Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line

M.-L. Cariou*, J.-F. Silvain*†, V. Daubin*, J.-L. Da Lage* and D. Lachaize*

*Laboratory Populations, Genetics & Evolution, CNRS, 91198 Gif-sur-Yvette Cedex, France; MIRD, 213 rue La Fayette, 75480 Paris cedex 10, France

Abstract

We have used two paralogous genes (*Amyrel* and *Amy*) of the amylase multigene family to reconstruct the phylogeny of the nine *Drosophila melanogaster* subgroup sister species, including *D. santomea*, the newly discovered endemic from São Tomé island. The evolutionary divergence of these genes is of special interest as it is suspected to result from physiological evolution via gene duplication. This paper describes the relationship between the geographical origin of the various strains and the patterns of mating and phylogeny, focusing on the evolution of *D. santomea* and its relationship to other species and their niches. The *Amyrel* and *Amy* data indicate that, contrary to expectations, the sympatric insular *D. yakuba* population is less closely related to *D. santomea* than allopatric mainland ones, suggesting that the extant insular *D. yakuba* population on São Tomé results from a recent secondary colonization. Data for sympatric and allopatric *D. yakuba* suggest that *D. santomea* arose from a mainland *D. yakuba* parental stock when montane habitats of the Cameroon volcanic line extended to lower altitudes during colder and less humid periods. Despite their different modes of evolution and different functions, the *Amyrel* and *Amy* genes provide remarkably consistent topologies and hence reflect the same history, that of the species.

Keywords: amylase multigene family, *Drosophila santomea*, *D. yakuba*, Gulf of Guinea, phylogeny, São Tomé

Received 30 June 2000; revision received 19 October 2000; accepted 19 October 2000

Introduction

The molecular evolution of multigene families has recently received considerable attention (Liao 1999; Kogan et al. 2000). It has been argued that reconstructing the evolution of multigene families is central to understanding the evolutionary meaning of duplication and divergence of genes as a result of physiological adaptation through gene duplication (Da Lage et al. 1998). This study was carried out to verify the extent to which the evolutionary patterns of homologous (paralogous) genes of a multigene family mirror the phylogeny of sister species.

The amylase gene is well suited to studies on evolution of multigene families and for testing how adaptation and speciation are relevant to each other. The amylase gene (*Amy*) has been studied in a number of *Drosophila* species and a duplicate structure was found in *D. melanogaster* and all the related species within the subgroup (Boer & Hickey 1986; Dainou et al. 1987; Shibata & Yamazaki 1995). The number and arrangement of the genes, two close copies divergently transcribed, are conserved throughout the subgroup (Payant et al. 1988; Shibata & Yamazaki 1995).

*Amyrel*, a paralogous gene of the amylase family (so-called because it is ‘Amylase-related’) was recently identified in *D. melanogaster* and other species of the *Sophophora* subgenus, *D. ananassae* and *D. pseudoobscura* (Da Lage et al. 1998). Unlike *Amy*, *Amyrel* is a single copy gene, which has a specific intron position and is physically distinct from the *Amy* cluster on chromosome 2R (*Amy* maps at 54A and
Amyre1 at 53D). But although the two genes are closely set, they are strikingly different from one another because their divergence is 40% at both the nucleotide and the amino acid levels. The function of Amyre1 remains unknown. However, in contrast to the Amy genes, Amyre1 is transcribed in the larvae of D. melanogaster but not in adults, suggesting that the Amyre1 gene is akin to the breeding sites used. Amyre1 is thought to evolve approximately twice as fast as Amy in the various species studied by Da Lage et al. (1998). These authors, therefore, suggested that this gene could result from a duplication of Amy followed by accelerated and selected divergence toward a new function.

We have now carried out a comparative analysis of Amy and Amyre1 in a pair of sister species, one of which is suspected to have arisen through strong selection and possibly rapid adaptation. Our analysis has focused on a presumably suitable pair of sister species of Drosophila, D. yakuba and D. santomea, which are two melanogaster siblings of the afrotropical yakuba complex within the melanogaster species subgroup.

D. yakuba is widely distributed throughout the tropical African mainland from the Sahel to Swaziland and has spread to inland Madagascar and the neighbouring Ste-Marie island to the east of Tamatave, Zanzibar island off Tanzania in the Indian Ocean, and São Tomé Island off Gabon in the Atlantic Ocean. Its capacity to colonize new insular environments is probably related to its ability to cope with open, even semiarid habitats. D. santomea is a newly discovered insular species endemic to São Tomé, one of the four Gulf of Guinea islands, and the nearest relative of D. yakuba (Lachaise et al. 2000). Both siblings occur on the São Tomé volcano, but they are segregated because they live at different altitudes. D. yakuba inhabits lowland secondary habitats and D. santomea the montane mist forest above 1100 m. The two species come into contact at that elevation and form a hybrid zone. Whether this altitude difference is due to a change in elevation (that is to climatic change) or to a dramatic change in the vegetation type is still controversial. However, given the great elevational range of D. yakuba, from sea level to 3000 m in East Africa, for instance on Mt Elgon in Kenya (Lachaise et al. 1988), it appears that D. yakuba is not limited by elevation but by the unsuitable mist forest habitat. The confinement of D. santomea to higher altitudes could likewise be more due to the suitability of the type of vegetation (the mist Podocarpus forest) than to elevation per se.

The speciation of D. santomea from a mainland or insular D. yakuba stock may, therefore, be the result of adaptation to new resources in the submontane mist forest. If so, and considering that one gene (Amyre1) is only expressed in larvae while Amy is expressed in both larvae and adults, we anticipate that there should be some differences in the evolutionary patterns of these paralogous genes in the two Drosophila sister species. We tested this prediction by sequencing the Amy and Amyre1 genes, which have different patterns of evolution (Da Lage et al., 1998), for the new species D. santomea and the other eight sister species of the melanogaster subgroup. We analysed Amy and Amyre1 in the remote insular population of D. santomea and in a diversity of sympatric (insular) and allopatric (mainland) populations of D. yakuba, all originating from the Gulf of Guinea. Our purpose was to compare gene sequences from between and within species to infer the phylogenetic relationships of D. santomea and make comparisons to the phylogeny previously derived from the ‘clock’ gene, period, which is located on the X chromosome. A long series of between-species hybridization tests was carried out using sympatric and allopatric strains to assess further the species identity and affinity. Finally, migration scenarios that might account for the origin of D. santomea are discussed.

Materials and methods

Samples in the Gulf of Guinea

Samples of the isofemale lines of the three siblings of the yakuba complex (Drosophila santomea, D. yakuba and D. teissieri) was obtained by one of us (D.L.) in March–April 1998 in the Gulf of Guinea in central West Africa, with a special focus on the offshore portion of the Cameroonian volcanic line (CVL) more especially São Tomé. Table 1 summarizes the origin (location, country), the habitats and resources where the foundresses of the isofemale lines were caught. The references of the lines used are indicated. Molecular data were based on six D. santomea isofemale lines from one population (São Tomé) and five D. yakuba isofemale lines from four populations (one insular, São Tomé, and three continental) from the Gulf of Guinea (Fig. 1).

Origin of the other melanogaster sister species

We used the other seven sister species of the melanogaster subgroup to reconstruct the phylogeny. We retained one population line from the Gulf of Guinea whenever possible, except for the two Indian Ocean endemic species. Thus, we used one isofemale line (the type line 154.1) of D. erecta from Lamto in the Ivory Coast, where it breeds exclusively in the syncarps of at least three species of Pandanus. We also analysed D. orena, which was recorded only once in one location (at Bafut N’Guémba at 2100 m elevation) and has, therefore, long been considered to be endemic to the Szyngu submontane forest of Mt Lefo on the Baméléké Plateau in western Cameroon (Lachaise et al. 1988). D. orena would then be endemic to one volcano of CVL, and D. santomea and D. orena can be seen as the insular and mainland counterparts of CVL endemics. No additional isofemale lines have ever been collected since the discovery of D. orena.

© 2001 Blackwell Science Ltd, Molecular Ecology, 10, 649–660
Table 1: Origins of the suid ancestral lines

<table>
<thead>
<tr>
<th>Origin</th>
<th>Date</th>
<th>Altitude</th>
<th>Country</th>
<th>Habitat/Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>1881</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
in 1973. As a result, all published data on that species are based on the single (188.1) isofemale line founded at the time of the finding of the species. Finally, we included in the analysis the D. yakuba Y115 type strain (one isofemale line) which originated (1967) from Mangum near Foubot at 1100 m on the Koumden Plateau, part of the Bamiléké Plateau (Bamenda-Banse block) that is within CVL.

Outside the Gulf of Guinea, we used the type strain (128.2) of D. teissieri which originates from an high Chirinda forest reserve at the top of Mt Silinda in southeastern Zimbabwe near the border with Mozambique at an elevation of 1100–1200 m (H.E. Paterson, personal communication, 26th May, 1977), one strain of the three species of the simulans clade originating from the Indian Ocean; D. sechellia from Cousin islands in the Seychelles, D. mauritiana from Mauritius island, and D. simulans of unknown origin. D. melanogaster Canton-S was also added.

Hybridization tests

Two series of interspecific crosses were made using either multifemale strains (pools) or isofemale lines (iso ?). The D. santomea pool was started by mixing 50 females and 50 males of each of 10 isofemale lines (STO.1, 2, 3, 4, 5, 6, 7, 9, 11, 12). The D. yakuba São Tomé pool was made from a mix of 20 females and 20 males from each of four isofemale lines (SA.1, 2, 3, 4). The D. yakuba pool from Libreville, coastal Gabon was made by mixing 20 females and 20 males from each of two isofemale lines (LBV.1, 2). The D. yakuba pool from Léop. Central Gabon was made by mixing 50 females and 50 males from each of 10 isofemale lines (LO.1, 2, 3, 4, 5, 6, 7, 8, 9, 10).

A second series of crosses was made using isofemale lines separately. A total of 10 (STO.1, 2, 3, 4, 5, 7, 11, 12, 30, 37) and 12 (idem plus STO.9, STO.16) isofemale lines from the same population were used for D. santomea females and males, respectively. Four isofemale lines from one sympatric (SYM) irregular population (SA.1, 2, 3, 4) and five isofemale lines from three allopatric (ALLOP) mainland populations, LBV (2), LO (2), 115 (1) of D. yakuba, were used.

The fertility of the male and female hybrids was tested by backcrossing approximately 20 per cent of a total of 565 crosses.

Analyses of Amyr eld and Amy

D. santomea amylase genes were amplified using primers designed from published sequences of D. yakuba and D. teissieri (Shibata & Yamazaki 1995) to match conserved sequences at the 3' and 3' flanking regions. The 1526 bp
AMYLASE MULTIGENES AND D. SANTOMEA PHYLOGENY 653

Fragment amplified included the full coding Amy sequence: forward primer 5’CAGAGTGAACCTAAGTCCG3’, reverse primer 5’CCAGCTGTATACACTGCG3’. Single fly DNA extracts were amplified using standard polymerase chain reaction (PCR) conditions and the following thermal cycler conditions: 3 min at 94 °C; 30 cycles of 30 s at 94 °C, 1 min at 57–58 °C and 2 min at 72 °C; followed by a final 10 min at 72 °C. As two genes have been reported for all species of the melanogaster subgroup, the PCR products were purified and cloned in a PGEM-T cloning vector (Promega). Three to five clones were sequenced for each individual from each experimental run.

The primers for the Amyr1 locus were designed to match the conserved 5’ and 3’ flanking regions reported by Da Lage et al. (1998): forward primer 5’TGGACCTTTACACCTGCG3’; reverse primer 5’CATACATTGTCGCTTGCG3’. The method used for Amyr1 was similar to that of Amy with the following thermal cycler conditions: 3 min at 94 °C; 35 cycles of 30 s at 94 °C, 1 min at 55 °C and 2 min 30 s at 65 °C; followed by a final 10 min incubation at 72 °C. The PCR products were also cloned in the PGEM-T cloning vector before sequencing. The 1597 bp fragment amplified included the full Amyr1 gene.

All of the sequencing was performed by the dye terminator chemistry using an ABI 373 sequencer (Perkin Elmer Applied Biosystems Inc.). The sequences obtained from this work are available in GenBank accession nos for Amy: AF280880–280885 santomea, D17737 yakuha-dist, D17738 yakuha-prox, AF280889 yakuha LO, AF280888 yakuha-LB, AF280886–280887 yakuha-SA, AF280890 teissieri T2, D17735 teissieri-dist, D17736 teissieri-prox, X04569 melanogaster 1 prox, D17734 simulans-prox, D17730 mauritiana-prox, D17732 schellina-prox, D17728 erecta-prox, D21129 ornea prox; and for Amyr1: AF280868–280873 santomea, AF039561 yakuha Y115, AF280876 yakuha LO, AF280876–280877 yakuha-LB, AF280874–280875 yakuha-SA, AF039557 teissieri, U69607 and AF022713 simulans, U69159 simulans, U69157 mauritiana, AF039558 schellina, AF039562 erecta, U69158 ornea).

Data analysis
Sequence data from the nine D. melanogaster subgroup species were analysed with sealign for Macintosh (Gilbert 1992) and aligned with clustal w (Thompson et al. 1994). Reconstruction of phylogenetic trees was based on the full coding sequence of Amy (1482 bp) and Amyr1 (1479 bp) and derived by either distance (neighbour-joining, NJ) or maximum likelihood (ML) methods. The NJ reconstructions used pairwise evolutionary distances estimated using the Kimura 2-parameter (K2P) (Kimura 1980) and the Tajima and Nei (TN) model of substitutions (Tajima & Nei 1984) as implemented in phylo_win (Galtier et al. 1996). Bootstrapping was based on 2000 replicates. For ML tree inference we used the Tamura and Nei (TrN) (Tamura & Nei 1993) substitution model from the TREE-PZUZZLE 4.02 program (Strimmer & Von Haeseler 2000). Quartet puzzling support values were based on 2000 iterations. D. orena was used as an outgroup in both Amyr1 and Amy phylogenetic trees. Using such a related outgroup is possible considering that more distant outgroups like D. ananassae, a distantly allied species of the melanogaster group (Lemeunier et al. 1986), provided topologies where D. orena consistently occupies a basal position with regard to the remaining D. melanogaster subgroup species. There is also considerable evidence that D. orena arose from the most ancient split within the melanogaster species subgroup (Lemeunier et al. 1986; Cariou 1987; Lachaise et al. 1988; Russo et al. 1995). Published sequences from proximal (prox) and/or distal (dist) Amy genes (Shibata & Yamazaki 1995) were also used. Among the Amy genes we were not able to distinguish between proximal and distal copies. However, the sequence divergence between such closely linked copies is generally very low due to concerted evolution (Brown et al. 1990; Popadic & Anderson 1995; Shibata & Yamazaki 1995), making the tree topology not very sensitive to the use of one duplicate or the other.

Results
Hybridizations
Table 2 summarizes our present knowledge of the reproductive relationships between the nine sister species of the melanogaster subgroup, including Drosophila santomea. Although the basic data implicating D. santomea have been published (Lachaise et al. 2000), they are given here as further evidence on a larger scale. Table 2 highlights the novelty of the reproductive pattern of D. santomea, which offers the possibility of producing fertile female hybrids in both directions of cross between two sister species of the melanogaster subgroup outside the melanogaster complex. The number of successful interspecies and intraspecies crosses as against the number of crosses were tested (N1/N2) between D. santomea from São Tomé island and either sympatric insular or allopatric mainland (Gabon, Cameroon) populations of D. yakuba. The ‘success’ of hybridizations was assessed by the production of female plus male F1 hybrids regardless of whether they were sterile or fertile. Crosses between isofemale lines involving D. santomea females and D. yakuba males gave a similar proportion of successful crosses regardless of whether the population of the latter species was sympatric or allopatric with that of the former species (29/57 vs. 26/46, $U = 0.571$, $P = 0.284$, NS). Similarly, crosses in the other direction gave results that were irrelevant to the origin of the D. yakuba population (213/268 vs. 210/263, $U = 0.848$, $P = 0.198$, NS).

© 2001 Blackwell Science Ltd, Molecular Ecology, 10, 649–660
The fertility of the male and female interspecific hybrids was tested by backcrosses with some of the various F1 progenies — and invariably gave sterile F1 males and fertile F1 females in both directions of cross. F1 hybrids between sympatric and allopatric D. yakuba were fertile.

As a consequence, there was no difference in the postzygotic isolation between D. santomea and sympatric or allopatric populations of D. yakuba. In contrast, when all D. yakuba populations were pooled [D. santomea × D. yakuba (55/103) as against D. yakuba × D. santomea (423/531)] there was a significant difference in the success of hybridization, depending on the direction of cross ($U = 5.663, P = 0.010^{**}$).

**Amy and Amyrel genes**

The amylose gene of *D. santomea* has the same structure as that of all the other species. The gene is 1482 bp long, has no intron and the putative protein is 494 amino acids (aa) long. The Amyrel of *D. santomea* has the typical structure of this gene. There is a short intron (57 bp) at position 655, the coding sequence is 1479 bp long and the putative protein has 495 aa. This structure is conserved for all the species within the melanogaster subgroup, yet the length of the intron varies from 57 to 60 bp.

For both Amy and Amyrel, there was greater variation in synonymous sites than replacements and for Amyrel greater variation in introns than in exons, as expected under a hypothesis of reduced constraint, although we analysed a limited sample of isofemale lines for the three species *D. santomea*, *D. yakuba* and *D. teissieri*.

The six Amyrel introns of *D. santomea* have identical sequences, but they are differentiated from the other species. Their alignment (Fig. 2) separates the intron of *D. santomea* from that of *D. yakuba* including the Y115 reference strain.

The three species, *D. santomea*, *D. yakuba* and *D. teissieri*, showed a similar trend in that Amyrel protein appeared to
AMYASE MULTIGENES AND D. SANTOMEA PHYLOGENY

Fig. 3 Phylogenetic reconstruction of Amy sequences from D. santomea and D. yakuba from the Gulf of Guinea and their seven sister species of the D. melanogaster subgroup. The tree was constructed by maximum likelihood (ML) to/lv = 2; Tamura and Nei model of substitutions. In likelihood = 3229 as implemented in PRUZLE 4.02 (Strimmer & Von Haeseler 2000). Quartet puzzling support values are based on 2000 iterations. Species designations are on the right. Sequences from proximal (pro) and/or distal (dist) genes are from Shibata & Yamazaki (1995).

neutral and the three others correspond to loss of positive charges, which explains the peculiar electrophoretic mobility of the santomea amylase variant. We found that D. santomea had a fast variant with a mobility similar to that of AMY1 of D. melanogaster and very distinct from those in D. yakuba and D. teissieri, which both had two electromorphs, AMY 3.4 is predominant in D. teissieri and AMY 4.4 in D. yakuba (Dainou et al. 1987).

Gene tree estimates

Figures 3 & 4 show the ML reconstruction of the phylogenetic tree of the nine D. melanogaster subgroup species, including six D. santomea and five D. yakuba isofemale lines from the Gulf of Guinea, based on the 21 nucleotide sequences of the Amy genes and the 20 sequences of Amyr. We obtained the same overall topology from the two genes, Amy and Amyr, regardless of the method (ML or NJ). The Quartet puzzling support values were very consistent with the NJ bootstrap values. All the nodes on the gene trees had high quartet puzzling support values.
close or equal to the maximum, except for the node leading to D. yakuba, indicating a clear gene differentiation between species. The sequence data based on Amy and Amphyl supported the present species designation and relationships. All the species of the D. melanogaster complex were clustered together and their positions were in agreement with the generally admitted relationships: D. erecta was near D. orena and the four melanogaster siblings were clustered together, giving further evidence of the validity of the Amphyl and Amy trees. There was, however, one exception — the position of D. melanogaster branching within the simulans clade in the Amy tree.

The sequences of D. santomea, D. yakuba and D. teissieri samples clustered with their respective groups and each of these groups had no Amphyl or Amy sequences from the other groups.

Apart from these intraspecies groupings, the trees always supported sister taxa relationships among species. In particular, D. santomea consistently formed a sister group to D. yakuba. D. teissieri also appeared to be a sister group to a clade that included the two former species for both Amy and Amphyl genes, confirming the reality of the yakuba species complex.

The tree topologies for D. yakuba suggested different patterns. The Amy tree supported substantial amounts of gene flow within the species, as there does not appear to be a tendency for the different lines to cluster according to their geographical origin. Instead, the D. yakuba isofemale lines or strains were interspersed (Fig. 3). It should however, be noted that the Amy gene evolves more slowly than Amphyl and the lower number of phylogenetic informative sites within D. yakuba makes the Amy tree less reliable at the intraspecific level.

The Amphyl tree showed the opposite tendency, with a good agreement with geographical affinities. The D. yakuba Y115 line from the mainland CVL (Koundou Plateau, West Cameroon) was in the most basal position, followed by the line from inland Gabon, the two lines from coastal Gabon and then, terminally, the two lines from São Tomé island (Fig. 3). Although weakly supported (bootstrap value = 60%), this Y115 line clustered with the D. santomea group in the NJ construction (tree not shown). Given the terminal positions in the D. yakuba cluster of both SA lines, these two D. yakuba lines from the insular santomea population did not appear to be the closest to the D. santomea cluster, but the most recently evolved lines.

Discussion

Further evidence for Drosophila santomea as a good species

Large scale crosses between Drosophila santomea and D. yakuba invariably produce sterile male and fertile female hybrids, with a significant difference in the success rates of hybridization depending on the direction of cross. This supports our earlier findings (Lachaise et al. 2000). Although premating isolation tests were not performed, these data suggest that there is a marked asymmetry in behavioural isolation between the two siblings. There has been considerable debate as to whether such asymmetry indicates the direction of evolution. The initial hypothesis of Kaneshiro (1976) was that the larger ancestral D. yakuba population retained the more elaborate mate recognition system and discriminated against D. santomea males with the simplified system, whereas females with the simplified requirements accepted equally readily both derived and ancestral males. However, our crosses were definitely easier between D. yakuba females and D. santomea males than vice versa. Which supports the contrary prediction of Watanabe & Kawanishi (1979) that ancestral species show less pre-mating isolation from derived species than vice versa (see Powell 1997 pp. 240–241 for review and references therein). Nevertheless, we still do not know how long the allegedly derived species (D. santomea) has been isolated from the mainland stock (D. yakuba). We agree with Powell (1997) that any such asymmetry resulting from simplified mate recognition systems in derived populations is presumably temporary. New complex systems probably arise and this may be a major event in speciation.

The sterility of male hybrids definitely indicates that the two forms are distinct yet closely related species. And the two paralogous genes (Amy and Amphyl) also support the validity of D. santomea as a good species, the sister species relationship between D. santomea and D. yakuba, and the monophyly of the D. santomea, D. yakuba and D. teissieri group, whichever method is used to construct the phylogenetic trees. These data are fully consistent with the stereotyped obtained with the period gene and a range of other data, including the morphology of terminalia, chromosomes and allozymes (Lachaise et al. 2000). Thus, at least three nuclear genes on various chromosomes, Amy and Amphyl on chromosome 2, and period on the X chromosome, provide consistent evidence that D. santomea and D. yakuba are a pair of siblings within the yakuba species complex, which is invariably identified as a monophyletic group.

Two endemic siblings in the CVL

The finding of D. santomea in São Tomé island indicates that the CVL is the historical home range of the melanogaster species subgroup. Two of the 12 volcanic centres of the Cameroon line are now known to harbour one endemic melanogaster sibling, D. orena is confined to one volcano (Mt Lefo) of the continental sector of CVL and D. santomea to one volcano of the offshore sector of CVL. Both live in fragmented submontane forest patches, and CVL as a whole, therefore, appears to be an archipelago made up of
oceanic and continental rainforest ‘islands’. These findings emphasize the role of specific montane areas as evolutionary centres (Fjeldså & Lovett 1997; Cobb et al. 2000). Both Drosophila endemics are confined to remote altitudinal submontane forests. Even D. erecta, which is commonly found in lowland Pandanus swamps, has been found at 2100 m in the CVL, together with D. orena, D. teissieri and D. yakuha, at Bafut N’Guemba on Mt Lefo (Lachaise et al. 1988). Therefore, CVL was probably the major corridor involved in the speciation of melanogaster sister species. A similar phenomenon occurred more recently among the simulans clade through the Indian Ocean archipelagos.

**Sympatric vs. allopatric D. yakuha**

We used a diversity of D. yakuha lines from São Tomé island (i.e. the offshore CVL) and various neighbouring places of the Gulf of Guinea mainland in an attempt to determine which D. yakuha parental stock gave rise to D. santomea. We tested D. yakuha lines from coastal Gabon and inland Gabon and one line from the continental sector of CVL (Koundou Plateau in West Cameroon). Assuming that speciation of D. santomea resulted from evolution on the São Tomé volcano, there should be less molecular divergence between D. santomea and sympatric D. yakuha than between D. santomea and any allopatric D. yakuha. There should also be greater divergence between the insular D. yakuha population and any mainland ones than between the mainland populations. Our data for the Amyr and Amy genes clearly show that D. santomea is not more closely related to the sympatric D. yakuha population than to any continental one. On the contrary, the Amyr phylogenetic tree indicates that the insular (SA) D. yakuha lines branch distally rather than basally within the D. yakuha cluster. Thus, contrary to expectation, they appear to be the most distantly related D. yakuha lines with regard to D. santomea, suggesting that the extant insular D. yakuha population on São Tomé results from a very recent secondary colonization.

In agreement with this, there is no difference in the postzygotic isolation between D. santomea and sympatric or allopatric populations of D. yakuha. But it is still possible that the prezygotic behavioural isolation with sympatric D. yakuha, especially in the hybrid zone, was reinforced; this is currently being tested. Taking into account the above arguments, and if reinforcement is to be expected in such a case, then the secondary colonization of São Tomé by D. yakuha may not have had enough time for any such reinforcement to evolve.

This poses the question of where D. santomea arose, assuming double colonization. Although the evidence is still weak, both the Amyr and Amy phylogenetic trees point to the onshore portion of CVL. The D. yakuha (Y115) line from the Koundou Plateau, within the larger Bamiléké Plateau in western Cameroon, which was used for the Amyr tree, branches in a very basal position among the D. yakuha lines, regardless of the algorithm used. ML indicates that Y115 clusters with the Gabonese and Santomé D. yakuha lines, but as an outlier. NJ enhances this tendency, and puts Y115 outside the D. yakuha species-specific cluster, but its place between the D. yakuha and the D. santomea clusters is poorly supported (low bootstrap value of 60), whereas the bootstrap is near 100 for all the other clusterings.

Such lineage sorting indicates that some forms within the D. yakuha gene pool have kept some ancestral characters common to D. yakuha and D. santomea. Another possibility is that the branching of D. yakuha lines, especially the reference strain Y115, also corresponds to a pattern expected if the sequences of D. yakuha and D. santomea have histories that include some recombination events. This single D. yakuha line from mainland CVL is at least an indication where populations with ancestral traits could be found if they still exist. We are currently analysing DNA sequence variations in D. santomea, D. yakuha and D. teissieri to determine more precisely the amount and pattern of variation, as these estimates reflect the histories of populations. There are also some ecological arguments suggesting that the D. yakuha-santomea ancestor had opportunities to migrate through CVL.

**The Podocarpus migration route**

The stepping-stone colonization of a mainland D. yakuha-santomea ancestor into the Gulf of Guinea islands may have been coincident with the spread of Podocarpus during some colder and less humid periods in oxygen isotope stages 5d and Sb (Frédouix 1994). This occurred around 150 000–130 000 years ago (Van Andel & Tzedakis 1996), when the African rainforest became fragmented and montane habitats extended to lower altitudes (Maley 1987). Frédouix (1994) stated that the abundance of Podocarpus pollen during that cooling period in deep-sea cores taken in the Gulf of Guinea suggests there was a broad geographical distribution of this conifer in western Africa. That abundance is in contrast to its present-day restriction to montane habitats of CVL between 800 and 2000 m and São Tomé Island above 1400 m. Its wide occurrence during the transitional phases of cooling was favoured by coolings in the equatorial zone (Maley et al. 1990). The very restricted species ranges of the two montane endemics, D. orena and D. santomea, match quite nicely two such scattered fragmented ranges of Podocarpus. The CVL shown in Fig. 1 could be the route by which Podocarpus pollen, and also the D. yakuha-santomea ancestor, were transported to the Gulf of Guinea islands. The tropical mountains, which are cooler than the lowlands, with mists replacing rain (Maley

1987), may have favored the speciation of *D. santomea*. However, there is still no evidence of an association between *D. santomea* and *Podocarpus*. An alternative pathway would be the fig tree route.

**The Ficus chlamydocarpa migration route**

*F. santomea* has so far been found only on fallen figs of the hemiepiphyte fig tree, *Ficus chlamydocarpa*, at 1200–1400 m in the São Tomé submontane forest. This fig tree species is endemic to CVL, where there are three subspecies (Berg & Wiersinga 1992). *F. chlamydocarpa* ssp. *chlamydocarpa* is found in the submontane forests of Cameroon and Bioko, *F. chlamydocarpa* ssp. *ferrandisiensis* (Berg 1988) occurs in São Tomé at 340–1400 m (Lejoly 1995; D. Lachaise, personal communication), and *F. chlamydocarpa* ssp. *latifolia* (Berg 1988) grows in Annobon (Pagalu) (15 km) at low altitudes. The mainland fig tree subspecies is rare, growing only in a few submontane forests at altitudes of 1300–2000 m. It has been recorded in the following localities: 40 km east-south-east Nkambe; Mt Tabenken, 10 km south-east Nkambe; Dschang, Mts Bambutos near Dschang, Mbom, Fotouni, 20 km north-north-east Bafang, Mt Cameroon, Tonja (Berg et al. 1985). These locations are on either side (north-east and south-west) of Bafut N’Guemba on Mt Lelo where *D. yakuba* was found up to 2400 m, and the mainland subspecies of this fig tree happens to occur at 2000 m. Hence, the *D. santomea*/*D. yakuba* split could have occurred at the same time as, and be related to, *F. chlamydocarpa* subspeciation.

Whichever host plant is implicated, there seem to have been repeated opportunities for *D. yakuba* to spread along the CVL. However, while the fragmentation of the host-plant range and hence of the *D. yakuba-santomea* ancestor's home range is plausible for the mainland CVL and Bioko, it is not for São Tomé.

The island of Bioko, formerly Fernando Po, lies on the continental shelf 32 km off the coast of Cameroon, and is therefore part of the continental sector including Mount Cameroon (4005 m). Bioko is separated from the shore of Cameroon by a strait only 60 m deep. Data from cores off the west coast of Africa, which record changes in pollen flux from the continent, indicate that the sea level was 70 m below the present level some 39 000–36 000 years ago during the oxygen isotope Stage 3 (van Andel & Tzedakis 1996). The last glacial aridity maximum occurred some 22 000–13 000 Carbon 14 (14C) years ago (approximately 23 000–14 500 ‘real’ years ago). Bioko became an island only after a postglacial rise in sea-level around 13 000–12 000 years ago. This largest (800 km²) and highest (2850 m) island of the four Gulf of Guinea islands, is still unexplored for *Drepanofila* and could well harbour a *D. yakuba* population from which *D. santomea* might have arisen, possibly by a founder effect.

The other three islands, Annobon, São Tomé and Principe, were probably never connected to the African mainland or to each other. Consequently, the colonization of São Tomé by the *D. yakuba-santomea* ancestor requires a long migration of 280 km from the Gabon coastline. It if came from the Cameroon mainland through the CVL it would require a stepwise 220 km leap from Bioko to Principe and then a 146-km one from Principe to São Tomé. The colonization of Principe from the mainland would also require an airborne transport over 220 km.

Any attempt to determine the time of the split between *D. yakuba* and *D. santomea* must take into account these palaeobiogeographic and molecular clock data. However, the use of molecular clocks requires independent estimates from one or two former sources to calibrate the clock (Powell 1997). There are no reliable estimates of the rates of substitution in *Amyrál*, which is evolving faster than *Amy* (Da Lage et al. 1998). A synonymous rate of 6.8 × 10⁻⁸ substitutions per nucleotide per year has been proposed for the distal amylase duplicate of *Amy* (E.N. Moriyama unpublished, in Li 1997). We applied this to the synonymous site divergence (K2P distance) and estimated that the split between *D. santomea* and *D. yakuba* occurred some 0.45 Ma. This absolute time of divergence is not too inconsistent with the *Podocarpus* pathway scenario (considering all the uncertainty of molecular clock calibration, it could be within the same order of magnitude), and may possibly match the time required for *F. chlamydocarpa* subspeciation.

In conclusion, although the products of the two genes of the amylase family studied differ markedly, and the genes have evolved differently and are differently expressed, there is, nonetheless, an overall consistency within their phylogenetic patterns. Both genes may have been implicated in these adaptations toward new resources in the submontane forests of the CVL. These resources could be *Podocarpus* or any *Podocarpus*-associated plant at higher altitudes or some *Ficus* at lower altitudes. The many amino acid replacements in the amylase protein of *D. santomea* may indicate resource specialization. We, therefore, postulate that the evolutionary changes occurring at both the *Amyrál* and the *Amy* loci have some common feature that is relevant to the adaptation of *D. santomea* to local resources.

**Acknowledgements**

We thank the Government of São Tomé e Príncipe and J Rosseel, T. d’Espiney and J-F Fontaine, ECOFAC-São Tomé for permission to work in the Obo National Park and for assistance in the field. We thank the Government of Gabon through Mr E. Mamoundoubi Kombila, Ministère des Eaux et Forêts for permission to work in the Lopé Forest Reserve. Our particular thanks go to K. Abernethy, R. Osilany, C.E.G. Tutin, Centre International de Recherche Médicale de Franceville, Station d’Etude des Gorilles et Chimpanzés.

and L.J.T. White, The Wildlife Conservation Society, New York, for providing valuable scientific and logistic assistance. We thank F. Maczkovics for technical assistance and J. Lepoy for botanical information on São Tomé. The English text was edited by Dr Owen Parkes.

References


Gilbert DG (1992) sqapp, a philosophy editor and analysis application (http://bio.indiana.edu/molbio).


This paper is a contribution to a broad study on speciation and molecular polymorphism in insular and continental sister species of the *Drosophila melanogaster* subgroup in the Afrotropical region. Marie-Louise Cariou is Head of the Laboratory ‘Populations, Genetics & Evolution’, of the CNRS at Gif-sur-Yvette, and has a broad interest in gene evolution and molecular polymorphism. Jean-François Silvain is from the Institut de Recherche pour le Développement (IRD) and is deeply involved in molecular phylogeny and ecology, and plant-insect relationships. Vincent Daubin is a PhD student and has contributed the data on the amylase multigene study. Jean-Luc Du Lage is a geneticist at the CNRS and is especially interested in the molecular genetics and evolution of multigene families. Daniel Lachaise is an evolutionary ecologist at the CNRS and has a special focus on speciation of Afrotropical *Drosophila*. He discovered *D. santomea* in São Tomé Island.