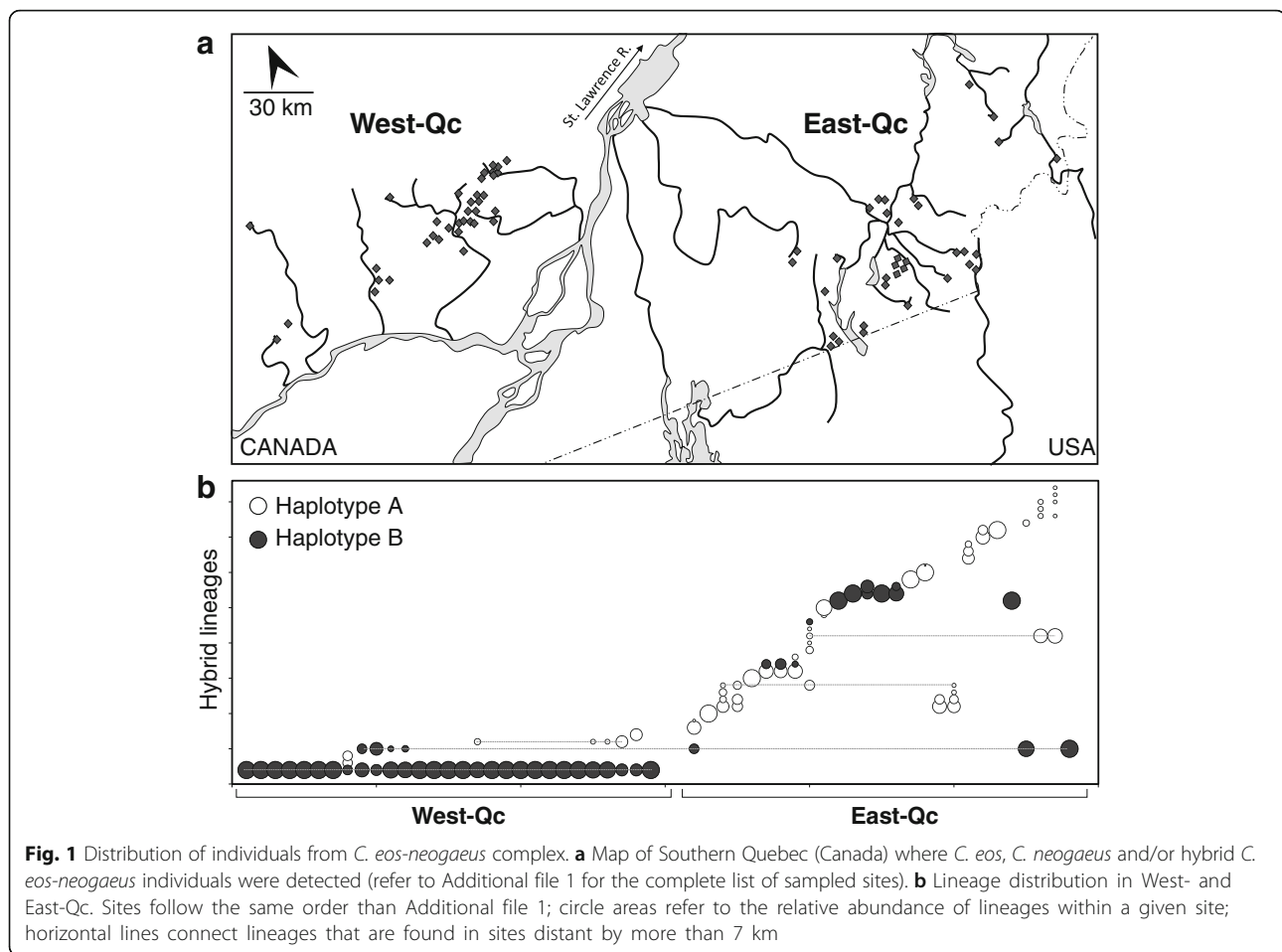
Methods

Sampling and identification

A total of 132 sites were sampled throughout two regions in Southern Quebec (Canada) including 63 sites in the Laurentians region (West-Qc) and 69 sites in the Eastern Townships (East-Qc) that span areas of approximately 3,800 km² and 4,000 km² respectively (Fig. 1a, Additional file 1). Sampled individuals were identified visually as *Chrosomus sp.* according to external morphological characteristics [57]. Morphological identification was confirmed using genetic markers. DNA was extracted from the upper lobe of the caudal fin and genetic identification of the different biotypes of the *C. eos-neogaeus* complex was performed according to Binet and Angers [58].

Analysis of hybrid lineages

For each individual, a multilocus genotype was defined using the size variation of an intron of the PEG1/MEST gene as well as six microsatellite loci (Pho-1, Pho-2, Pho-60, Pho-61, Ca-12 and Seat-412 [50, 58, 59]) amplified according to Angers and Schlosser [50]. Pho-1, Pho-2, Pho-60, Pho-61 are specific to the *C. eos* genome while PEG1/MEST, Ca-12 and Seat-412 amplify both *C. eos* and *C. neogaeus* genomes leading to a total of ten



nuclear loci for hybrid genotypes. *Chrosomus eos* and *C. neogaeus* alleles of hybrid genotypes could be distinguished for PEG1/MEST and Ca-12 loci, but not for Seat-412 for which allelic size ranges overlap between parental species. This locus was then excluded from the analyses that required *C. eos* genotypes.

A hybrid lineage is defined as all individuals originating from a given hybridization event [50]. In the absence of recombination and segregation, individuals of a given lineage were thus expected to display the same multilocus genotype. However, variants could be detected due to the high mutation rate of microsatellite loci. They generally differ by a single mutation at one or a few loci [50]. A given lineage is therefore characterized by a consensus genotype defined by the allele of invariant loci and the more abundant allele for the variable loci.

Triploid hybrids are abundant in the *Chrosomus eos-neogaeus* complex [49]. They are exclusively produced *de novo* following the incorporation of the spermatozoid genome into a diploid hybrid egg [48]. Since genotypes of triploid individuals include the consensus multilocus genotype of a given hybrid lineage in addition to the spermatozoid haplome, it is possible to unambiguously

identify the lineage from which triploid individuals derived. In the absence of a match between a triploid genotype and consensus lineages, this genotype is considered as an additional lineage. However, it becomes impossible to discriminate the *C. eos* multilocus hybrid haplotype from the spermatozoid haplome in lineages represented by a single triploid individual.

The hybrid lineages were further characterized at the mitochondrial level using a fragment of the cytochrome oxidase I (COI) gene to test the hypothesis of a unique *vs* multiple hybridization events. Amplifications were carried out using 5'-CCAGTGTAGCAGCCGGAAT and 5'-GGGTGTCTACGTCTATGCC primers. PCR amplicons were differentiated using the SSCP (Single-Stranded Conformation Polymorphism) method [60] and variants were sequenced.

We infer genetic relationships among hybrid genotypes to assess whether they are derived from single or multiple hybridization events. The clonal distance [61] was used to assess the number of pairwise differences over all microsatellite loci, except Seat-412, according to a stepwise mutation model (SMM). An unrooted Neighbor-Joining tree was constructed using the derived

distance matrix. Assuming a SMM model, if the distinct hybrid lineages derived from a single clone, the unrooted tree should have a star-like conformation with short external branches.

Regional diversity of hybrid lineages was assessed by calculating α - and β -diversity. The degree of uniqueness in terms of their lineage composition was assessed with the Local Contribution to Beta Diversity (LCBD) index [62] for each site. LCBD values were computed using abundance-based Jaccard coefficient [63] and tested using 999 permutations to identify sites that contributed the most to the beta diversity index throughout the sampling area [62].

Mantel correlograms were also constructed to establish the spatial distribution of distinct hybrid lineages for each region (East-Qc and West-Qc). This analysis was used to determine whether sites closer geographically displayed higher similarity in their lineage composition [64, 65]. Historical processes are not well characterized by elements of the current landscape [66], especially when the sampling sites are located in the headwaters of distinct hydrographic networks. Moreover, temporal interconnections may exist between two sites from distinct hydrographic networks [66, 67]; we therefore used straight line distance between sites in addition to waterway distance. Binary (presence/absence) dissimilarity values among sites were used as a dependent variable and geographic distance matrix as an independent variable. Number of distance classes were determined according to Sturges' rule and Mantel statistics were tested with 999 permutation using the correction for multiple tests proposed by Holm [68]. Dissimilarity calculations, Mantel correlograms, and permutations tests were performed using the *vegan* package [69] for the R 3.2.2 software.

Analysis of parental species

Populations of parental species *C. eos* were analyzed using the same nuclear markers than those used for hybrids. For each population, 7–25 individuals were genotyped (Additional file 2). Allele frequencies for each locus and genetic diversity (H_E) were calculated using FSTAT v.2.9.3 [70]. Linkage disequilibrium between locus and departure from Hardy-Weinberg equilibrium was tested for each and all nuclear loci for each population and significance was assessed by 10,000 iterations using GENEPOP v.4.2 [71, 72]. Sequential Bonferroni corrections were applied to all multiple comparisons [73].

A Principal Component Analysis (PCA), as implemented in the *adegenet* package [74, 75] for the R 3.2.2 software was realised to assess the organization of the genetic diversity among *C. eos* populations. To infer population differentiation, allelic differentiation over all

loci and between populations was quantified by computing Weir and Cockerham's [76] estimator of pairwise $F_{ST}(\theta)$ based only on the allelic diversity and Michalakis and Excoffier's [77] estimator of pairwise $Rho_{ST}(\rho_{ST})$ taking into account allelic size using ARLEQUIN v. 5 [78]. Significance of pairwise differences was assessed by 1,000 permutations. The comparison of F_{ST} and ρ_{ST} , which behave differently with regard to mutation, allows to make historical and demographical inferences on population differentiation processes [79, 80]. We estimated $pRho_{ST}$ (ρ_{ST} computed after allele size permutation, [80]) to test if the mutation process has contributed to population differentiation using the allele size permutation test with 20,000 randomization assuming that the loci analyzed follows an SMM-like model as implemented in SpaGeDi v.1.5 [81]. A hierarchical analysis of genetic variation (AMOVA) implemented in ARLEQUIN v. 5 [78] was applied using both F_{ST} and ρ_{ST} to quantify the amount of genetic variance explained by the subdivision in regions (West-Qc and East-Qc) over the total genetic variance. Mantel correlograms were also constructed for *C. eos* populations from each region (East-Qc and West-Qc) by using Cavalli-Sforza and Edward's D_{CE} genetic distance [82, 83] as a function of straight line or waterway distances.

Assignment of the hybrid lineages

Assignments of hybrid lineages to *C. eos* populations was performed using all *C. eos* nuclear loci, except Seat-412. The likelihood that the *C. eos* multilocus haplotype of hybrid lineages originated from a given *C. eos* population was computed using the method based on allele frequencies [84]. The frequency of missing alleles was set at 10^{-5} and the probability was estimated by simulating 10,000 individuals based on the algorithm by Cornuet et al. [85], using the method implemented in GeneClass2 [86]. A percentage score for each population was calculated by dividing the assignment likelihood of each population to the sum of likelihoods across all populations tested.

Results

Distribution of biotypes

Members of the *Chrosomus eos-neogaeus* complex (*C. eos*, *C. neogaeus* and/or hybrids) were captured in 70 of the 132 sampled sites, more specifically in 36 (57.14%) sites in West-Qc and 34 (49.28%) sites in East-Qc. The maternal species *Chrosomus neogaeus* is nearly absent from both regions as successful captures of this species were restricted to three sites (AS-3, SF-12 and CH-1, Additional file 1). Because *C. neogaeus* has been sampled at multiple occasions at sites AS-3 and SF-12, we can exclude particular behaviours or habitat specificity that could hamper its capture and assume a very low

abundance of this species through the sampled area. The paternal species *C. eos* and hybrid individuals were respectively sampled in 27 and 29 sites in West-Qc and 24 and 27 sites in East-Qc (Fig. 1, Additional file 1).

The paternal species *C. eos* and hybrid individuals were sympatric in 37 sites (20 in West-Qc and 17 in East-Qc). A total of 14 sites (7 in each region) contained exclusively *C. eos* individuals. Hybrids were detected in absence of parental species in 18 sites (9 in each region). However, this should be interpreted with caution as the low number of captures at those sites (<6 individuals) reduces the probability of detecting parental species.

Diversity of hybrid lineages

The genetic survey of hybrid individuals revealed a high diversity of genotypes that could be clustered in 41 distinct consensus genotypes. As observed in previous studies [50], several genotypes differed from a consensus by a single or a few mutations (Additional files 1 and 3). Two distinct mitochondrial haplotypes are detected over all hybrid lineages. Sequencing allowed recovery of the A and B haplotypes found by Angers and Schlosser [50] (NCBI accession numbers [EU014286] and [EU014287] respectively). These haplotypes differ by a single bp out of 685 bp which corresponds to a divergence of 0.15%. Haplotypes A and B are associated to 34 and 7 distinct hybrid genotypes respectively (Fig. 1b, Additional file 1).

Of the 41 hybrid lineages, 15 are represented by a single triploid individual (Additional file 3) for which it was impossible to characterize the hybrid genotype. Analysis of the remaining 26 hybrid multilocus genotypes specific to the paternal *C. eos* species (Additional file 3) revealed a high allelic diversity, ranging from 9 (Ca-12) to 15 (Pho-1 and Pho2) alleles. This represents an average of 10.5 alleles/loci and a Nei's gene diversity ranging from

0.815 (Ca-12) to 0.954 (Pho1), values that are similar to the ones measured within *C. eos* populations. The Neighbor-Joining tree of the multilocus genotypes (6 loci) of these 26 hybrid lineages shows long internal branches and does not present a star-like topology (Fig. 2). Lineages did not cluster according to regions, mitochondrial haplotypes or alleles at the intron of the PEG1/MEST gene. There was an average of 50.65 (± 21.73) stepwise mutations among consensus genotypes and up to 122 mutations between the most distant genotypes. On the other hand, the average number of pairwise differences within-lineage estimated on all lineages with more than a single individual was 3.65 (± 2.50) stepwise mutations. In addition, variants of the mitochondrial and PEG1/MEST loci appeared to segregate independently and this topology requires four homoplasious mutations occurring on each of these markers. Altogether, this confirms that the diversity of hybrids is not the result of the diversification of one (or few) clone(s) via mutation process but originated from distinct hybridization events.

The diversity of hybrid lineages is highly different between the two regions: sites in West-Qc (0.0083 ≤ LCBD index ≤ 0.0192) are less unique and diverse than those in East-Qc (0.0243 ≤ LCBD index ≤ 0.0269) and only sites from East-Qc contribute significantly (LCBD, *P* < 0.05) to the total β-diversity. Only 6 lineages were detected in West-Qc; 1 lineage is widespread and was detected in all of the 29 sampling sites containing hybrids. Other lineages are limited to a single site or a few geographically close sites (Figs. 1 and 3). This leads to a very low α-diversity (1.38) as well as β-diversity (4.35). On the other hand, 36 lineages were detected in the 27 sampling sites containing hybrids in East-Qc (Figs. 1 and 3). These lineages displayed a narrow geographic distribution and up

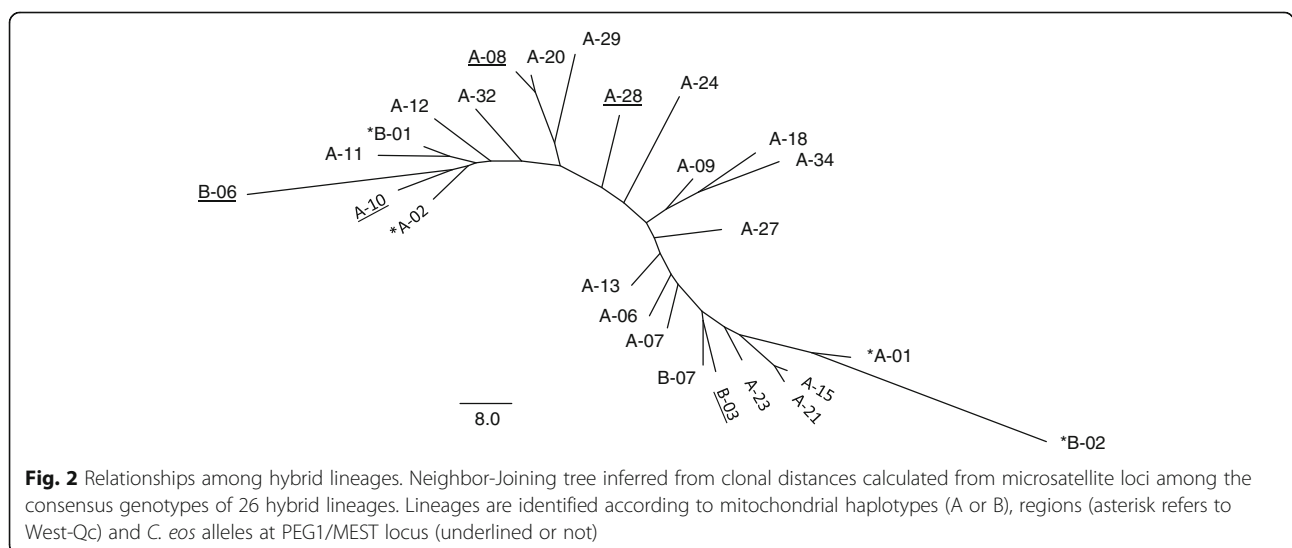
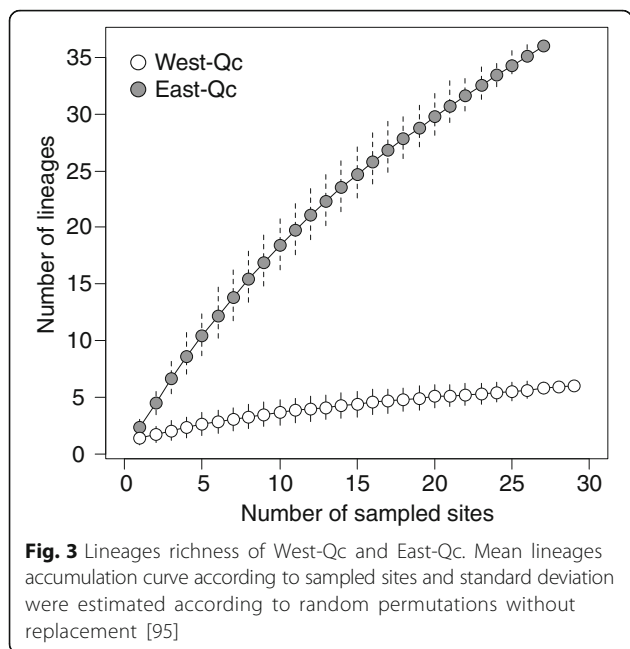


Fig. 2 Relationships among hybrid lineages. Neighbor-Joining tree inferred from clonal distances calculated from microsatellite loci among the consensus genotypes of 26 hybrid lineages. Lineages are identified according to mitochondrial haplotypes (A or B), regions (asterisk refers to West-Qc) and *C. eos* alleles at PEG1/MEST locus (underlined or not)



to six lineages were found in sympatry. This leads to a higher α -diversity (2.30) and β -diversity (15.68) than observed in West-Qc. The geographically close sites harbouring similar hybrid diversity could be grouped into 15 assemblages of hybrid. Only 3 lineages are detected in distinct assemblages (A-09; A-15; B-02). Interestingly, the B-02 hybrid lineage is found in both West-Qc (sites

AS-12 and AS-13; AS-14; AS-15; AS-16) and East-Qc (sites YA-1, SF-19, and SF-22) (Additional file 1).

The assemblage of hybrids is similar among geographically close sites (e.g. within a given minor drainage), but are completely different as distance increases in East-Qc. Therefore, strong and significant positive spatial autocorrelations were observed for the first distance class using straight as well as waterway distances (Fig. 4), indicating that similar lineages are observed in close sites (< 7 km for straight line distance, 15 km for waterway distance) compared to more distant sites.

Significant positive autocorrelations are also detected in West-Qc for the first distance classes ((< 7 km for straight line distance, 16 and 47 km for waterway distance) Fig. 4). The difference observed between straight line and waterway distances suggests the presence of historical and/or contemporary interconnections between drainages for the West-Qc.

The correlation value is significantly higher for straight line distance in East-Qc ($r = 0.703$) compared to West-Qc ($r = 0.248$) (Fisher's $z = 4.1918$, $P < 0.0001$), as well as for waterway distance in East-Qc ($r = 0.6977$) compared to West-Qc (first distance class: $r = 0.0901$; Fisher's $z = 4.1119$, $P < 0.0001$; second distance class: $r = 0.2463$, Fisher's $z = 2.8310$, $P = 0.0023$ [87]).

Characterization of *Chrosomus eos* populations

Chrosomus eos populations were characterized according to 16 sites for which more than 7 individuals were

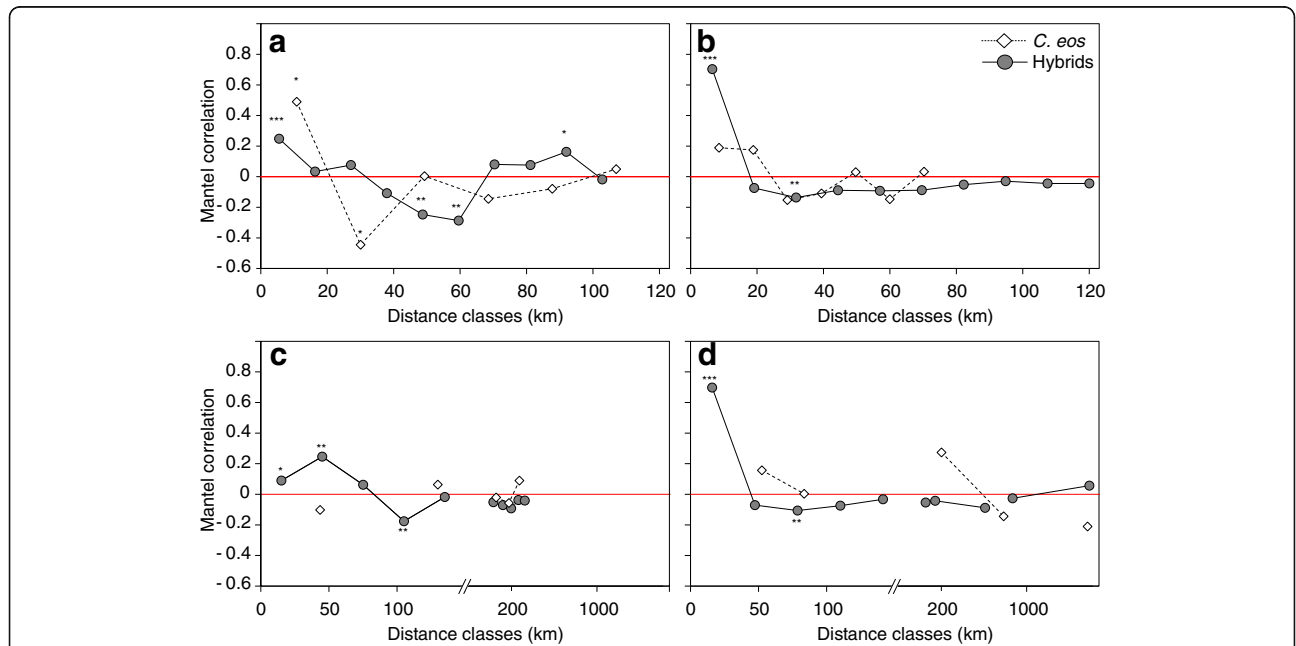


Fig. 4 Mantel correlograms of the distribution of the hybrid lineages and *C. eos*. Jaccard distance and Cavalli-Sforza and Edward's D_{CE} genetic distance were used to assess the differences among sites respectively for hybrid lineages and *C. eos* populations. Straight line distance in (a) West-Qc and (b) East-Qc and waterway distance in (c) West-Qc and (d) East-Qc. Stars indicate significant autocorrelation values: ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$ after Holm corrections

sampled. High allelic diversity is observed for those populations, with the number of alleles per locus ranging from 25 (Ca-12) to 32 (Pho-1). An average of 10.76 alleles/locus/site and 0.88 for Nei's gene diversity (H_E) are observed (Table 1). The number of alleles and Nei's gene diversity are not significantly different between regions ($P > 0.240$).

No significant linkage disequilibrium between loci has been detected (Fisher's method, $P > 0.937$). Global H-W test revealed four populations from West-Qc (AS-5, AS-17, BA-1, RO-1) showing a significant deviation from H-W expectations (using sequential Bonferroni corrections). This is mainly due to a deficiency of heterozygotes at the microsatellite locus Pho-2. When this locus is excluded from the analysis, no further significant deviation from H-W expectations is detected.

Populations are genetically different as only 15 of the 120 pairwise comparisons were not significant (using sequential Bonferroni corrections). All of them involved populations from East-Qc. The PCA analysis revealed a very important overlap among all populations (Fig. 5), as a result of a low global F_{ST} (0.0579) and ρ_{ST} (0.1663). Three populations from West-Qc (AS-17, AS-1, NO-10; Fig. 5) are different from the rest. The difference of F_{ST} between regions (West-Qc: 0.0739; East-Qc: 0.0327) is not significant ($P = 0.180$) but allele size permutation tests performed on ρ_{ST} (West-Qc: 0.2300; East-Qc: 0.0734) provides a significant result ($P < 0.0005$).

The lack of spatial patterning between regions is supported by hierarchical analyses of genetic variation (AMOVA) using either allele diversity ($F_{CT} = 0.00291$, $P = 0.23949$) or allele size diversity ($\rho_{CT} = 0.01620$, $P = 0.23656$). Moreover, Mantel tests revealed no significant isolation-by-distance patterns, neither for the whole area

($R^2 = 0.0137$, $P = 0.132$) nor within each region (West-Qc: $R^2 = 0.0284$, $P = 0.266$; East-Qc: $R^2 = 0.0337$, $P = 0.146$). A significant positive spatial autocorrelation signal is detected in the first distance class in West-Qc but only for straight line distance (Fig. 4a, c). This result is mostly due to the populations AS-1 and NO-10 that are genetically similar one to each other but highly divergent from the other populations; both populations are geographically close but in distinct hydrographic networks (Figs. 4a and 5). On the other hand, no significant spatial autocorrelation was observed for the East-Qc populations (Fig. 4b, d). This result reflects a uniform distribution of alleles in *C. eos* populations, that strongly contrasts with the distribution of hybrid assemblages.

Assignment of hybrid lineages to *Chrosomus eos* populations

Because assignment tests are sensitive to sample size, sites geographically close and harboring the same assemblage of hybrids were pooled together. A total of 18 hybrid lineages are found at these sites (Table 2). However, 3 of them are represented by a single triploid individual and were excluded from this analysis because of our incapacity to infer the *C. eos* haplotype (Additional file 3). Ten of the 12 lineages exclusive to East-Qc are assigned with a high probability to their respective sympatric *C. eos* population and 9 of them with the highest score over all sites (Table 2). Lineage A-01 exclusive to West-Qc is assigned with the higher score to East-Qc populations while lineage B-01 is assigned to one of the three sympatric *C. eos* populations with the highest score. The B-02 lineage detected in both regions is assigned to its sympatric *C. eos* population in East-Qc but not in West-Qc.

Discussion

The joint analysis of *C. eos-neogaeus* hybrid diversity and the paternal species *C. eos* populations genetic diversity in two regions of southern Quebec revealed two striking results. First, a highly distinct geographic organization of the assemblages of hybrids is observed between the two regions. Second, in spite of the sperm dependence of hybrids, there is a lack of correspondence between the spatial organization of hybrid lineages and the genetic diversity of paternal species *C. eos*, this is especially clear in East-Qc.

Uniform distribution of *Chrosomus eos*

The results revealed a lack of population structure for the *C. eos* species in the studied areas. Even the separation of both regions by the St-Lawrence River has a limited impact on the phylogeographic structure of *C. eos* populations as less than 2% of the total genetic variance is explained by the subdivision in regions. Drainage

Table 1 Genetic diversity in hybrid *Chrosomus eos-neogaeus* and in the sexual species *Chrosomus eos* between West-Qc and East-Qc regions

Hybrids <i>C. eos-neogaeus</i>					
	N sites (n)	N lineages	α -diversity	β -diversity	
West-Qc	29 (346)	6	1.38	2.30	
East-Qc	27 (340)	36	4.35	15.68	
Total	56 (686)	41	1.82	22.51	
<i>Chrosomus eos</i> populations					
	N sites (n)	k	H_5	F_{ST}	ρ_{ST}
West-Qc	7 (132)	23.33	0.8653	0.0740	0.2302
East-Qc	9 (118)	25.27	0.8913	0.0327	0.0734
Total	16 (250)	28.33	0.8799	0.0579	0.1663

Number of sites and total number of individuals analysed is provided. For hybrids, diversity is represented by the number of lineages as well as α - and β -diversity. Diversity of the sexual species is given in number of alleles (k), mean of Nei's gene diversity, F_{ST} and ρ_{ST}

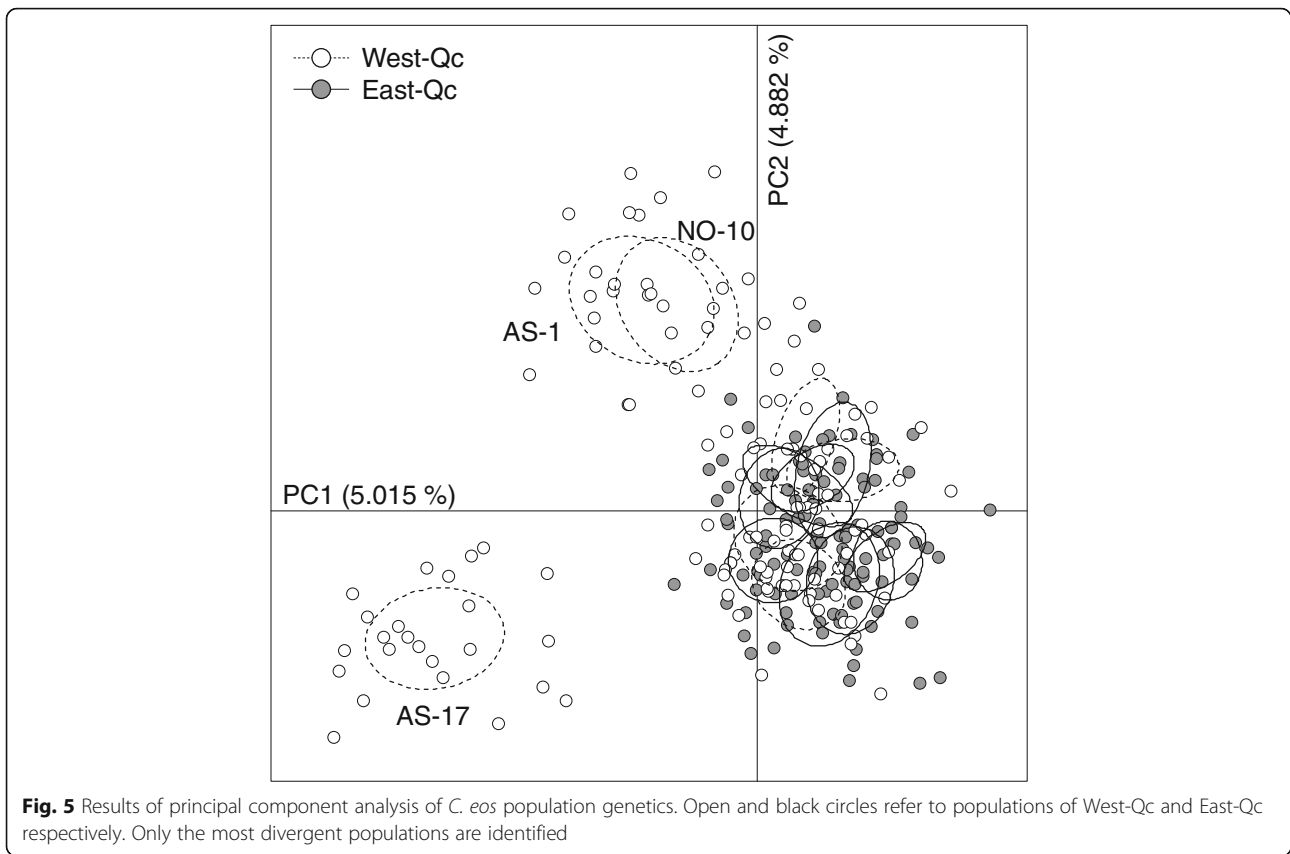


Table 2 Results of the assignment of hybrid lineages to the *Chrosomus eos* populations

Lineages	East-Qc							West-Qc						
	CO-1 SF-14 n=23	RI-2 RI-4 n=22	SF-12 n=17	SF-2 n=11	YA-1 n=18	SF-4 n=17	SF-10 n=10	RO-1 n=11	AS-1 n=16	AS-16 n=12	NO-10 n=19	AS-5 n=25	AS-17 n=24	BA-1 n=25
A-06	99.60											0.40		
A-07	50.62			19.53	7.13	22.67								0.02
A-08		4.92		12.26		82.80								
A-09	89.65	0.08		10.22										
A-11				12.83	87.17									
A-12		10.81		18.86	24.03	12.22				23.67				
B-03	0.08	99.92												
A-18			70.65	13.01						16.33				
A-27				70.34							29.64			
A-34		9.04			73.69				15.79					1.28
B-06						99.94		0.03						
B-07	33.94				15.78	50.17				0.05				
B-02				11.23	46.93									41.85
A-01	47.35	0.04				52.49								
B-01	40.33					6.62				52.78	0.25			

Assignment score of the *C. eos* genotype for the 15 hybrid lineages detected to a given *C. eos* population (Additional file 1). Shaded indicated site(s) where hybrid lineage was detected. Bold number refers to a probability of 100 % that a given lineage belongs to a reference population

basins do not seem to explain population genetics structure of *C. eos* either. Although allele frequencies of most populations are significantly different, the differentiation is limited and only three populations from West-Qc are clearly distinct from the other populations. Such organization of the genetic diversity suggests that most of the *C. eos* populations likely dispersed through temporary hydrologic bridges from one single founding group after the ice-sheet retreated [13, 52].

Contrasting patterns in hybrid diversity

The hybrids sampled in this study are highly diverse with 41 distinct hybrid lineages. The number of stepwise mutations among genotypes is 14 times larger than the average number of pairwise difference within lineage. In addition, phylogenetic relationships among hybrids revealed that lineages do not cluster according to variants of the mitochondrial and PEG1/MEST loci. These results rule out the probability of derived clones from a unique hybridization event and suggest hybrid lineages have been produced through distinct hybridization events (e.g. [88–90]). This high diversity contrasts with previous studies on *Chrosomus eos-neogaeus* reporting a single or a few lineages within a given region [48, 50, 51].

The striking difference in the distribution of hybrid lineage diversities between the two regions likely reflects distinct scenarios in the origin and colonization of hybrid lineages in West-Qc and East-Qc. In West-Qc, the widespread distribution of the B-01 hybrid lineage leads us to conclude, as previously stated by Angers and Schlosser [50], that this lineage occurred during the Pleistocene period and has colonized the sites of West-Qc via pro-glacial temporary drainages. This pattern has been observed previously, including in the *Chrosomus eos-neogaeus* complex [50, 51] and is common in sperm-dependent unisexual hybrid species [53–55].

The narrow distribution of the other lineages specific to West-Qc (A-02, A-03 and A-04) is however more difficult to interpret in a context of postglacial colonization. If a postglacial colonization appears unlikely, these hybrid lineages are good candidates for the hypothesis of locally produced hybrids in West-Qc. These lineages have a narrow distribution similar to the one observed in the assemblages of hybrids of East-Qc, they are always in sympatry with the B-01 lineage and two of these additional lineages occurred in the same site while most of the West-Qc sites harbored only a single lineage.

Contrasting with the homogeneity of West-Qc, 15 distinct assemblages of hybrids composed of 36 distinct lineages are observed in East-Qc. Except for three lineages that can be found in assemblages distant by more than 7 km using straight line distance (15 km using

waterway distance), most assemblages harbor specific lineages and have a narrow distribution leading to highly significant positive spatial autocorrelation patterns. Assignment tests revealed a close genetic relationship between hybrids from East-Qc and sympatric *C. eos* populations as most of the hybrids tested were assigned to their respective *C. eos* population. In addition to the large diversity and the positive spatial autocorrelations detected in hybrids compared to the lack of organization in the genetic diversity of the paternal species *C. eos*, these results give weight to the hypothesis that multiple hybridization events between *C. eos* and *C. neogaeus* occurred locally in several sites in East-Qc. Given the high waterway distances among drainages, positive spatial autocorrelations detected at short distances using straight line as well as waterway distances also support the hypothesis that hybridization events occurred after disappearance of pro-glacial temporary bridges.

Post-glacial colonization alone cannot explain the observed distribution of hybrid lineages in both regions. The wide distribution of the B-01 lineage in West-Qc suggests a post-glacial colonization. However, hybridization events likely occurred *in situ*, in order to yield the current distribution of the alternative lineages. In East-Qc, the distribution of few lineages in different assemblages is likely the result of dispersal although numerous local hybridization events are responsible for the observed hybrid assemblages. *In situ* hybridization events appear to be rare in most sperm-dependent unisexual hybrids (but see [56]) and our results provide the first strong evidence of locally produced hybrid lineages during the Holocene in the *Chrosomus eos-neogaeus* system.

Diversity and landscape elements

In spite of a higher number of sites sampled in this survey compared to the one performed by Angers and Schlosser (2007), no additional lineages was detected in the lakes of West-Qc. Interestingly, all of the additional hybrid lineages detected in this survey were found in small streams. This leads us to associate the difference in diversity of the assemblages of hybrids between regions to the different geomorphologic characteristics of their respective hydrographic networks. West-Qc is characterized by thousands of lakes and ponds whereas a dendritic system of ephemeral and permanent streams is observed in East-Qc; lakes and ponds are nearly absent from this last region.

The very low diversity of the West-Qc region may be explained by the presence of the widespread B-01 lineage. The high abundance of this lineage in most of the lakes could be related to extensive postglacial colonization. Its presence in most of the lakes may prevent the persistence of other hybrid lineages in lakes as expected under the neutral theory of biodiversity [91],

or be the result of competitive exclusion of alternative lineages in lake environments.

Demographic decline of *Chrosomus neogaeus*

The results of this survey are consistent with previous studies reporting low abundance of the maternal species *Chrosomus neogaeus* [50]. In addition, the presence of *C. eos-neogaeus* × *eos* as triploid hybrids, never *C. eos-neogaeus* × *neogaeus*, indirectly confirmed the near absence of *C. neogaeus* in both regions. Gynogenetic *C. eos-neogaeus* require the sperm of either parental species to reproduce, the spermatozoid genome is occasionally incorporated into the diploid hybrid egg leading to the formation of triploid hybrids. These triploid hybrids are observed in high proportion in every hybrid populations but are only produced *de novo* [48]. The presence of an additional *C. eos* genome in triploid individuals indicates that hybrids use the sperm of the *C. eos* species, not *C. neogaeus* to trigger the development of their eggs.

The reasons for the scarcity of the *C. neogaeus* species in these regions remains to be investigated, however, multiple local hybridization events may be indicative. According to the Hubbs principle, hybridization is more likely to occur if one of the sympatric species becomes rare [92]. Hybridization events involving now extinct species have been reported in multiple species including fishes [93, 94]. Environmental changes following the end of the Pleistocene may have triggered the demographic decline of *C. neogaeus* populations, increasing the probability of interspecific mating by *C. neogaeus* females. Production of gynogenetic unisexual hybrids may have accelerated the disappearance of *C. neogaeus* populations by outcompeting them due to the alleviation of the “double cost of sex” in hybrids. While this scenario is hypothetical, it provides an explanation to the high diversity of hybrids as well as to the near lack of maternal species.

Conclusions

This study supported the hypothesis of multiple hybridization events between *Chrosomus eos* and *C. neogaeus* occurring during the Holocene in East-Qc. This explains the difference in the diversity and distribution of hybrid lineages between regions separated by the St Lawrence River in absence of marked differences among *C. eos* populations. This highlights distinct historical processes that act in shaping the diversity in the assemblage of hybrids in each region. This unusually high diversity of hybrid lineages in East-Qc coupled to the near absence of the maternal species *Chrosomus neogaeus* suggests that the decline of this species could be the trigger event at the origin of the high rates of spontaneous hybridization in this region.

Additional files

Additional file 1: Characteristics of the different sampled sites. List of sampled sites with their geographical coordinates and biotypes composition. E, N and H respectively refer to *Chrosomus eos*, *C. neogaeus* and the hybrids. Hybrid lineages of each site are identified. (PDF 73 kb)

Additional file 2: Multiloci genotype of the parental species *C. eos* individuals. Nuclear genotype of every *C. eos* individuals used to assess the paternal species genetic diversity and structure. For each individual, region and sampled sites are provided. (PDF 348 kb)

Additional file 3: Consensus nuclear genotype and mitochondrial haplotype of the different lineages. Nuclear genotype and mitochondrial haplotype of the different lineages. Allele size refers to the more common genotype and bold characters to variable loci within a given lineage. (PDF 46 kb)

Acknowledgments

We thank the anonymous reviewers for their constructive comments that improved the manuscript. The authors are extremely grateful to the numerous students who assist us in this survey and biologists from Quebec minister of natural resources and wildlife who provide us with fish from several sampling sites. CL was supported by *Fonds de recherche du Québec - Nature et Technologies* (FRQNT), J-A Bombardier and Étienne-Magnin scholarships. This research was supported by a research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to BA.

Availability of data and material

The datasets supporting the conclusions of this article are included within the article and its additional files.

Authors' contributions

RV and CL contributed equally to this work. Research was designed by RV, CL and BA. Sampling was conducted by CL and BA. Laboratory experiments and data analyses were conducted by RV and CL. All authors wrote and edited the manuscript draft.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

This research was performed under institutional animal care guidelines (permit #13-084 delivered by the *Université de Montréal*), and conforms to the mandatory guidelines of the Canadian Council on Animal Care. Sampling permits were provided by the Quebec Ministry of Natural Resources and Wildlife (MRNF).

Received: 22 February 2016 Accepted: 24 August 2016

Published online: 06 September 2016

References

- Dionne M, Caron F, Dodson JJ, Bernatchez L. Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Mol Ecol*. 2008;17(10):2382–96.
- Wofford JE, Gresswell RE, Banks MA. Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecol Appl*. 2005;15(2):628–37.
- Whiteley AR, Hastings K, Wenburg JK, Frissell CA, Martin JC, Allendorf FW. Genetic variation and effective population size in isolated populations of coastal cutthroat trout. *Conserv Genet*. 2010;11(5):1929–43.
- Coulon A, Guillot G, Cosson JF, Angibault J, Aulagnier S, Cargnelutti B, Galan M, Hewison A. Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Mol Ecol*. 2006;15(6):1669–79.
- Manel S, Schwartz MK, Luikart G, Taberlet P. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol*. 2003;18(4):189–97.
- Neville HM, Dunham JB, Peacock MM. Landscape attributes and life history variability shape genetic structure of trout populations in a stream network. *Landscape Ecol*. 2006;21(6):901–16.

7. Hewitt GM. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc.* 1996;58(3):247–76.
8. Bernatchez L, Wilson CC. Comparative phylogeography of Nearctic and Palearctic fishes. *Mol Ecol.* 1998;7(4):431–52.
9. Wilson AB, Veraguth IE. The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Mol Ecol.* 2010;19(20):4535–53.
10. Poissant J, Knight TW, Ferguson MM. Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Mol Ecol.* 2005;14(5):1321–31.
11. Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM. Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. *Mol Ecol.* 2009;18(12):2668–86.
12. Arbogast BS, Kenagy G. Comparative phylogeography as an integrative approach to historical biogeography. *J Biogeogr.* 2001;28(7):819–25.
13. April J, Hanner RH, Dion-Côté A-M, Bernatchez L. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Mol Ecol.* 2013;22(2):409–22.
14. Laurin-Lemay S, Angers B, Beney B, Brodeur J. Inconsistent genetic structure among members of a multitrophic system: did bruchid parasitoids (*Horismenus spp.*) escape the effects of bean domestication? *Bull Entomol Res.* 2013; 103(02):182–92.
15. Poulin R, Krasnov BR, Mouillot D. Host specificity in phylogenetic and geographic space. *Trends Parasitol.* 2011;27(8):355–61.
16. Cyr F, Peres-Neto P, Angers B, Tonn WM. Allopatric distribution of generalist parasites: interplay between postglacial dispersal and host species. *Can J Fish Aquat Sci.* 2012;69(9):1491–8.
17. Prugnolle F, Théron A, Pointier JP, Jabbar-Zahab R, Jarne P, Durand P, Meeùs T. Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. *Evolution.* 2005;59(2):296–303.
18. Dowling TE, Secor CL. The role of hybridization in diversification of animals. *Annu Rev Ecol Syst.* 1997;28:593.
19. Kearney M. Hybridization, glaciation and geographical parthenogenesis. *Trends Ecol Evol.* 2005;20(9):495–502.
20. Vrijenhoek RC. Genetic and ecological constraints on the origins and establishment of unisexual vertebrates. In: Dawley RM, Bogart JP, editors. *Evolution and ecology of unisexual vertebrates.* Albany: New York State Museum; 1989. p. 24–31. Bulletin 466.
21. Bengtsson BO. Asex and evolution: a very large-scale overview. In: *Lost Sex.* Edited by Schön I, Martens K, van Dijk P. Dordrecht, Netherlands: Springer; 2009:1–19.
22. Simon J-C, Delmotte F, Rispé C, Crease TJ. Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. *Biol J Linn Soc.* 2003;79(1):151–63.
23. Vrijenhoek R, Dawley R, Cole C, Bogart J. A list of the known unisexual vertebrates. In: Dawley RM, Bogart JP, editors. *Evolution and ecology of unisexual vertebrates.* Albany: New York State Museum; 1989. p. 19–23. Bulletin 466.
24. Beukeboom L, Vrijenhoek R. Evolutionary genetics and ecology of sperm-dependent parthenogenesis. *J Evol Biol.* 1998;11(6):755–82.
25. Lamatsch DK, Stöck M. Sperm-dependent parthenogenesis and hybridogenesis in teleost fishes. In: *Lost sex.* Edited by Schön I, Martens K, van Dijk P. Dordrecht, Netherlands: Springer; 2009:399–432.
26. Neaves WB, Baumann P. Unisexual reproduction among vertebrates. *Trends Genet.* 2011;27(3):81–8.
27. Quilodrán CS, Montoya-Burgos JI, Currat M. Modeling interspecific hybridization with genome exclusion to identify conservation actions: the case of native and invasive *Pelophylax* waterfrogs. *Evol Appl.* 2015;8:199–210.
28. Rhymer JM, Simberloff D. Extinction by hybridization and introgression. *Annu Rev Ecol Syst.* 1996;27:83–109.
29. Maynard Smith J. *The evolution of sex.* Cambridge: Cambridge Univ Press; 1978.
30. Schultz RJ. Special adaptive problems associated with unisexual fishes. *Am Zool.* 1971;11(2):351–60.
31. Schlupp I. The evolutionary ecology of gynogenesis. *Annu Rev Ecol Evol Syst.* 2005;36:399–417.
32. Schlupp I, Parzefall J, Scharl M. Biogeography of the Amazon molly, *Poecilia formosa* J Biogeogr. 2002;29(1):1–6.
33. Stebbins GL. Polyploidy, Hybridization, and the Invasion of New Habitats. *Ann Missouri Bot Gard.* 1985;72(4):824–32.
34. Lewontin RC, Birch LC. Hybridization as a source of variation for adaptation to new environments. *Evolution.* 1966;20(3):315–36.
35. Mavárez J, Linares M. Homoploid hybrid speciation in animals. *Mol Ecol.* 2008;17(19):4181–5.
36. Stebbins GL. The role of hybridization in evolution. *Proc Am Phil Soc.* 1959; 103(2):231–51.
37. Allendorf FW, Leary RF, Spruell P, Wenburg JK. The problems with hybrids: setting conservation guidelines. *Trends Ecol Evol.* 2001;16(11):613–22.
38. Seehausen O. Conservation: losing biodiversity by reverse speciation. *Curr Biol.* 2006;16(9):R334–7.
39. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science.* 2003; 301(5637):1211–6.
40. Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC. Homoploid hybrid speciation in an extreme habitat. *Science.* 2006;314(5807):1923–5.
41. Schwenk K, Brede N, Streit B. Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. *Phil Trans R Soc B.* 2008;363(1505):2805–11.
42. Stelkens R, Seehausen O. Genetic distance between species predicts novel trait expression in their hybrids. *Evolution.* 2009;63(4):884–97.
43. Yakimowski SB, Rieseberg LH. The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. *Am J Bot.* 2014;101(8):1247–58.
44. Seehausen O. Hybridization and adaptive radiation. *Trends Ecol Evol.* 2004; 19(4):198–207.
45. Otto SP, Whitton J. Polyploid incidence and evolution. *Annu Rev Genet.* 2000;34(1):401–37.
46. Mable B. Why polyploidy is rarer in animals than in plants: myths and mechanisms. *Biol J Linn Soc.* 2004;82(4):453–66.
47. Linder CR, Rieseberg LH. Reconstructing patterns of reticulate evolution in plants. *Am J Bot.* 2004;91(10):1700–8.
48. Goddard KA, Dawley RM, Dowling TE. Origin and genetic relationships of diploid, triploid, and diploid-triploid mosaic biotypes in the *Phoxinus eos-neogaeus* unisexual complex. In: RM D, JP B, editors. *Evolution and ecology of unisexual vertebrates.* New York: New York State Museum; 1989: 466, pp. 268–280. Bull.
49. Goddard K, Megwinoff O, Wessner L, Giaino F. Confirmation of gynogenesis in *Phoxinus eos-neogaeus* (Pisces: Cyprinidae). *J Hered.* 1998;89(2):151–7.
50. Angers B, Schlosser IJ. The origin of *Phoxinus eos-neogaeus* unisexual hybrids. *Mol Ecol.* 2007;16(21):4562–71.
51. Elder JF, Schlosser IJ. Extreme clonal uniformity of *Phoxinus eos-neogaeus* gynogens (Pisces: Cyprinidae) among variable habitats in northern Minnesota beaver ponds. *Proc Natl Acad Sci USA.* 1995;92(11):5001–5.
52. Gagnon M-C, Angers B. The determinant role of temporary proglacial drainages on the genetic structure of fishes. *Mol Ecol.* 2006;15(4):1051–65.
53. Janko K, Culling M, Rab P, Kotlik P. Ice age cloning - comparison of the Quaternary evolutionary histories of sexual and clonal forms of spiny loaches (Cobitis; Teleostei) using the analysis of mitochondrial DNA variation. *Mol Ecol.* 2005;14(10):2991–3004.
54. Cunha C, Coelho M, Carmona J, Doadrio I. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Mol Ecol.* 2004;13(9):2807–17.
55. Quattro JM, Avise JC, Vrijenhoek RC. An ancient clonal lineage in the fish genus *Poeciliopsis* (Atheriniformes: Poeciliidae). *Proc Natl Acad Sci USA.* 1992;89(1):348–52.
56. Quattro J, Avise J, Vrijenhoek R. Molecular evidence for multiple origins of hybridogenetic fish clones (Poeciliidae: Poeciliopsis). *Genetics.* 1991;127(2):391–8.
57. New JG. Hybridization between two cyprinids, *Chrosomus eos* and *Chrosomus neogaeus*. *Copeia.* 1962;147–152.
58. Binet M-C, Angers B. Genetic identification of members of the *Phoxinus eos-neogaeus* hybrid complex. *J Fish Biol.* 2005;67(4):1169–77.
59. Skalski GT, Grose MJ. Characterization of microsatellite loci in the creek chub (*Semotilus atromaculatus*). *Mol Ecol Notes.* 2006;6(4):1240–2.
60. Orita M, Suzuki Y, Sekiya T, Hayashi K. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics.* 1989;5(4):874–9.
61. Meirmans PG, Van Tienderen PH. genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes.* 2004; 4(4):792–4.
62. Legendre P, Cáceres M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol Lett.* 2013;16(8):951–63.

63. Chao A, Chazdon RL, Colwell RK, Shen TJ. Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics*. 2006;62(2):361–71.
64. Legendre P, Legendre L. *Numerical ecology*. 2nd ed. Amsterdam: Elsevier Science; 1998.
65. Oden NL, Sokal RR. Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Syst Biol*. 1986;35(4):608–17.
66. Boizard J, Magnan P, Angers B. Effects of dynamic landscape elements on fish dispersal: the example of creek chub (*Semotilus atromaculatus*). *Mol Ecol*. 2009;18(3):430–41.
67. Cyr F, Angers B. Historical process lead to false genetic signal of current connectivity among populations. *Genetica*. 2011;139(11-12):1417–28.
68. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6(2):65–70.
69. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara R, Simpson GL, Solymos P, Stevens MHH, Wagner H. *vegan*: Community Ecology Package. 2015. <http://CRAN.R-project.org/package=vegan>.
70. Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered*. 1995;86(6):485–6.
71. Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour*. 2008;8(1):103–6.
72. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered*. 1995;86(3):248–9.
73. Rice WR. Analyzing tables of statistical tests. *Evolution*. 1989;43(1):223–5.
74. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008;24(11):1403–5.
75. Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011;27(21):3070–1.
76. Weir B, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984;38:1358–1370.
77. Michalakis Y, Excoffier L. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*. 1996;142(3):1061–4.
78. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online*. 2005;1:47.
79. Rousset F. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*. 1996;142(4):1357–62.
80. Hardy OJ, Charbonnel N, Fréville H, Heuertz M. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*. 2003;163(4):1467–82.
81. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes*. 2002;2:618–20.
82. Cavalli-Sforza LL, Edwards AW. Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet*. 1967;19(3 Pt 1):233.
83. Edwards A. Distances between populations on the basis of gene frequencies. *Biometrics*. 1971;873–881.
84. Paetkau D, Calvert W, Stirling I, Strobeck C. Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol*. 1995;4(3):347–54.
85. Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*. 1999;153(4):1989–2000.
86. 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(6):536–9.
87. Fisher RA. *Statistical methods for research workers*. Edinburgh: Genesis Publishing Pvt. Oliver and Boyd Ltd; 1925.
88. Vollmer SV, Palumbi SR. Hybridization and the evolution of reef coral diversity. *Science*. 2002;296(5575):2023–5.
89. Delmotte F, Sabater-Muñoz B, Prunier-Leterme N, Latorre A, Sunnucks P, Rispe C, Simon J-C. Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species. *Evolution*. 2003;57(6):1291–303.
90. Vergilino R, Markova S, Ventura M, Manca M, Dufresne F. Reticulate evolution of the *Daphnia pulex* complex as revealed by nuclear markers. *Mol Ecol*. 2011; 20(6):1191–207.
91. Hubbell SP. *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ: Princeton University Press; 2001.
92. Hubbs CL. Hybridization between fish species in nature. *Syst Zool*. 1955; 4(1):1–20.
93. Wilson CC, Bernatchez L. The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Mol Ecol*. 1998;7(1):127–32.
94. Bernatchez L, Glémet H, Wilson CC, Danzmann RG. Introgression and fixation of Arctic char (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Can J Fish Aquat Sci*. 1995;52(1):179–85.
95. Gotelli NJ, Colwell RK. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett*. 2001;4(4):379–91.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

