UPPSALA
UNIVERSITET

Cytonuclear co-evolution in the mitochondrial OXPHOS system

Christina Classson

Degree project in biology, Bachelor of science, 2009
Examensarbete i biologi 15 hp till kandidatexamen, 2009
Biology Education Centre and Department of Evolutionary Biology, Uppsala University
Supervisor: Johan Lindell

Contents

Abstract

The energy producing system of enzyme units in the inner mitochondrial membrane control the oxidative phosphorylation (OXPHOS) pathway and are thereby responsible for the energy production of a cell. Since the enzymes of the system are encoded partly from nuclear genes and partly from mitochondrial genes there is believed to be selection on interacting units of the system.

Imperfect interaction has been proposed as a cause of hybrid breakdown, which is often observed in hybrid zones. Since F_2 hybrids have different nuclear and mitochondrial genomes it is believed that their observed loss of fitness is due to incompatibilities between the nuclear and mitochondrial encoded subunits.

In this study I investigate if there is any observable pattern of selection on two interacting OXPHOS genes in populations of the black tailed brush lizard that inhabits the entire length of the peninsula of Baja California.

Introduction

Speciation is a process that involves the gradual build-up of genetic incompatibilities between diverging populations (Coyne & Orr 2004). This leads to reproductive isolation as signified by negative selection of hybrids formed in cross-breeding of diverging populations. The diverging genomes accumulate incompatibilities until the two populations are unable to interbreed and exchange genetic material (fig 1 ; Wu & Ting 2004).

Particular genomic regions or genes may initiate and drive the differentiation process. For example, it has been shown that genes linked to metabolism plays a crucial part in the fitness of an individual and selection of such genes may therefore be important in the differentiation process of diverging populations (Blier *et al*. 2001). Interpopulational differences in such genes have been linked to the reduced fitness observed in F_2 hybrids (Burton *et al.* 2006).

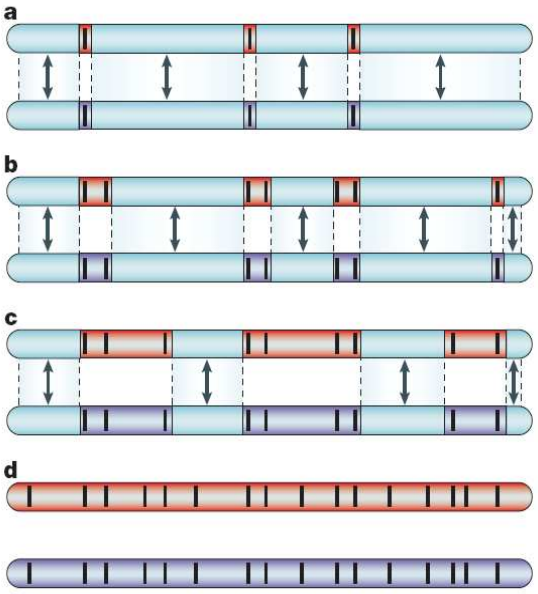

Figure 1. The process speciation through diminishing gene flow between diverging genomes. Initially, two diverging genomes (exemplified by horizontal bars) acquire mutations that cause some incompatibility between them (a). Through time there is a gradual decrease in gene flow between linked regions, causing the incompatible regions to expand (b & c). Eventually, no gene flow is possible and complete reproductive isolation is reached (d). (Adapted from Wu & Ting 2004.)

Oxidative phosphorylation and cytonuclear co-evolution

In eukaryotes, mitochondria produce the great majority (>90%) of the energy available for cellular metabolism (Rand *et al.* 2003). In the inner membrane of mitochondria resides a system of five protein complexes that are involved in the energy production of the cell. The enzymatic pathway mediated by this system is known as oxidative phosphorylation (OXPHOS) and produces ATP, the main energy unit of the cell (fig 2.). Consequently, the function of this pathway plays a key part in the energy production of the cell and, by extension, the fitness of the organism. Interestingly, approximately 73 of the polypeptide subunits of the complexes are encoded by genes residing in the nucleus, whereas 13 subunits are the products of genes of the mitochondrial genome. Thus, the OXPHOS complexes result from a close interplay of two different genomes.

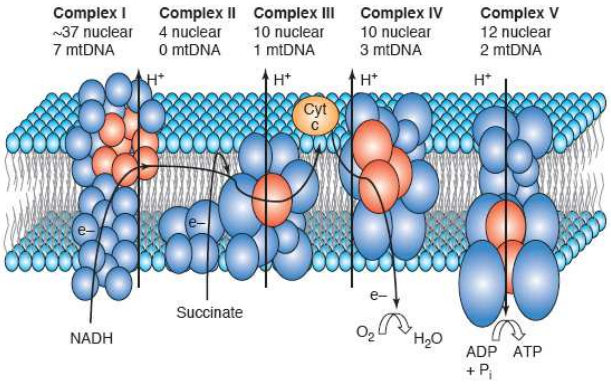

Figure 2. The enzyme complexes of the oxidative phosphorylation (OXPHOS) pathway, which is the main source of energy production in eukaryotic cells. The OXPHOS system results from an intergenomic collaboration; blue subunits are encoded by nuclear genes whereas red subunits are encoded by genes in the mitochondrial genome. (Adapted from Rand *et al.* 2004.)

Given the importance of energy production to individual fitness, there is strong selection for optimal function of the OXPHOS pathway. This means that each subunit has to function well together with other subunits with which it is physically interacting, typically neighbouring subunits of the same complex. For that reason, it has been proposed that genes encoding interacting subunits are under strong selection for co-functionality (Blier *et al.* 2001). This type of co-evolution is known as cytonuclear co-evolution, indicating a co-evolution between genes located in the nucleus and genes of organelles located in the cytoplasm (mitochondria and chloroplasts). Consequently, any mutation that renders this cytonuclear interaction less efficient will be selected against.

Mitochondrial DNA and the speciation process

Maladaptive combinations of mitochondrial and nuclear genes lead to hybrid breakdown and might therefore be an important part of the speciation process (Burton *et al.* 2006; Dowling *et al.* 2008). During the speciation process mating of individuals from diverging populations result in hybrids. Since hybrid lines harbour novel combinations of nuclear and mitochondrial genes, any negative effects on the OXPHOS pathway will cause reduced fitness of hybrids and contribute to reproductive isolation. This is usually not clearly manifested in the first generation of hybrids since F_1 hybrids still inherit one set of co-adapted alleles through the maternally inherited gene copies. In the second generation of hybrids, however, a decrease in fitness is commonly observed. Since these F_2 hybrids may carry the mitochondrial genome of one line and be homozygous for some nuclear

genes of the other population they carry a novel combination of OXPHOS genes that may not function properly (Ellison *et al.* 2008). The potential for negative effects are greater when mtDNA divergence between hybridizing populations are large. Differences in mtDNA may then drive differentiation of co-evolving nuclear genes. Consequently, as many species show great intraspecific differentiation in mtDNA, mtDNA may have a more important role in speciation than currently recognized (Dowling *et al.* 2008).

This study

The main objective of this study is to assess the potential for differences in mtDNA to drive differentiation in coevolving nuclear OXPHOS genes and hence affect the speciation process.
The black-tailed brush lizard (*Urosaurus nigricaudus*), which populates the peninsula of Baja California (fig. 3), is used as a model system. The peninsula of Baja California has a very complex geological history that has greatly affected the regional biota (Lindell 2007). Importantly, temporary seaways formed across the peninsula on multiple occasions throughout its history (Carreño *et al.* 2002). These seaways isolated populations which were subsequently reunited when the seaways disappeared.
Multiple mtDNA breaks that are congruent among species are now evident as a result of these temporary isolation events (Lindell *et al.* 2008). Such patterns have been revealed for the Isthmus of La Paz break (formed >7 Mya), the mid-peninsular break (~7 Mya), the Cabo break (~5 Mya) and the Loreto break (2-3 Mya). Moreover, dispersal of mitochondrial lineages across hybrid zones is limited, suggesting that there is selection against mtDNA introgression (Lindell 2007, Lindell & Murphy 2008, Lindell *et al.* 2008). These patterns of mtDNA breaks are strong even though mitochondrial lineages came back into secondary contact millions of years ago.

Figure 3. Map of the peninsula of Baja California. The range of *U. nigricaudus* indicated in dark grey. The dotted lines represent hybrid zones that have been indicated in previous studies (Lindell *et al.* 2008) The three breaks studied are indicated as Cabo, Isth. (Isthmus of La Paz) and Lor. (Loreto).

This study aims to investigate if the pattern of differentiation in a mitochondrial gene, *cytochrome b (*CYTB*)* is also seen in a nuclear gene, *cytochrome* c_1 *(*CYC1*).* These two genes code for two closely interacting polypeptide subunits known to co-evolve (Willet & Burton 2003, Willet 2006). Thereby it can be assessed whether cytonuclear coevolution affects the OXPHOS pathway and, by extension, speciation.

Materials and methods

Sampling and DNA extraction

Tail tissue samples from 9-10 (17 San Bartolo) individuals were taken from populations approximately 30 km north and south of each of the three hybrid zones. Samples from El Cien (n = 10) and 43 km NW of La Paz (n = 10) represented populations north and south of the Isthmus break, respectively. Samples from San Bartolo (n = 17) and San José del Cabo (n = 9) represented populations north and south of the Cabo break, respectively. Finally, samples from 30 km N of Loreto (n = 10) and Ligui (n = 10) represent populations north and south of the Loreto break, respectively.

For each sample a small slice (1-2 mm) from the thickest end of the tail was cut off. DNA was extracted by a standard phenol-chlorophorm protocol as follows. Cells were digested with 15 µl of proteinase K (20 mg/µl) in 500 µl of Laird's buffer. Tubes were incubated at 37°C over night with gentle rotation and thereafter placed in fridge for storage. In order to remove protein from the samples, 500 µl of phenol-chloroform-isoamylalcohol 25:24:1 was added to each tube. Tubes were vortexed to mix contents and thereafter centrifuged at 13000 rpm for 15 min. The supernatant was subsequently mixed with 500 µl of chloroform to remove any phenol residue and the contents were vortexed and centrifuged at 13000 rpm for 10 min. The supernatant (now containing only buffer and DNA) was transferred to into new microcentrifuge tubes. In order to precipitate the DNA, 40 µl (approximately 0.1 volumes) of 3M NaAc and 800 µl (approximately 2 volumes) of ice cold 95% ethanol was added to each tube. Tubes were inverted a few times to mix everything and thereafter stored at -20°C over night. The tubes were then centrifuged at 13000 rpm for 15 min to concentrate the DNA into pellets. The solution was removed whereafter 200 µl of ice cold 70% ethanol was carefully added to each tube to wash the pellet of remaining salt. Following centrifugation at 13000 rpm for 10 min the ethanol was removed. The pellets were dried by allowing any remaining ethanol to evaporate at room temperature, and later resuspended in 100 µl of ddH_2O. The DNA concentration was measured using Nanodrop™ spectrophotometer. Finally, the samples were diluted to 10 ng/µl as working concentration for subsequent PCR amplification.

Marker development

Conserved primers for the amplification of CYTB and CYC1 in *Urosaurus* were designed based on sequence data available from GenBank. For CYTB, full mitogenomic sequences for *Sceloporus occidentalis*, *Iguana iguana*, and *Anolis carolinenesis* were aligned using BioEdit (Hall 1999). Conserved regions upstream and downstream of CYTB were determined by visual inspection. A forward primer (CYTB-F; 5'-GAA AAA CCA CCG TTG TTA TTC-3') was designed in the glutamine-tRNA gene and a reverse primer (CYTB-R; 5'-CCA TCT TTG GTT TAC AAG ACC-3') was designed in the threonyl-tRNA gene. For the nuclear gene CYC1, conserved exonic regions were determined by visual inspection following determination of orthologous sequences of *Gallus gallus*, *Homo sapiens*, and *Anolis carolinensis* (the closest relative to *Urosaurus* for which large-scale genomic data is available) with Exonerate (Slater & Birney 2005). Two primers, CYC1-F (5'-GGG AAC ACT GTG AAC AGA-3') and CYC1-R (5'-TTT

ACT TCA GTT TTG GGA AGA-3') were designed to amplify a region containing an intron of appropriate size for DNA sequencing (403 bp in *Anolis carolinensis*).

Molecular methods

PCR conditions were tested for optimal amplification of CYTB and CYC1 The PRC mixes (15 µl) contained 4 µl 10 ng/µl DNA, 1.5 µl each of AmpliTaq-Gold 10x buffer and AmpliTaq-Gold $MgCl_2$, 0.75 µl of each primer, 0.15 µl each of 20 µM dNTP and AmpliTaq-Gold enzyme and 6.2 µl ddH_2O per reaction. The PCR program for CYTB started with 5 min at 95°C followed by 39 cycles of 30 seconds at 95°C, 45 seconds at 50°C and 1 min at 72°C, followed by a prolonged extension step of 7 min at 72°C. The PCR program for CYC1 started with 5 min at 95°C followed by 39 cycles of 30 seconds at 95°C, 45 seconds at 52°C and 1 min at 72°C, followed by a prolonged extension step of 7 min at 72°C. PCR reactions were cooled at 4°C following program completion. PCR amplifications were performed on a PTC-225 tetrad thermocycler.

ExoSAP cleanup of samples for sequencing

The samples were cleaned according to the ExoSAP cleanup protocol to remove unincorporated dNTPs and primers. To the wells of a PCR plate, 5 µl of amplified fragments (CYTB or CYC1) was added followed by 1 µl of ExoSAP mix containing 0.2 µl Exo, 0.4 µl SAP and 0.4 µl MilliQ water. The plate was sealed with flat strip lids[1]. The PCR plate was loaded onto PCR a machine and run for 15 min at 37°C followed by 15 min at 80°C.

BigDye sequence reaction

Since the samples were too many to fit forward and reverse sequencing in a 96 well plate a few were excluded since they showed weak bands on the gel or the PCR only worked for one of the markers.
The BigDye sequencing reaction mix consisted of 0.125x Ready reaction mix BigDye, 0.875x BigDye Sequencing Buffer, 0.16x Primer and 2 µl DNA solution to a final reaction volume of 10 µl. The sequence reaction was run on a PTC 225 Tetrad thermal cycler according to a standard program of 30 seconds at 95°C followed by 35 cycles of 25 seconds at 94°C, 15 seconds at 50°C and 2 min at 60°C. The plate was stored in the machine at 4°C over night.

Cleaning of the sequencing reactions with BigDye® XTerminator™ Purification Kit

Sequencing reactions were cleaned using the BigDye® XTerminator™ Purification Kit. A master mix was made by mixing 4.5 ml SAM™ Solution with 1 ml BigDye® XTerminator™ Solution. The mix was vortexed vigorously to ensure optimal mixing. A repeater pipette was used to add 50 µl mix into each well on the PCR plate and the mix was vortexed between each refill of the pipette to to avoid sedimentation of the mixture.
The plate was sealed with aluminium foil, which was tightened using a roller and fingertips, whereafter the plate was vortexed for 30 min at 1800 rpm to make the

[1] ref nr 65.1998.002

solutions mix and make sure all contaminations were removed. The plate was thereafter centrifuged in a Hermle centrifuge at 1000 RCF for 2 min to make the solution sediment forming a top layer containing the cleaned solution and a bottom layer containing all possible contaminants. The plate was stored in a fridge (4°C) until it was run on the in-house ABI 3730 XL sequencer of the Department of Evolutionary Biology.

Sequence analysis

The sequences obtained were checked and edited in Sequencher and thereafter exported as sequence alignments to BioEdit (Hall 1999).
Single sequences were obtained by fusing the forward and reverse sequences. These sequences were thereafter aligned with the reference sequences (*S. occidentalis*, *I. iguana*, and *A. carolinensis* in the case of CYTB and exonic regions of *A. carolinensis*, *G, gallus*, and *H. sapiens* in the case of CYC1.
The first 339 bp were cut and the 340th base was confirmed to be at the right codon position (i.e. correctly aligned with the amino acid sequence). A total of 51 fragments, 642 bp in length, were obtained and translated into amino acid sequences that could also be aligned to the reference sequences of Anolis and Chicken. The two files were saved in FASTA format and opened in MEGA (Tamura *et al.* 2007) where they were converted into MEGA format (.MEG).In MEGA the overall nucleotide variation (π) was calculated for both the nucleotide and the amino acid sequences.

Results

Sequence analysis

CYTB

Nucleotide sequences

The overall variation of the nucleotide sequences was calculated to 0.0767 (S.E. = 0.0068) and the within and between means according to tables 1 and 2, respectively.

Table 1. Within group means for the nucleotide sequences.

Within group means		
Population	**d**	**S.E.**
Cabo N	0.0123	0.0027
Cabo S	0.0065	0.0018
Isthmus N	0.0044	0.0016
Isthmus S	0.0100	0.0024
Loreto N	0.0147	0.0027
Loreto S	0.0097	0.0021

There is no great variation within the groups.

Table 2. Within group means for the nucleotide sequences.

Between group means	
Populations	
Cabo N - Cabo S	0.0797
Isth. N - Isth. S	0.1174
Lor. N - Lor. S	0.0416

The variation is higher between the groups.

Amino acid sequences

The overall variation for the amino acid sequence was calculated to 0.0182 (S.E. 0.0057) and the within and between means according to tables 3 and 4, respectively.

Table 3. Within group means for the protein sequences.

Within group means		
Population	d	S.E.
Cabo N	0.0033	0.0025
Cabo S	0.0012	0.0012
Isthmus N	0.0012	0.0012
Isthmus S	0.0071	0.0040
Loreto N	0.0065	0.0038
Loreto S	0.0019	0.0013

There is no great variation within the groups.

Table 4. Between group means for the protein sequences.

Between group means	
Compared populations	
Cabo N - Cabo S	0.0229
Isth. N - Isth. S	0.0310
Lor. N - Lor. S	0.0069

The variation is higher between the groups.

CYC-1

Nucleotide sequences

The overall variation of the nucleotide sequences was calculated to 0.0172 (S.E. = 0.0027) and the within and between means according to tables 5 and 6, respectively.

Table 5. Within group means for the nucleotide sequences.

Within group means		
Population	d	S.E.
Cabo N	0.0118	0.0034
Cabo S	0.0125	0.0026
Isthmus N	0.0099	0.0027
Isthmus S	0.0123	0.0029
Loreto N	0.0156	0.0038
Loreto S	0.0213	0.0042

There is no great variation within the groups.

Table 6. Between group means for the nucleotide sequences.

Between group means	
Compared populations	
Cabo N - Cabo S	0.0120
Isth. N - Isth. S	0.0125
Lor. N - Lor. S	0.0188

There is no great variation between the groups.

F_{st} results

The values of F_{st} were calculated to compare the differentiation across the three hybrid zones (table 7).

Table 7. Differentiation across hybrid zones (F_{st})

Populations	F_{st}
Cabo N - Cabo S	-0.0141
Isth. N - Isth. S	0.10343
Lor. N - Lor. S	0.0496

The Isthmus break shows the highest differentiation followed by the Loreto break and least differentiation is found at the Cabo break.

Discussion

CYTB

The calculations for nucleotide variation for CYTB show differences between populations, whereas the sequences are rather preserved within each population. As the differences are mirrored in the amino acid sequences it can be concluded that the nucleotide differences give rather large and significant differences in protein structure. It can also be concluded that the differences are maintained within the population and cause differences between populations living as little as a few kilometres apart.

CYC-1

Since there was an evident selection on the CYTB gene there should be a similar selection on the CYC-1 gene since the two genes interact. However, the calculations of CYC-1 give no conclusive results. Nothing indicates that CYC-1 coevolves with the CYTB gene. This may be due to the fact that a non-coding region of the CYC-1 gene was sequenced. Since introns are non-coding there should be only minimal selection pressure on the sequence. Future analysis of a coding and CYTB interacting part is needed to determine if the CYC-1 gene is indeed affected by coevolution.

F_{st}

As seen in table 8, the Isthmus break shows the highest differentiation, as expected. The Cabo break was expected to come next, but instead the Loreto break shows higher differentiations than the Cabo break.

Conclusions and future prospects

It seems necessary to perform the analysis on a greater number of genes in order to determine if the OXPHOS system is affected by cytonuclear co-evolution. It may well be that some genes are under harder selective constraint than others, which is why I did not find as clear differences when examining only small parts of the genes. It is also likely that the effects of the co-evolution is more evident in exons than introns.
Finally it can be concluded that the analyses showed signs of co-evolution, but not as evident as was initially expected.

References

Blier, P.U., Dufresne, F. and Burton, R.S. 2001. Natural selection and the evolution of mtDNA-encoded peptides: Evidence for intergenomic co-adaptation. Trends in Genetics 17: 400-406.

Burton, R.S., C.K. Ellison, & J.S. Harrison. 2006. The sorry state of F_2 hybrids: Consequences of rapid mitochondrial DNA evolution in populations. American Naturalist 168: S14-S24.

Carreño, A.L., Helenes, J. 2002 Geology and ages of the islands. In: Case TJ, Cody ML, Ezcurra E (eds) A New Island Biogeography of the Sea of Cortés. Oxford University Press, New York, New York, p 14-40.

Coyne, J.A., Orr, H.A. 2004. Speciation. Sinauer Associates, Sunderland, Massachusetts.

Dowling, D.K., Friberg, U., Lindell, J. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. Trends in Ecology & Evolution 23: 546-554.

Ellison, C.K., Niehuis O. & Gadau, J. 2008. Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. Journal of Evolutionary Biology 21: 1844-1851.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98

Lindell, J.S. 2007 Genetic Consequences of Plate Tectonics: Cytonuclear Discordance in Phrynosomatid Lizards of Baja California, Mexico. Doctoral thesis. University of Toronto, Toronto, Canada.

Lindell, J., Méndez-de la Cruz, F.R. & Murphy, R.W. 2005. Deep genealogical history without population differentiation: Discordance between mtDNA and allozyme divergence in the zebra-tailed lizard (*Callisaurus draconoides*). Molecular Phylogenetics and Evolution 36: 682-694.

Lindell, J., Méndez-de la Cruz, F.R. & Murphy, R.W. 2008. Deep biogeographical history and cytonuclear discordance in the black-tailed brush lizard (*Urosaurus nigricaudus*) of Baja California. Biological Journal of the Linnean Society 94: 89-104.

Rand, D.M., Haney, R.A., Fry, A.J. 2004. Cytonuclear coevolution: The genomics of cooperation. Trends in Ecology & Evolution 19: 645-653.

Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.

Tamura, K., Dudley, J., Nei, M., & Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Upton, D.E. & Murphy, R.W. 1997. Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) Based on mtDNA sequences: Support for a midpeninsular seaway in Baja California. Molecular Phylogenetics and Evolution, 8: 104-113.

Willett, C.S. & Burton, R.S. 2003. Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. Molecular Biology and Evolution 21:443-453.

Willett, C.S. 2006 Deleterious epistatic interactions between electron transport system protein-coding loci in the copepod *Tigriopus californicus*. Genetics 173: 1465-1477.

Wu, C-I. & Ting, C-T. 2004. Genes and speciation. Nature Reviews Genetics 5: 114-122.