doi: 10.1111/j.1420-9101.2008.01637.x

Divergent selection and phenotypic plasticity during incipient speciation in Lake Victoria cichlid fish

I. S. MAGALHAES,*,† S. MWAIKO,†,‡ M. V. SCHNEIDER† & O. SEEHAUSEN*,†

*Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of Bern, Bern, Switzerland
†Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Center of Ecology, Evolution and Biogeochemistry, Kastaninenbaum, Switzerland
‡Tanzanian Fisheries Research Institute, Mwanza Centre, Mwanza, Tanzania

Keywords:

cichlids;
divergent selection;
F_{ST};
phenotypic plasticity;
P_{ST};
Q_{ST}.

Abstract

Divergent selection acting on several different traits that cause multidimensional shifts are supposed to promote speciation, but the outcome of this process is highly dependent on the balance between the strength of selection vs. gene flow. Here, we studied a pair of sister species of Lake Victoria cichlids at a location where they hybridize and tested the hypothesis that divergent selection acting on several traits can maintain phenotypic differentiation despite gene flow. To explore the possible role of selection we tested for correlations between phenotypes and environment and compared phenotypic divergence (P_{ST}) with that based on neutral markers (F_{ST}). We found indications for disruptive selection acting on male breeding colour and divergent selection acting on several morphological traits. By performing common garden experiments we also separated the environmental and heritable components of divergence and found evidence for phenotypic plasticity in some morphological traits contributing to species differences.

Introduction

Understanding how biological diversity is generated and maintained is a central issue to both evolutionary and conservation biology. Evolutionary forces shaping genetic variation are mutation, migration, genetic drift and selection. Theory predicts that even low levels of gene flow between populations will prevent neutral divergence (Slatkin, 1987), but that divergent or disruptive selection can generate and maintain adaptive divergence despite gene flow (Dobzhansky, 1970; Dieckmann & Doebeli, 1999; Higashi *et al.*, 1999; Kondrashov & Kondrashov, 1999; Drossel & McKane, 2000; Gavrilets, 2004). An increasing number of empirical studies suggest that divergent selection can indeed lead to speciation with gene flow (Schluter, 1996a, b; Lu & Bernatchez, 1999; Rundle *et al.*, 2000; Boughman, 2001; Naisbit *et al.*, 2001; Schliewen *et al.*, 2001; Bolnick, 2004; Emelianov *et al.*, 2004; Rogers & Bernatchez, 2007; Grant & Grant, 2008). Under which conditions and how commonly divergent selection despite gene flow leads to complete speciation is a debated issue though (Nosil & Harmon, in press). Very recently it has been suggested that two different factors may explain variation in the completeness of ecological speciation: (1) the intensity of divergent selection on one trait or (2) the multifariousness of selection, i.e. the number of genetically independent traits that are under selection (Nosil & Sandoval, 2008; Nosil *et al.*, in press).

Utilizing this theoretical framework, we studied a pair of sister species of Lake Victoria cichlids at a location where they hybridize to investigate the potential role of disruptive or divergent selection in the maintenance of multivariate phenotypic differentiation despite gene flow. The sister species *Pundamilia pundamilia* and *Pundamilia nyererei* are widely and sympatrically distributed at rocky islands in Lake Victoria (Seehausen & van Alphen, 1999). The species differ primarily in male breeding coloration, which is used in their identification. Males of both species have blackish underparts and blackish vertical bars on the

Correspondence: I. S. Magalhaes, Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland.
Tel.: +41 413492146; fax: +41 413492168;
e-mail: isabel.magalhaes@eawag.ch
Present address: M. V. Schneider, The European Bioinformatics Institute, Welcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK.

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

flanks, but *P. nyererei* males have crimson red dorsum and dorsal fin and yellow flanks, whereas *P. pundamilia* males have blue-grey dorsum, a blue dorsal fin and blue-grey flanks (Seehausen, 1997). Females of both species are cryptically coloured and either more yellowish (*P. nyererei*) or more greyish (*P. pundamilia*). At places in the lake where the water transparency is high and the light environment heterogeneous, these phenotypes are ecologically and genetically differentiated sister species, divergent in several dimensions: male breeding colour (Seehausen *et al.*, 1997; Seehausen, in press), opsin genes that affect colour vision (Carleton *et al.*, 2005), depth distribution (Seehausen, 1997; Seehausen *et al.*, 1998), photic environment (Maan *et al.*, 2006) and feeding ecology (Bouton *et al.*, 1997). However, at other places where the water transparency is low, intermediate colour phenotypes are found or even dominate the population (Seehausen *et al.*, 1997; Seehausen, in press). At these places, the two phenotypes are merely extremes in a single panmictic population. A recent study analysed these phenotypes at five different islands along five light gradients of different steepness (Seehausen *et al.*, 2008). The study found that in populations living on moderately shallow to moderately steep light gradients there were strong correlations between ambient light colour, male nuptial colour, visual pigments and female mating preference for male nuptial colour, which provided evidence for speciation through sensory drive. The correlations become weaker and eventually disappear altogether as the light gradient becomes steeper. Along this turbidity transect the two species go from being one species with phenotypic variation to two incipient species, and finally two sister species, in what has been called a speciation or a 'speciation in reverse' transect (Seehausen *et al.*, 1997, 2008).

Here, we analysed in detail the incipient species from the speciation transect that showed the weakest yet significant bimodality in the distribution of male nuptial coloration and the lowest significant differentiation at neutral loci (Seehausen *et al.*, 2008) and investigated traits not studied so far in this population. We investigated neutral genetic differentiation for two sampling years, analysed morphological and colour variation for 1 year and investigated if differences found among wild fish are heritable or phenotypically induced. Specifically, we ask the following questions: (1) is divergence in male coloration associated with differences in water depth? (2) Is there any evidence for differences in morphology along water depth and between colour phenotypes? (3) Is divergence in quantitative phenotypic traits larger than expected from divergence at neutral loci and can we invoke divergent or disruptive selection? (4) Are the phenotypic differences heritable or are they environmentally induced? Finding eco-morphological divergence between the two colour phenotype classes despite low genetic differentiation would support the hypothesis that divergent selection acting on several independent traits is sufficiently strong to maintain differences in the presence of gene flow.

We first investigated associations between water depth, male coloration and morphology. Then we compared differentiation in phenotypic traits as defined by P_{ST}, the phenotypic surrogate for Q_{ST} (Spitze, 1993), with differentiation at neutral marker genes (F_{ST}; Wright, 1951) between groups of male defined by nuptial colour or water depth. Finally, we performed a common garden experiment to determine whether: (1) phenotypic differentiation could be explained by environmentally induced plasticity and (2) phenotypic variation could be explained by segregation of alleles with complete dominance at a single locus.

Materials and methods

Study area and sampling

We collected samples at Kissenda Island, located in the western part of the Mwanza Gulf, in southern Lake Victoria, Tanzania. The water around the island is turbid, with an average secchi depth of 80 cm. With increasing water depth, the ambient light spectrum around this island shifts towards longer wave lengths, resulting in an increasingly red-shifted environment (Seehausen *et al.*, 2008). Phenotypically distinctly differentiated red and blue phenotypes are common around this island, with intermediate forms occurring at lower frequencies.

Using hook and line and gill nets, we randomly collected males in 2 years, 2000 and 2005. We took fin clips for genetic analysis in both years and collected whole fish, took photographs and recorded water depth at which individuals were collected in 2005. We collected 59 males in 2000 and 140 males in 2005. Water depth at which individuals were collected was recorded by measuring the length of the angling line (vertically down into rocky crevices) or by recording the depth of the sinker line of gill nets. Immediately after capture males were photographed in a standard photo cuvette for colour analysis. Fin clips were taken for genetic analysis and after sedating the fish on MS222, they were fixed in 4% formaldehyde solution. After transfer to Switzerland the fish were transferred in three steps to 30%, 50% and 70% ethanol for storage.

In 2005, we also collected live fish, which were transferred to aquaria in Switzerland for the common garden experiment.

Phenotypic analysis

Colour analysis

For photography, males were placed in a perpex cuvette with water immediately upon capture, and squeezed gently between a grey PVC sheet and the front window. A total of 118 wild males were photographed using a digital camera. Individuals with no picture taken and

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

those that did not show full nuptial coloration were excluded from further analysis. Laboratory-bred individuals were photographed in the holding tanks in full nuptial coloration. Once the most colourful, dominant male was photographed, it was taken out of the holding tank so that another male would become dominant and colour up fully. Males were assigned to a colour class on a five-point scale (Fig. 1a) (Dijkstra *et al.*, 2006; van der Sluijs *et al.*, 2008). The photographs were scored independently by two different observers. The correlation of the phenotype scoring of the two observers was calculated and the average of both scores was used for each fish in further analysis.

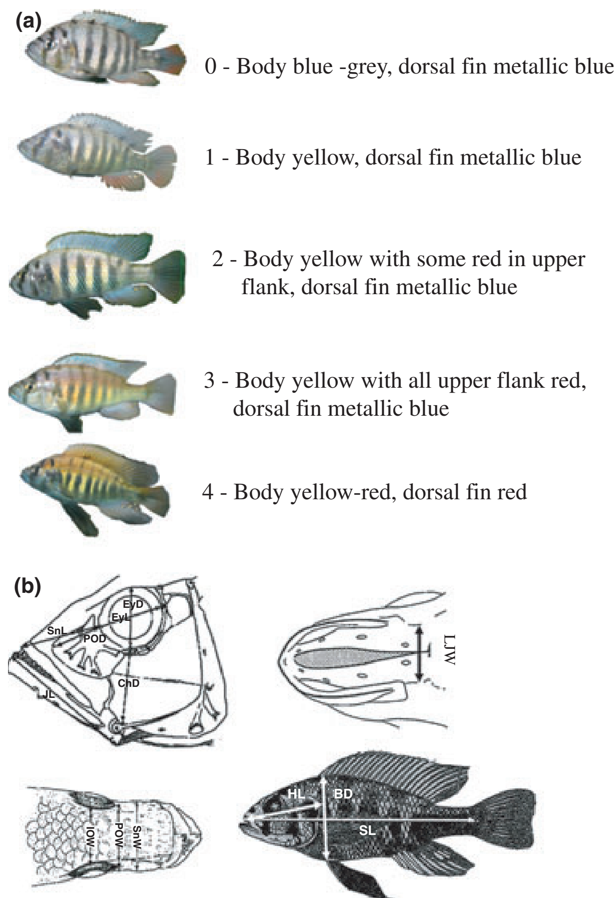

Fig. 1 (a) Scale of male nuptial colour: 0 = blue or grey flanks with spiny part of dorsal fin blue, 1 = a yellow flank but no red, spiny part of dorsal fin is blue, 2 = yellow flank with some red on the flank along the upper lateral line, spiny dorsal fin is blue, 3 = yellow flank with a partially red dorsum upwards from the upper lateral line, but a grey body crest and largely blue spiny dorsal fin, 4 = yellow flank with a completely red dorsum between the upper lateral line and the body crest, red spiny dorsal fin (after van der Sluijs *et al.*, 2008); (b) eco-morphological distances taken by a digital caliper (adapted from Barel *et al.*, 1977): standard length (SL), head length (HL), head width (HW), body depth (BD), lower jaw length (LJL), lower jaw width (LJW), snout length (SnL), snout width (SnW), cheek depth (ChD), pre-orbital depth (PoD), inter-orbital width (IoW), eye length (EyL) and eye depth (EyD).

Morphological analysis

Using a digital caliper, we took 13 standard morphometric distances that are powerful in detecting eco-morphological variation among haplochromine cichlids (Barel *et al.*, 1977) (Fig. 1b). We used the univariate residual method to adjust each morphometric distance for size heterogeneity among individuals (Fleming & Gross, 1994): raw data were log transformed and used to establish regression lines describing the relationship between each character and standard length. Standardized residuals from these regressions were used as variables for further analyses. Multivariate morphological variation among individuals was reduced to principal components using SPSS 14.0 (SPSS Inc., Chicago, IL, USA). The relative percentage of the variance in each variable that loaded to each PC axis was estimated by first squaring the component loadings, then multiplying them by the percentage of variance explained by an axis and finally dividing this value by the sum of the percentages of variance in a trait that load to the first three PC axes.

Statistical analysis

As visual representation frequency histograms were plotted for the three PCs of eco-morphological variation and for male nuptial colour scores. Using SPSS we tested if principal components one, two, three and colour scores had a normal distribution and fitted alternative regression models to the frequency histograms. Additionally, we also used DISCMIXTUREPROGS v0.4 (Brewer, 2003) and estimated Akaikes information criteria (AIC) corrected for sample size (AIC_c; Burnham & Anderson, 2002), which tested statistically whether each sample was better represented by a single normal distribution or by a mixture of two normal distributions. We then estimated AIC_c differences (ΔAIC_c) as AIC_c for the single normal distribution minus AIC_c for the fitted mixture of two distributions. We followed established guidelines for the interpretation of ΔAIC_c (Burnham & Anderson, 2002) and used the interpretation of Hendry *et al.* (2006): $\Delta AIC_c < -8$ is considered as strong support for a single normal distribution, $-8 < \Delta AIC_c < -5$ as moderate support for a single normal distribution, $-5 < \Delta AIC_c < 5$ as roughly equivalent support for a single normal distribution or a mixture of two normal distributions, $5 < \Delta AIC_c < 8$ as moderate support for a mixture of two normal distributions and $\Delta AIC_c > 8$ is considered as strong support for a mixture of two normal distributions.

DNA extraction and microsatellite amplification

DNA was extracted from fin tissue of 59 individuals collected in 2000 and of 107 individuals collected in 2005 using a QIAGEN® (Basel, Switzerland) Biosprint 96™ extraction robot with the corresponding standard

extraction method. We analysed 10 microsatellite loci developed for these species (Ppun 5, Ppun7, Ppun17 and Ppun32; Taylor *et al.*, 2002) and for other African cichlid species (OSU20d, OSU19T and OSU16d: Wu *et al.*, 1999; TmoM5: Zardoya *et al.*, 1996 and Pzeb3 and Pzeb5: Van Oppen *et al.*, 1997). Two different sets of markers were used for multiplexing, avoiding overlapping allele ranges. The first set included loci Ppun5, Ppun7, Ppun17, Ppun32, TmoM5 and Pzeb5 and the second set included loci OSU20d, OSU19T, OSU16d and Pzeb3. We used the QIAGEN Multiplex PCR kit for PCR amplification according to the manufacturer's protocol and PCR program. A quantity of 1 μL of a 1 : 2 dilution of the PCR was added to a volume of sample loading solution (deionized formamide) and 400-bp DNA size standard for analysis of fragments up to 400 nucleotides (Beckman coulter, Fullerton, CA, USA). Denaturated fragments were resolved on an automated DNA sequencer (CEQ 8000; Beckman coulter). We then determined genotypes manually.

Estimation of divergence

Tests for phenotypic and genetic differentiation of individuals collected in 2005 were performed dividing the data in two different ways. In the first analysis we divided the entire sample by catching depth, creating two depth classes, from 0 to 2 m and below 2 m of depth. The 2-m breakpoint divided the sample into two groups of similar size. The group of individuals caught between 0 and 2 m and the group of individuals caught below 2 m of depth will be referred to as 'shallow' and 'deep' respectively. Second, we divided the sample by male nuptial colour (colour classes 0 + 1 vs. 3 + 4). This division represented the coloration of the two sister species at islands where they are well differentiated. Therefore, groups of colour classes 0 + 1 and 3 + 4 will be referred to as 'blue' and 'red' respectively. The single male of colour class 2 in our sample was excluded from this analysis.

Information on depth distribution of individuals collected in 2000 was not available; so, tests of genetic differentiation and divergence of these individuals were only performed for when samples were divided by male nuptial colour. Therefore, groups will also be referred to as 'blue' and 'red' respectively.

Neutral genetic variation and divergence

CONVERT (Glaubitz, 2004) was used to create input files for other programs and to create allele frequency tables. FSTAT was used to calculate allelic richness at each locus (Goudet, 1995). We used ARLEQUIN 3.11 (Excoffier *et al.*, 2005) to test each locus in each population for departure from Hardy–Weinberg equilibrium (HWE), to calculate observed (H_O) and expected (H_E) heterozygosities and to perform tests for genotypic linkage disequilibrium. Statistical significance in the above tests was adjusted for multiple comparisons using sequential Bonferroni adjustments (Rice, 1989). Multilocus and single locus F_{ST} estimates (Wright, 1951) were calculated over 1000 permutations, as implemented in ARLEQUIN 3.11 (Excoffier *et al.*, 2005). We used the allele frequency-based program DOH (http://www.biology.ualberta.ca/jbrzusto/Doh.php) to calculate genetic assignment for each phenotypically assigned individual (Paetkau *et al.*, 1995). Significance of deviation from random assignments was assessed using 2 × 2 contingency tests.

Model-based assignment tests were performed too, as implemented in the computer program STRUCTURE 2.2 (Pritchard *et al.*, 2000). MCMC simulations were run with 500 000 replicates and a burn-in of 50 000 replicates for K (number of populations) = 1–4 and applying the admixture model, in which individuals may share portions of the genome assigned to more than one population as a result of mixed ancestry (Pritchard *et al.*, 2000). Finally, we performed a hierarchical AMOVA, implemented in ARLEQUIN (Excoffier *et al.*, 2005) to investigate how genetic variability was distributed between species and years. Phenotypes were nested inside years.

Phenotypic divergence

We used independent samples *t*-tests to compare mean PC scores between depth groups and male colour groups and to compare mean male coloration between depth groups. We performed a discriminant function analysis to test for differences between groups. We also performed analyses of variance to estimate the divergence in phenotypic traits and to obtain the components of phenotypic variances for the calculation of the P_{ST} values. P_{ST} values are equivalent to Q_{ST} values (Spitze, 1993) but can be influenced by environmental and nonadditive genetic effects.

The theory that population structure can be inferred by comparing the genetic differentiation of presumably selected traits with that of presumably neutral molecular markers (Wright, 1951; Lande, 1992; Spitze, 1993) predicts that the genetic differentiation of a neutral polygenic trait is identical to the genetic differentiation of a single, neutral locus. Hence, the genetic differentiation of neutral markers represents the null hypothesis for the expected amount of differentiation of quantitative traits due to migration or drift alone. If for a given trait the value of Q_{ST}, or in its absence its surrogate $P_{ST,}$ significantly exceeds F_{ST}, divergent selection is invoked, whereas similar values are consistent with neutral evolution and significantly smaller values of Q_{ST} than F_{ST} suggest purifying selection (Merilä & Crnokrak, 2001; McKay & Latta, 2002). The approach relies on a number of assumptions, the most important being that phenotypic variance is mostly additive genetic variance. Several recent studies have pointed out that differences in F_{ST} and Q_{ST} can arise due to biased estimations of differentiation, which can occur for both quantitative traits and molecular markers (Hendry, 2002; O'Hara & Merilä, 2005; Goudet & Buchi, 2006; Leinonen *et al.*, 2008; Whitlock, 2008). However, although we are aware of its constraints and caveats, the $Q_{ST}/P_{ST} - F_{ST}$

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

comparison is a very useful method and perhaps the best currently available, for the inference of selection on phenotypes, other than by experimentation.

P_{ST} values were quantified as the proportion of variance in quantitative traits attributable to differences among populations

$$P_{ST} = \frac{\sigma^2_{gb}}{\sigma^2_{gb} + 2h^2\sigma^2_{gw}},$$

where σ^2_{gb} and σ^2_{gw} are the among-population and within-population variance components for a phenotypic trait, respectively, and h^2 is the heritability. As commonly performed on studies on morphological traits (Merilä, 1997; Leinonen *et al.*, 2006), we have assumed a heritability of 0.5, which means that environmental and nonadditive effects each account for half of the phenotypic variation. Because of the nonindependence of the morphological structures and the associated morphometric distances taken, P_{ST} values for morphology were estimated from principal components. Confidence intervals were estimated by bootstrapping over individuals, using code in R (http://www.r-project.org/). P_{ST} values were contrasted with multilocus F_{ST} values and considered significantly different when their 95% confidence intervals did not overlap. Additionally, we used an individual-based approach to test for a correlation between neutral genetic variation and morphological divergence. We computed morphological divergence between individuals as multidimensional Euclidean distances based on the residuals of the morphological measurements. Matrices of genetic distances ($\hat{a}$) (Rousset, 2000) were computed using the program SPAGeDI (Hardy & Vekemans, 2002). The relationship between the matrices was then tested using a Mantel test based on 1000 permutations implemented in ARLEQUIN (Excoffier *et al.*, 2005).

The distribution of phenotypic and genetic variation over water depth

We tested for an effect of water depth on the distribution of male colour phenotypes and on morphology of individuals collected in 2005 by performing Spearman rank correlations between depth and the five colour classes and between depth and principal components one, two and three of morphology.

To test for an effect of water depth on neutral genetic variation and morphological divergence we used an individual-based approach equivalent to testing for isolation by distance. Depth distances between individuals were computed with the program SPAGeDI (Hardy & Vekemans, 2002). We tested the relationships between depth distances and genetic distances ($\hat{a}$) (Rousset, 2000) and between depth distances and multidimensional Euclidean distances with Mantel tests based on 1000 permutations as implemented in ARLEQUIN (Excoffier *et al.*, 2005).

Common garden experiment and analysis of laboratory-bred individuals

Individuals of the parental generation were collected at Kisssenda Island in November 2005. Individuals were assigned to either incipient species on the basis of phenotype, and were kept in separate tanks at 24 ± 1 °C and 12 L : 12 D cycle. All fish were fed twice a day with dry commercial cichlid pellets in the morning and fresh shrimps and peas in the afternoon. Males and females within each incipient species population were allowed to interact and mate freely. During the first 2 months of 2006 six F1 families were produced. Families were raised in separate tanks. *Pundamilia* are female mouth brooders and we removed brooding females with their broods from the breeding tank. Each brood had a different wild-caught mother. To assign broods to a sire, we assigned paternity by genotyping fry, their mother, and all possible sires at five microsatellite loci. We used one dinucleotide (Ppun32) and four tetranucleotide (Ppun5, Ppun7, Ppun17 and Ppun21) loci, with specific primers designed for *P. pundamilia* and *P. neyererei* (Taylor *et al.*, 2002). Each brood had a different wild-caught father. Hence, every brood was a different and unrelated family. Ten *P. nyererei* males from three families and 25 *P. pundamilia* males from three families were raised to maturity (at least 12 months of age). Parents and F1 males from families of *P. pundamilia* and *P. nyererei* will be referred to as 'blue' and 'red' respectively.

Parental individuals were not available for phenotypic analysis because we could not preserve them before they showed strong phenotypic effects of aging. To detect possible family effects on the phenotypic variation observed among the experimental families, a nested ANOVA was carried out [two factors: family (six families) nested in species (*P. pundamilia* and *P. nyererei*)]. To separate environmental from heritable components of phenotypic variance between the incipient species, we compared phenotypic variance between males of 'blue' and 'red' raised in a common garden environment with that observed between the males of these populations in the wild. We used these data to address the question if adaptive phenotypic plasticity contributes to the divergence found between the wild populations.

To test if differences between the P_{ST} values of wild and common garden experimental fishes could be explained by sampling error, we resampled 50 times from the two wild populations to the sample size of the parents of the common garden experimental fish (i.e. six individuals in each species). We made pairwise comparisons for the first three principal components between the resampled populations (25 of each species) obtaining 625 P_{ST} values for each trait. We then compared the P_{ST} value from our common garden experiment with the distribution of P_{ST} values from wild fish. To see if the variance observed within the common garden populations was due to accidental crossing of the incipient species, we compared

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

the variances of the wild populations with the variance of the common garden populations. An increase in variance of laboratory-bred populations might suggest mixing of the incipient species during breeding. If phenotypic differences observed in the wild have a genetic basis, differences should be maintained when individuals are raised in a common environment. On the other hand, if phenotypic plasticity was the cause behind the observed differences, we expected that differences would collapse when fish were raised in a common environment.

Results

Phenotypic variation

Male nuptial coloration

We assigned 31 individuals to colour class 0, 26 to class 1, one individual to class 2, nine to class 3 and 40 individuals to class 4. Strong bimodality in male nuptial coloration of the wild population was evident from obvious discontinuity in the frequency histogram of colour scores (Fig. 2a). A cubic regression made a very good ($r^2 = 0.945$), yet not significant fit to the data ($P = 0.296$). A ΔAIC_c value of 120.47 ($\Delta AIC_c = 357.83 - 237.36$) was considered as strong support for a mixture of two normal distributions ($\Delta AIC_c > 8$).

Morphology

The principal component analysis identified three major axes of morphological variation (axes that explain more than 10% of the variation) (Table S1). Head length, lower jaw length, snout length, snout width and cheek depth all had heavy positive loadings on PC1, eye length and depth had heavy negative loadings on PC2 and body depth had a heavy positive loading on PC3. When comparing the percentage of variance in each trait that load to each PC, seven of 12 traits had more than 80% of their variance loading to PC1. Pre-orbital depth was the only trait whose variance loaded heavily to PC2 (79%) and body depth was the only trait with its variance loading mostly to PC3 (58%). Eye length and eye depth had both close to 50% of their variance loading to PC1 and to PC2.

Frequency histograms of variation in PC1, PC2 and PC3 showed weak or no discontinuities (Fig. 2b–d). Variation in all three axes was normally distributed (Kolmogorov–Smirnov tests, $P > 0.05$). ΔAIC_c values of 1 (192.6 − 191.6) for PC1 and −1.5 (188.7 − 190.2) for PC2 fitted equally well for single normal distributions and mixtures of two normal distributions ($-5 < \Delta AIC_c < 5$). A ΔAIC_c of −5.4 for PC3 was moderate support for a single normal distribution.

Genetic and phenotypic divergence

Neutral genetic variation and divergence

Most microsatellite markers exhibited a high degree of polymorphism (Table S2). A total of 210 alleles were found in the 166 individuals, ranging from 2 to 44 per locus.

For the sampling year of 2000, there was one locus deviating from HWE in the population 'red' and five loci deviated from HWE in the population 'blue'. Two of 45 tests of LD were significant in the population 'red' and 20 were significant in the population 'blue'. After sequential Bonferroni correction three loci were in Hardy–Weinberg

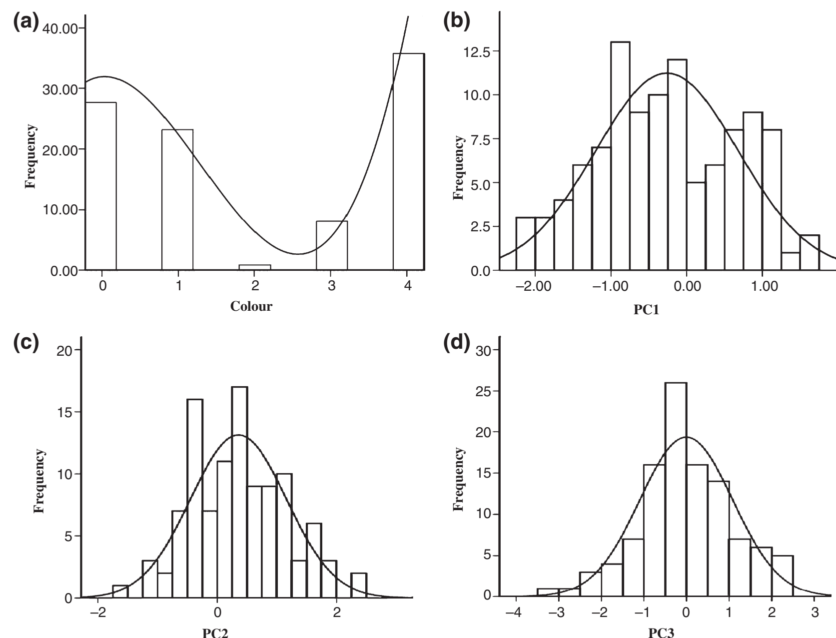

Fig. 2 Frequency distributions of (a) colour scores of wild-caught individuals with a cubic regression fitted to it, (b) principal component 1, (c) principal component 2 and (d) principal component 3 of wild-caught individuals with a normal distribution fitted to it.

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

disequilibrium and four tests showed linkage disequilibrium, all in the population 'blue'.

For the sampling year of 2005, before sequential Bonferroni correction, deviations from HWE were observed at one locus in one population, 'red'. One test out of 45 showed significant linkage disequilibrium in the same population. When populations were divided by depth, significant departures from HWE were observed at two loci and only in the population 'deep'. Now five tests showed linkage disequilibrium in the population 'deep'. After sequential Bonferroni correction no significant departures from HWE or linkage equilibrium were detected for the sampling year of 2005.

Differentiation at neutral genetic markers between depth classes was low ($F_{ST} = 0.001$) and not significant ($P = 0.25$). It was low between nuptial colour classes too but it was significantly different from zero, both for the year 2000 ($F_{ST} = 0.011$, $P = 0.01$) and for the year 2005 ($F_{ST} = 0.009$, $P = 0.008$). Single locus F_{ST} values had a very narrow distribution and few were significant, but single locus F_{ST} values were generally higher between nuptial colour classes than between depth ranges (Table 1). The assignment tests corroborated the F_{ST} values. Genetic assignment to classes defined by colour was significantly different from random, but assignment to classes defined by water depth was not (Fig. 3).

STRUCTURE 2.2 (Pritchard *et al.*, 2000) identified the most likely number of populations (K) in our sample of two incipient species and two years as being 1 (estimated −ln probability of data = −7632.3; $P > 0.99$). Nonetheless, plots for $K = 2$ showed clear differences between individuals phenotypically assigned to 'blue' and those phenotypically assigned to 'red', but no differences between individuals of the same phenotype collected in different years (Fig. 4). The hierarchical analysis of molecular variance (AMOVA) confirmed this result: there were significant differences between phenotypes within years ($P = 0.00$) but no differences between years ($P = 0.65$).

Table 1 Pairwise single locus and multilocus F_{ST} statistics between groups made based on depth and colour for the sampling year of 2005 and by colour for the year 2000.

Locus	Depth (2005)	Colour (2005)	Colour (2000)
Ppun5	0.001	0.011	0.003
Ppun7	0.000	−0.001	0.004
Ppun17	0.006	**0.016**	−0.002
Ppun32	−0.007	0.004	−0.008
OSU16d	0.003	**0.008**	0.011
OSU19t	**0.008**	0.020	**0.015**
OSU20d	0.022	**0.009**	**0.040**
Pzeb3	−0.004	0.019	0.000
Pzeb5	−0.006	−0.005	0.027
TmoM5	−0.001	0.004	0.001
multilocus	0.004	**0.009**	**0.011**

Significant F_{ST} values ($P < 0.05$) are shown in bold face.

Phenotypic divergence

Independently of how groups were formed mean PC scores and male nuptial coloration were all significantly different (Table 2). However, differentiation along morphometrics PC1 and PC2 was weaker when individuals were grouped by water depth than when they were grouped by male nuptial coloration (Fig. 3). The discriminant function analyses confirmed these results: groups were significantly different, both based on depth range (Wilk's $\lambda = 0.763$, $P = 0.02$) and based on male nuptial colour (Wilk's $\lambda = 0.303$, $P = 0.00$). However, 91.5% of the originally grouped individuals were correctly classified when individuals were grouped by colour, whereas the same figure was 77.6% when individuals were grouped by depth. The most important variables contributing to the differences between groups were head length, snout length and lower jaw length.

Phenotypic differentiation (P_{ST}) both between depth groups and between colour groups was higher than the neutral genetic marker differentiation (Table 3, Fig. 5). Confidence intervals, however, were large for all traits. When analysis was performed by depth group (Fig. 5a), male nuptial coloration had the highest P_{ST} value ($P_{ST} = 0.208$), followed by eco-morphometric PC1 ($P_{ST} = 0.135$), PC2 ($P_{ST} = 0.113$) and PC3 ($P_{ST} = 0.049$). The confidence intervals for male coloration, PC1 and PC2 laid outside the confidence interval of the multilocus microsatellite-derived F_{ST} value and were also higher than the highest of the 10 single locus F_{ST} values, indicating that differentiation along these axes of phenotypic variation was higher than expected by drift alone. Differentiation in eco-morphometric PC3 was close to the expectation from drift alone.

All P_{ST} values were larger when populations were grouped by male nuptial coloration, but the sequence of relative magnitudes was unchanged (Fig. 5b). It is not surprising that P_{ST} in male nuptial coloration was close to 1, because this trait was used to make the groups. Nonetheless, we show its P_{ST} for comparative purposes with the common garden populations. Among the morphometric PCs, PC1 had again the highest value ($P_{ST} = 0.375$), followed by PC2 ($P_{ST} = 0.183$) and PC3 ($P_{ST} = 0.156$). All confidence intervals laid outside those of the multilocus F_{ST} value and were higher than the highest single locus F_{ST} value, indicating that differentiation along several axes of morphological variation was higher than expected by drift alone.

Variation in neutral loci had no effect on morphological structure. Pairwise genetic distances among individuals did not significantly correlate with morphological distances ($r = 0.046$, $P = 0.12$).

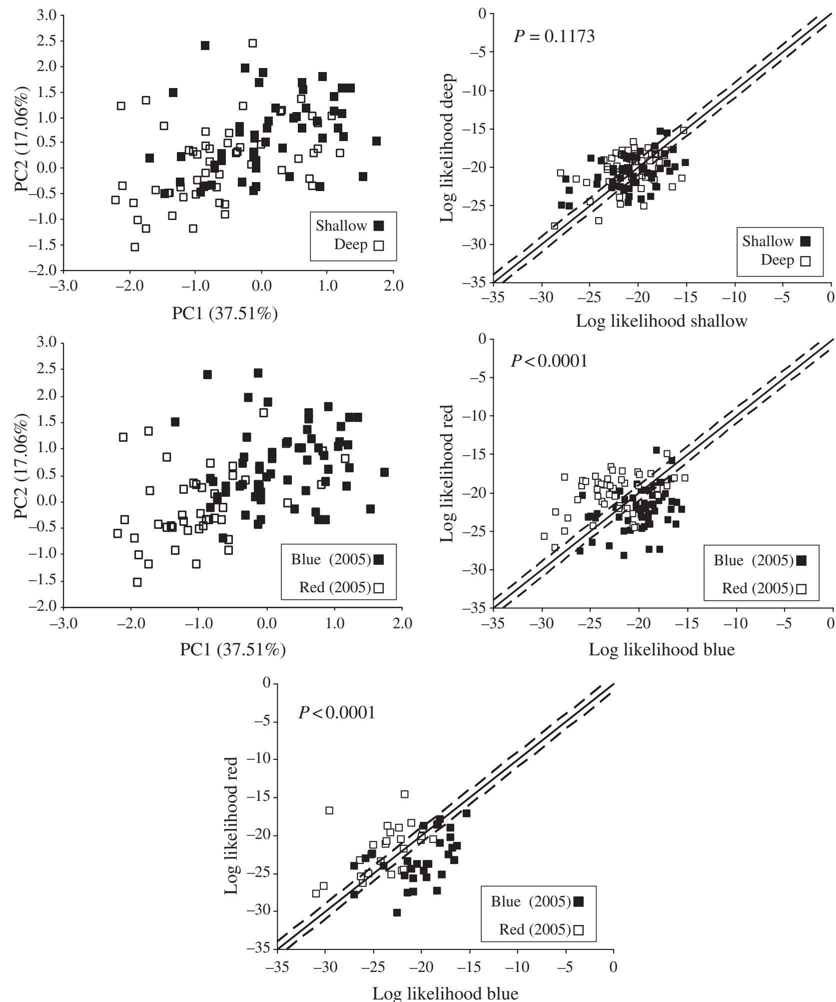

Fig. 3 Panels on the left: principal component analysis performed on morphometric distances measured in wild males, shown are PC1 and PC2 for individuals assigned to two groups; panels on the right: plots of likelihood of assignment tests showing the likelihoods of individuals belonging to each of the groups compared. From top to bottom: groups based on depth distribution from year 2005 and on colour from year 2005 and from 2000. *P*-values shown on the plots are the probabilities that the observed assignments were the result of chance. The solid diagonal line indicates the equal probability of being assigned to one population or the other. Points outside the dotted lines are assigned to one population with more than 10 times the likelihood as the other one.

The distribution of phenotypic and genetic variation over water depth

Male coloration significantly correlated with water depth (Fig. 6a). Individuals with bluish coloration were caught at shallower depth, whereas individuals with reddish nuptial coloration were distributed over the entire depth range. principal components one and two in eco-morphology showed a significant negative correlation with water depth, whereas PC3 showed a significant positive correlation (Fig. 6b–d). Individuals with lower PC1 and PC2 scores and higher PC3 scores tended to be from the lower end of the water depth range.

Variation in water depth clearly caused genetic and morphological structure. Pairwise genetic distances among individuals significantly correlated with the depth differential between them ($r = 0.190$, $P < 0.001$) as did morphological distances ($r = 0.219$, $P < 0.001$). However, depth differential between individuals only explained 3.5% of the genetic distance and 4.7% of the morphological distance.

Common garden experiments

Family did not explain any significant fraction of the variance in either colour ($F_{4,29} = 0.726$, $P = 0.58$) or morphological PC1 ($F_{4,29} = 0.24$, $P = 0.91$) but had significant effects on morphological PC2 ($F_{1,4} = 2.78$, $P = 0.05$) and PC3 ($F_{4,29} = 3.63$, $P = 0.02$). Parental phenotype class, on the other hand, was the main source of variation in male nuptial colour ($F_{1,29} = 28.78$, $P = 0.00$) and in morphological PC1 ($F_{1,29} = 12.99$, $P = 0.00$) and of some of the variation in PC3 ($F_{1,29} = 7.37$, $P = 0.01$). That neither parental phenotype class nor family explained much of the variation in PC2 indicates low heritability (Table S3).

The coloration of males bred from 'blue' parents ranged from zero to three on the colour scale, with most individuals of types zero, one or two. Males bred from 'red' parents ranged from one to four on the colour scale, with most individuals of type four (Fig. 7a). There was considerable overlap between laboratory-bred families of the two incipient species on morphology PC axis 1.

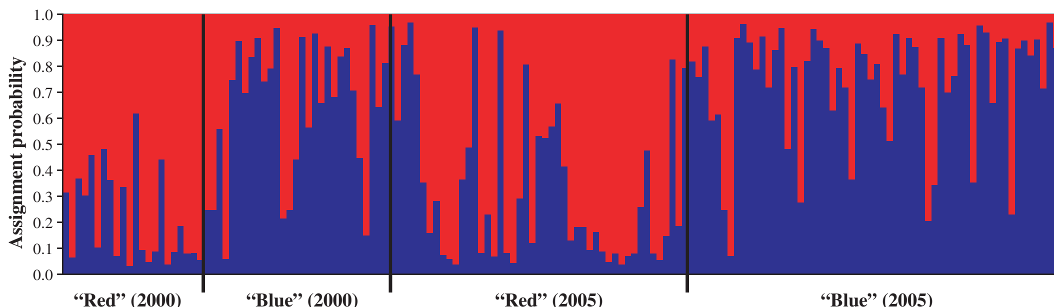

Fig. 4 Bayesian assignment probabilities of individuals to lineages estimated using STRUCTURE 2.2. Each bar represents an individual and colours indicate the proportion of an individual's genotype assigned to a particular population. Assignment proportions differ between the phenotype classes but not between sampling years.

Table 2 *t*-Tests comparing mean PC scores of groups based on male coloration and comparing mean PC scores and male coloration of groups based on depth range.

	t	d.f.	Sig. (two-tailed)	Mean difference	SE	2.50%	97.50%
Colour							
PC1	8.052	104	0.000	1.175	0.146	0.885	1.464
PC2	5.143	104	0.000	0.730	0.142	0.448	1.011
PC3	−4.37	104	0.000	−0.867	0.198	−1.260	−0.474
Depth							
PC1	4.003	104	0.000	0.685	0.171	0.346	1.025
PC2	3.567	104	0.001	0.529	0.148	0.235	0.824
PC3	−2.319	104	0.022	−0.483	0.208	−0.896	−0.070
Colour	−5.584	104	0.000	−1.601	0.287	−2.170	−1.033

Table 3 P_{ST} and F_{ST} values (their 95% confidence intervals) for the wild-caught populations divided by depth and coloration and for individuals from the common garden experiment.

	P_{ST}				
	Colour	PC1	PC2	PC3	F_{ST}
Depth	0.208 (0.087–0.377)	0.135 (0.038–0.27)	0.113 (0.024–0.241)	0.049 (0.001–0.146)	0.001 (0.000–0.003)
Colour	–	0.375 (0.240–0.514)	0.183 (0.072–0.321)	0.156 (0.055–0.303)	0.009 (0.004–0.013)
Laboratory bred	0.525	0.323	0.001	0.049	–

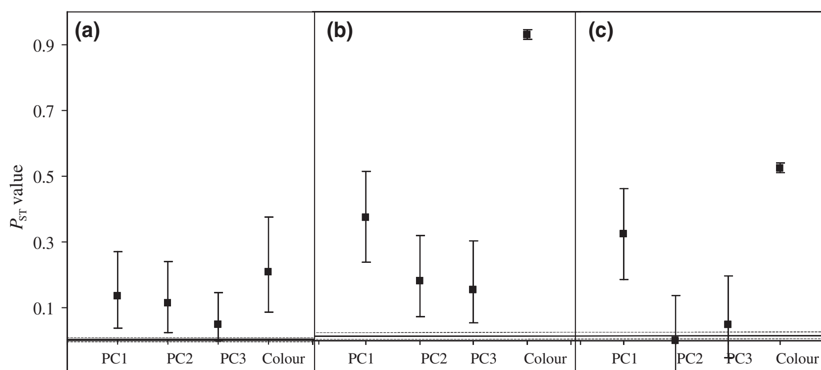

Fig. 5 Comparison of F_{ST} values calculated from 10 microsatellite loci with P_{ST} values (with their 95% confidence intervals) calculated from principal components and from colour classes between fish from 'shallow' and 'deep' waters (panel a), wild-caught 'blue' and 'red' (panel b) and laboratory-bred 'blue' and 'red' (panel c). Symbols from left to right in all panels refer to PC1, PC2 and PC3 and male coloration. The solid lines indicate the multilocus F_{ST} value. The dotted lines indicate the 95% confidence intervals of the F_{ST} value.

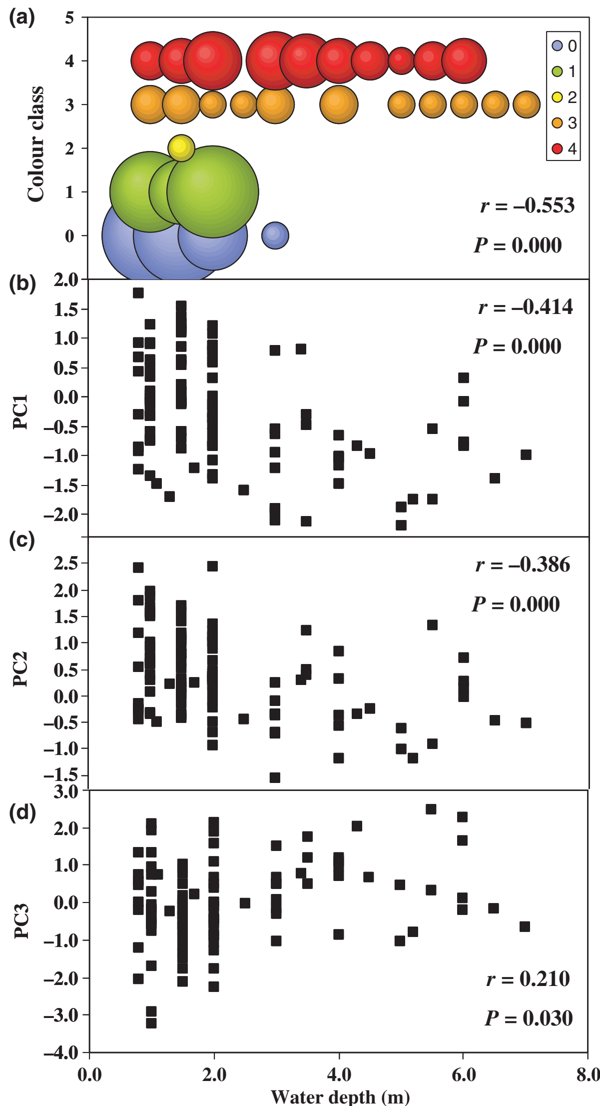

Fig. 6 Depth distribution (in metres) of individuals by (a) colour class; (b) principal component 1 scores; (c) principal component 2 scores; (d) principal component 3 scores.

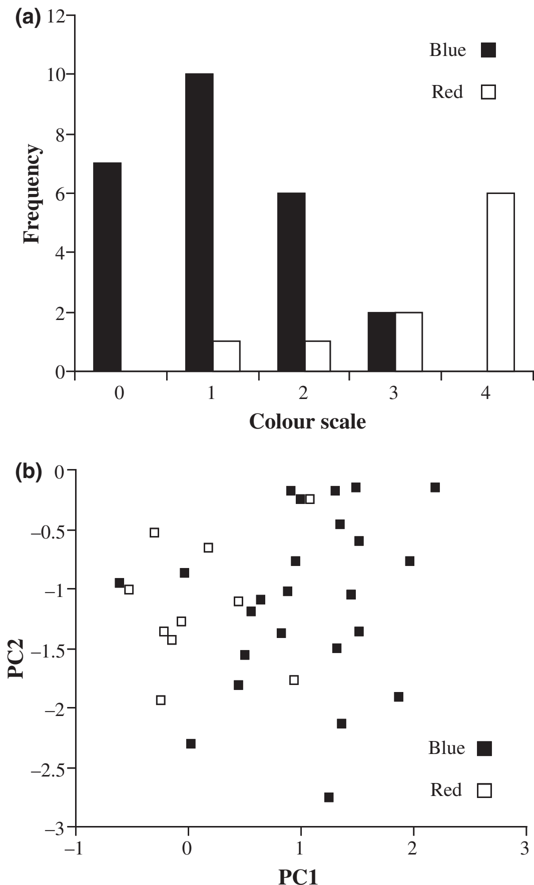

Fig. 7 Frequency distributions of (a) colour scores and (b) principal components 1 and 2 for laboratory-bred individuals.

Table 4 Mean values and confidence intervals of P_{ST} distributions generated from 625 pairwise comparisons for PC1, PC2 and PC3.

	P_{ST}		
	PC1	PC2	PC3
Mean	0.386	0.186	0.132
97.50%	0.811	0.661	0.628
2.50%	0.133	0.023	0.017

However, males bred from ‘red’ parents often had lower scores than males bred from ‘blue’ parents (Fig. 7b), resembling the situation in wild fish. There was complete overlap along the PC2 and the PC3 axes. Divergence between fish bred from ‘blue’ vs. ‘red’ wild phenotypes in the common garden experiment was the highest for coloration (P_{ST} = 0.525), followed by PC1 (P_{ST} = 0.323), PC3 (P_{ST} = 0.049) and PC2 (P_{ST} = 0.0005) (Table 3). Hence, all P_{ST} estimates among populations grown in a common environment were smaller than the P_{ST} among wild populations, but those for PC2 and PC3 were much smaller. Due to the small sample sizes, confidence intervals were large, and no P_{ST} estimate was significantly different from the neutral expectation except that for colour. If we assume that variation among the parental fish that we used for breeding was representative of the variation in the wild populations, we can apply the confidence intervals estimated from wild fish. Even with this approach it was clear that P_{ST} values from morphological PC2 and PC3 were close to the expectations from drift alone. By contrast, the PC1 and colour-derived P_{ST} values were significantly outside the neutral expectation also in the fish raised under identical conditions (Fig. 5c).

We took 50 random subsamples of six individuals from each of the two wild-caught incipient species groups defined by male nuptial coloration (‘blue’ and ‘red’). We

© 2008 THE AUTHORS. J. EVOL. BIOL. 22 (2009) 260–274
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

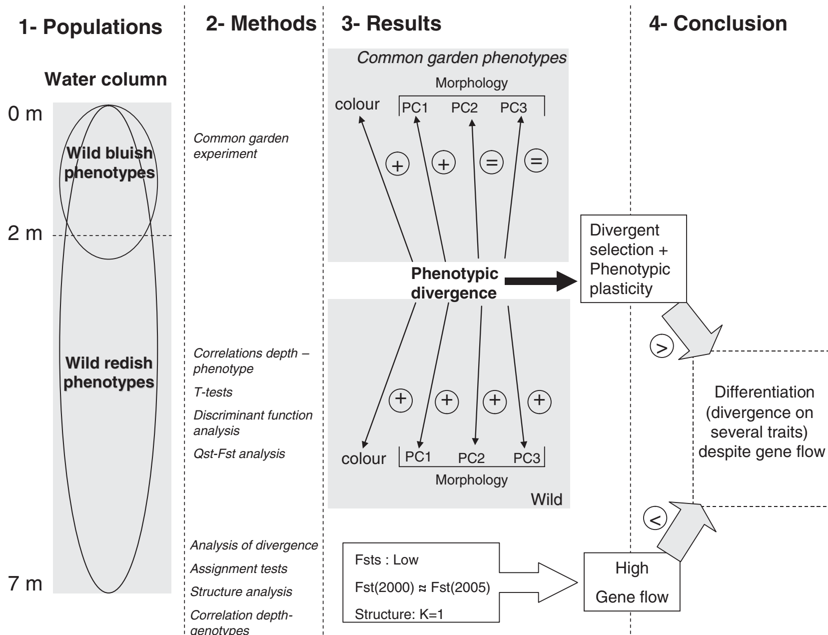

Fig. 8 Summary diagram showing: (1) the distribution of colour phenotypes across water depth; (2) main methods used to analyse phenotypic and genetic divergence and heritability of the phenotypic traits; (3) main results showing the level of divergence in quantitative traits for wild and laboratory-bred phenotypes and divergence at neutral markers. Symbols +, = and − are used if the trait divergence is higher, equal or lower than neutral expectations respectively; (4) conclusion showing how the different processes inferred from the results affect the measured characters and lead to their differentiation (>) or homogenization (<).

then made 625 pairwise comparisons and generated a distribution of P_{ST} values for each principal component. The P_{ST} values for PC2 and PC3 of our common garden populations laid below and very close to the lower confidence intervals of the P_{ST} distributions generated from the resampling of the wild populations respectively (Table 4). The P_{ST} value of PC1, however, was similar to the P_{ST} value generated from the resampling of wild fish, although its confidence interval was larger. Intrapopulation variances were not significantly different between wild and laboratory-bred populations for PC1, PC2 and PC3 of 'blue' phenotypes (Levene's test, PC1: $P = 0.263$, PC2: $P = 0.856$, PC3: $P = 0.226$) and 'red' (Levene's test, PC1: $P = 0.309$, PC2: $P = 0.156$, PC3: $P = 0.132$), but they were significantly larger for coloration of laboratory-bred populations (Levene's test, 'blue': $P = 0.002$; 'red': $P = 0.000$).

Discussion

Our study found that male nuptial coloration correlated with water depth and has a bimodal distribution and that divergence at neutral loci is extremely low but significant, which is consistent with previous findings (Seehausen *et al.*, 2008). We have also found that differentiation in morphological traits in wild fish correlated with water depth and is generally much larger than the degree of differentiation for neutral genetic makers, although never as large as differentiation in colours. However, several of the morphological differences were lost when fish were raised in a common laboratory environment. In the following section, we will discuss each of these findings and their general implications (see Fig. 8 for summary diagram of methods, results and conclusion).

Genetic structure and differentiation

Our results suggest that the phenotypically variable populations of *Pundamilia* around Kissenda Island are significantly structured at unlinked putatively neutral marker loci, but differentiation is weak and gene flow widespread. Water depth appears to be a structuring force, with significant isolation by depth in an individual-based analysis, even though the effect is small. Groups of individuals made based on catching depth were not significantly differentiated, but groups made based on male nuptial colour were significantly differentiated. Male nuptial colour in turn is significantly associated with water depth. The biological meaningfulness of small but significant F_{ST} values has rightly been questioned (Waples, 1998). However, in our case, biological meaningfulness of these low but significant measures of differentiation is supported by our result that similarly weak but also similarly significant differentiation was found in two independent samples taken with an interval of 5 years, whereas year had no significant effect on molecular variance. Hence, the differentiation between phenotype groups is unlikely just a spurious result of random from year-to-year variation. From this, we conclude that populations with blue or red male nuptial coloration are indeed incipient species at this location, even though there clearly is evidence for introgressive hybridization between them.

The more frequent occurrence of linkage disequilibrium in 'blue', when compared with 'red', suggests that 'blue' phenotypes are more strongly introgressed. 'Blue' phenotypes occupy predominantly shallower water depth, where they coexist syntopically with 'red' phe-

notypes. The latter, on the other hand, are distributed throughout the whole depth gradient, with maximum abundance at greater depth. They overlap with 'blue' only in part of their depth range, the shallow water. The considerable overlap of spawning habitat and the fact that genetic differentiation between the colour phenotypes was larger than that between the depth ranges, suggest that factors other than spawning site segregation must contribute to the partial isolation between these incipient species. The assignment tests confirmed the F_{ST} values; there is significantly nonrandom distribution of genotypes between the colour phenotype groups, suggesting either assortative mating within types or very strong selection against intermediates.

Phenotypic differentiation vs. neutral genetic divergence

Estimation of Q_{ST} or its analogous P_{ST} and F_{ST} values is challenging and subjected to a number of potential pitfalls (Hendry, 2002; O'Hara & Merilä, 2005; Whitlock, 2008). An error of most concern in the calculations of Q_{ST} which could also apply to P_{ST} values is the imprecision of the estimations when using low sample sizes, which tend to bias Q_{ST} downwards (O'Hara & Merilä, 2005). F_{ST} calculations and interpretations in high gene flow species can also be problematic as sampling error across loci may account for 1/2S of the obtained F_{ST} value (Waples, 1998). The observed F_{ST} value might, hence, be an overestimate. Hence, maximum caution is warranted when interpreting $Q_{ST}/(P_{ST} - F_{ST})$ comparisons. In our case, the above-mentioned potential sources of error seem conservative as we are searching for the signal of divergent selection.

Hendry (2002) also raised two important issues on the relationship between Q_{ST} and F_{ST}: (1) an assumption of this approach is that mutation rates are considerably lower than migration rates, which might be violated when markers with high mutation rates are used or populations with low gene flow analysed and (2) as divergence time increases and migration and mutation rates decrease demonstrating that Q_{ST} is higher than F_{ST} becomes increasingly harder, even if divergence of quantitative traits was driven by selection. Although we have used markers known to have high mutation rates, the F_{ST} values indicate extremely high gene flow between populations; therefore, migration rates are most probably considerably higher than mutation rates.

Whitlock (2008) recently suggested that comparing the Q_{ST} (or its equivalent value P_{ST}) for one trait with the mean multilocus F_{ST} was incorrect, and that it should instead be compared with the distribution of F_{ST} values. Our results are robust to this concern. We only considered evidence for selection the nonoverlap of the 95% confidence intervals, which are good representations of single locus F_{ST} distributions. Our P_{ST} estimates were in every significant case larger than the largest of our 10 single locus F_{ST} values.

All the P_{ST} values from wild fish were higher than the F_{ST} values. Confidence intervals were generally large, which is expected when using a small number of populations (O'Hara & Merilä, 2005). However, apart from one exception (PC3 when populations were divided by depth), they did not overlap with the confidence intervals of the F_{ST} values and had their lower end above the highest single locus F_{ST} values. These results therefore are consistent with the hypothesis of divergent selection on male nuptial colour and head morphology (PC1 and PC2).

Both the F_{ST} value and the P_{ST} values were higher when populations were divided based on colour than when populations were divided based on depth range. This difference can be explained by the fact that, even though there is a correlation between colour and water depth, there is considerable overlap between colour phenotypes in shallow water. Colour may be a better estimate of a male's long-term average water depth than the depth at which we caught the fish. However, this also implies that factors other than spawning site isolation by water depth contribute to partial isolation between phenotypes at Kissenda Island.

Divergent selection and phenotypic plasticity in morphology

Morphology was more divergent than expected by drift between fishes from different water depth and between fishes of different breeding coloration. The individual-based analysis also revealed a small but highly significant effect of depth distances between individuals on their morphology. However, in contrast to breeding coloration, variance in morphology was unimodally distributed with large overlap between the depth zones and colour phenotypes. Nonetheless, 91.5% of the individuals were correctly assigned in our discriminant functions analysis when groups were based on male colour, and all *t*-tests were significant, indicating clear differences in morphology between the incipient species.

We compared the phenotypic differences between wild-caught fish with those between animals raised in a 'common garden' laboratory environment. This allows us to estimate the relative contribution of phenotypic plasticity to differences found in the wild.

Coloration and principal component 1 of morphology were largely maintained when populations were raised in a common garden environment, whereas the differences in PC2 and PC3 dropped to the neutral F_{ST} line in laboratory-bred fish, suggesting that the differences in several morphological traits observed in wild fish are due to phenotypic plasticity. The P_{ST} values for laboratory-bred PC2 and PC3 were also significantly lower than expected by chance through random sampling of wild

fish for breeding. Therefore, it is unlikely that the loss of P_{ST} in PC2 and PC3 was due to a sampling bias. Hence, we refer to these axes as 'genetic' PC axis 1 and 'environmental' PC axes 2 and 3. Head length, lower jaw length and width, snout length and width, cheek depth and inter-orbital depth had more than 80% of their relative variance loading to PC1 and are therefore most likely traits with a high heritable component. These traits, all related to head morphology, were also the most important measures contributing to differences between groups in the discriminant functions analysis. The causes for divergence in these traits most likely arise from differences in feeding resources available at different water depths. Studies on the feeding ecology of cichlids from Lake Victoria have suggested that at islands where depth segregation between the sister species *P. pundamilia* and *P. nyererei* is strong, they have different diets (Bouton *et al.*, 1997), with individuals of *P. pundamilia* feeding more on benthic food, whereas *P. nyererei* feed more on plankton. These differences in diet are perhaps a consequence of the different microhabitats, shallow water and crevices in the blue *P. pundamilia* and greater average depth in the red *P. nyererei* (Seehausen, 1997; Seehausen *et al.*, 2008) and require phenotypic adaptation, some of which, we found to be achieved through heritable divergence.

On the other hand, divergence in the wild populations in the traits body depth, pre-orbital depth, eye length, eye depth and to a smaller extent head width was largely or partly due to phenotypic plasticity. With the exception of eye length and eye depth, these traits had positive loadings on genetic PC axis 1, as well as on environmental PC axes 2 and 3. From this we infer that environmentally induced divergence and genetic adaptation operate in the same direction, which is an additional argument for the action of divergent selection between the incipient species.

In conclusion, we found that the *Pundamilia* population at Kissenda Island in Lake Victoria is phenotypically and genetically structured along a gradient of water depth. Two incipient species with different male breeding colour occupy different but overlapping depth ranges and are weakly but consistently differentiated at neutral gene loci. This confirms previous findings of a correlation between ambient light colour along the water depth gradient, male nuptial colour and speciation. We show that male nuptial colour is heritable. Additionally divergent selection acts on ecologically relevant morphological traits along the same gradient. Response to this selection is partly accommodated by heritable adaptation and partly by phenotypic plasticity, where the latter appears to accentuate differences between the incipient species. Hence, divergent selection acts on several traits, contributing to the maintenance of differentiation despite high levels of gene flow.

Acknowledgments

We thank Bob O'Hara for writing the codes for R and Gaku Takimoto for implementing them. We also thank Denis Roy for the very useful discussions, Pascal Vonlanthen for comments on the manuscript, and the members of the Fish Ecology & Evolution laboratory for helpful feedback and discussion. We thank the Tanzanian Fisheries Research Institute for hospitality and access to facilities. Mhoja Kayeba and Mohammed Haluna provided crucial assistance in the field. This research was supported by Swiss National Science Foundation Project 3100A0-106573, to Ole Seehausen.

References

Barel, C.D.N., Vanoijen, M.J.P., Witte, F. & Wittemaas, E.L.M. 1977. Introduction to taxonomy and morphology of Haplochromine Cichlidae from Lake Victoria – manual to greenwoods revision papers. *Neth. J. Zool.* **27**: 333–338.

Bolnick, D.I. 2004. Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* **58**: 608–618.

Boughman, J.W. 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**: 944–948.

Bouton, N., Seehausen, O. & van Alphen, J.J.M. 1997. Resource partitioning among rock-dwelling haplochromines (Pisces: Cichlidae) from Lake Victoria. *Ecol. Freshw. Fish.* **6**: 225–240.

Brewer, M.J. 2003. Discretisation for inference on normal mixture models. *Stat. Comput.* **13**: 209–219.

Burnham, K.P. & Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. Springer, New York.

Carleton, K.L., Parry, J.W.L., Bowmaker, J.K., Hunt, D.M. & Seehausen, O. 2005. Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Mol. Ecol.* **14**: 4341–4353.

Dieckmann, U. & Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature* **400**: 354–357.

Dijkstra, P.D., Seehausen, O., Gricar, B.L.A., Maan, M.E. & Groothuis, T.G.G. 2006. Can male-male competition stabilize speciation? A test in Lake Victoria haplochromine cichlid fish. *Behav. Ecol. Sociobiol.* **59**: 704–713.

Dobzhansky, T. 1970. *Genetics and the Evolutionary Process*. Columbia University Press, New York and London.

Drossel, B. & McKane, A. 2000. Competitive speciation in quantitative genetic models. *J. Theor. Biol.* **204**: 467–478.

Emelianov, I., Marec, F. & Mallet, J. 2004. Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proc. R. Soc. B Biol. Sci.* **271**: 97–105.

Excoffier, L., Estoup, A. & Cornuet, J.M. 2005. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* **169**: 1727–1738.

Fleming, I.A. & Gross, M.R. 1994. Breeding competition in a pacific salmon (Coho, *Oncorhynchus kisutch*) – measures of natural and sexual selection. *Evolution* **48**: 637–657.

Gavrilets, S. 2004. *Fitness Landscapes and the Origin of Species*. Princeton University press, Princeton, NJ.

Glaubitz, J.C. 2004. CONVERT: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes* **4**: 309–310.
Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate *F*-statistics. *J. Hered.* **86**: 485–486.
Goudet, J. & Buchi, L. 2006. The effects of dominance, regular inbreeding and sampling design on Q(ST), an estimator of population differentiation for quantitative traits. *Genetics* **172**: 1337–1347.
Grant, P.R. & Grant, B.R. 2008. Fission and fusion of Darwin's finches populations. *Philos. Trans. R. Soc. B* **363**: 2821–2829.
Hardy, O.J. & Vekemans, X. 2002. SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **4**: 618–620.
Hendry, A. 2002. $Q_{ST} >= \neq < F_{ST}$. *Trends Ecol. Evol.* **17**: 502.
Hendry, A.P., Grant, P.R., Grant, B.R., Ford, H.A., Brewer, M.J. & Podos, J. 2006. Possible human impacts on adaptive radiation: beak size bimodality in Darwin's finches. *Proc. R. Soc. B Biol. Sci.* **273**: 1887–1894.
Higashi, M., Takimoto, G. & Yamamura, N. 1999. Sympatric speciation by sexual selection. *Nature* **402**: 523–526.
Kondrashov, A.S. & Kondrashov, F.A. 1999. Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**: 351–354.
Lande, R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* **46**: 381–389.
Leinonen, T., Cano, J.M., Makinen, H. & Merila, J. 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *J. Evol. Biol.* **19**: 1803–1812.
Leinonen, T., O'Hara, R.B., Cano, J.M. & Merilä, J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* **21**: 1–17.
Lu, G.Q. & Bernatchez, L. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. *Evolution* **53**: 1491–1505.
Maan, M.E., Hofker, K.D., van Alphen, J.J.M. & Seehausen, O. 2006. Sensory drive in cichlid speciation. *Am. Nat.* **167**: 947–954.
McKay, J.K. & Latta, R.G. 2002. Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.* **17**: 285–291.
Merilä, J. 1997. Quantitative trait and allozyme divergence in the greenfinch (*Carduelis chloris*, Aves: Fringillidae). *Biol. J. Linn. Soc.* **61**: 243–266.
Merilä, J. & Crnokrak, P. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol* **14**: 892–903.
Naisbit, R.E., Jiggins, C.D. & Mallet, J. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proc. R. Soc. B Biol. Sci.* **268**: 1849–1854.
Nosil, P. & Harmon, L.J. in press. Niche dimensionality and ecological speciation. In: *Ecology and Evolution* (R. Butlin, D. Schluter & J.R. Bridle, eds). Cambridge University Press, Cambridge.
Nosil, P., Harmon, L.J. & Seehausen, O. in press. Ecological explanations for incomplete speciation. *Trends Ecol. Evol.*
Nosil, P. & Sandoval, C.P. 2008. Ecological niche dimensionality and the evolutionary diversification of stick insects. *PLoS ONE* **3**: 1–11.
O'Hara, R.B. & Merilä, J. 2005. Bias and precision in *Q*(ST) estimates: problems and some solutions. *Genetics* **171**: 1331–1339.
Paetkau, D., Calvert, W., Stirling, I. & Strobeck, C. 1995. Microsatellite analysis of population-structure in Canadian polar bears. *Mol. Ecol.* **4**: 347–354.
Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
Rogers, S.M. & Bernatchez, L. 2007. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonas* sp Salmonidae) species pairs. *Mol. Biol. Evol.* **24**: 1423–1438.
Rousset, F. 2000. Genetic differentiation between individuals. *J. Evol. Biol.* **13**: 58–62.
Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**: 306–308.
Schliewen, U., Rassmann, K., Markmann, M., Markert, J., Kocher, T. & Tautz, D. 2001. Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. *Mol. Ecol.* **10**: 1471–1488.
Schluter, D. 1996a. Ecological causes of adaptive radiation. *Am. Nat.* **148**: S40–S64.
Schluter, D. 1996b. Ecological speciation in postglacial fishes. *Philos. Trans. R. Soc. B* **351**: 807–814.
Seehausen, O. 1997. Distribution of and reproductive isolation among color morphs of a rock-dwelling Lake Victoria cichlid (*Haplochromis nyererei*). *Ecol. Freshw. Fish.* **6**: 57–60.
Seehausen, O. in press. Progressive levels of trait divergence along a ''speciation transect'' in the Lake Victoria cichlid fish *Pundamilia*. In: *Ecology and Evolution* (R. Butlin, D. Schluter & J.R. Bridle, eds). Cambridge University Press, Cambridge.
Seehausen, O. & van Alphen, J.J.M. 1999. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* **2**: 262–271.
Seehausen, O., van Alphen, J.J.M. & Witte, F. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**: 1808–1811.
Seehausen, O., Witte, F., van Alphen, J.J.M. & Bouton, N. 1998. Direct mate choice maintains diversity among sympatric cichlids in Lake Victoria. *J. Fish Biol.* **53**(Suppl. A): 37–55.
Seehausen, O., Terai, Y., Magalhaes, I.S., Carleton, K.L., Mrosso, H.D.J., Miyagi, R., vander Sluijs, I., Schneider, M.V., Maan, M.E., Tachida, H., Imai, H. & Okada, N. 2008. Speciation through sensory drive in cichlid fish. *Nature* **455**: 620–626.
Slatkin, M. 1987. Gene flow and the geographic structure of natural-populations. *Science* **236**: 787–792.
van der Sluijs, I., van Alphen, J.J.M. & Seehausen, O. 2008. Preference polymorphism for coloration but no speciation in a population of Lake Victoria cichlids. *Behav. Ecol.* **19**: 177–183.

Spitze, K. 1993. Population-structure in *Daphnia obtusa* – quantitative genetic and allozymic variation. *Genetics* **135**: 367–374.

Taylor, M.I., Meardon, F., Turner, G., Seehausen, O., Mrosso, H.D.J. & Rico, C. 2002. Characterization of tetranucleotide microsatellite loci in a Lake Victorian, haplochromine cichlid fish: a *Pundamilia pundamilia* × *Pundamilia nyererei* hybrid. *Mol. Ecol. Notes* **2**: 443–445.

Van Oppen, M.J.H., Turner, G.F., Rico, C., Deutsch, J.C., Ibrahim, K.M., Robinson, R.L. & Hewitt, G.M. 1997. Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. R. Soc. B Biol. Sci.* **264**: 1803–1812.

Waples, R.S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* **89**: 438–450.

Whitlock, M.C. 2008. Evolutionary inference from *Q*(ST). *Mol. Ecol.* **8**: 1885–1896.

Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* **15**: 323–354.

Wu, L., Kaufman, L. & Fuerst, P.A. 1999. Isolation of microsatellite markers in *Astatoreochromis alluaudi* and their cross-species amplifications in other African cichlids. *Mol. Ecol.* **8**: 895–897.

Zardoya, R., Vollmer, D.M., Craddock, C., Streelman, J.T., Karl, S. & Meyer, A. 1996. Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid fishes (Pisces: Perciformes). *Proc. R. Soc. B Biol. Sci.* **263**: 1589–1598.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Loadings of morphological variables on the first three principal component axes, the per cent variance explained by each axis, and the relative per cent variance in each variable that load to each PC axis.

Table S2 Allelic variability at 10 microsatellite loci in *Pundamilia* from Kissenda island. Number of samples analysed in each population (N), number of alleles at each locus (A), allelic richness (AR), range of allele size (AS), observed heterozygosity (H_O) and expected heterozygosity (H_E) at each locus.

Table S3 Nested ANOVA performed on colour, PC1, PC2 and PC3 of laboratory-bred individuals of *Pundamilia* with brood nested inside species. Significant values ($P < 0.05$) are shown in bold face.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received 3 April 2008; revised 16 September 2008; accepted 28 September 2008