Vol. 18(4): 189–209
FOLIA MALACOLOGICA
ISSN 1506-7629
The Association of Polish Malacologists
& Faculty of Biology, Adam Mickiewicz University
Poznań 2010
doi:10.2478/v10125-010-0016-x

CHARACTERISTICS OF MITOCHONDRIAL DNA OF UNIONID BIVALVES (MOLLUSCA: BIVALVIA: UNIONIDAE). II. COMPARISON OF COMPLETE SEQUENCES OF MATERNALLY INHERITED MITOCHONDRIAL GENOMES OF *SINANODONTA WOODIANA* AND *UNIO PICTORUM*

MARIANNA SOROKA

Department of Genetics, University of Szczecin, Felczaka 3c, 71-412 Szczecin, Poland
(e-mail: marianna.soroka@univ.szczecin.pl)

ABSTRACT: The objective of the study was to sequence the complete female mitochondrial genomes of *Sinanodonta woodiana* and *Unio pictorum*, and to ascertain their intra- and inter-specific variation within all the genes and non-coding regions. Complete sequence of maternally inherited mitochondrial genome of *Sinanodonta woodiana* is 16,243 bp long and is one of the three longest mt genomes in unionids; mt genome of *Unio pictorum* with its 15,761 bp is one of the smallest in the family. Each genome includes 14 protein-coding genes (13 as in most Metazoa, plus one open reading frame F ORF), 2 ribosomal RNA genes and 22 genes for transfer RNA with typical secondary structures, except for tRNA for serine1. *S. woodiana* has 25, and *U. pictorum* 27 non-coding sequences, which constitute 7.8 and 4.9% of mt genome, respectively. The content of A+T bases is 65.8% in *S. woodiana* and 65.1% in *U. pictorum*; the difference between the two species is 21%. Intra-specific variation in *S. woodiana* is three times greater than in *U. pictorum*. Lower nucleotide diversity within the second species suggests a small effective population size of the Polish *U. pictorum*. The two species show a similar usage of the most frequent and least frequent codons; the differences pertain to the usage of both stop codons. Gene *trnS1* (serine, UGA) in unionids is very little variable with respect to tRNA secondary structure (absence of DHU arm), and nucleotide composition, while its inter-specific variation is half lower than in other mt-tRNAs.

KEY WORDS: Complete mitochondrial genome, intra- and inter-specific diversity, *Sinanodonta woodiana*, *Unio pictorum*, Unionidae, Poland

INTRODUCTION

Animal mitochondrial DNA (mtDNA) is small (from 14 to 18 kilobases) and very conservative with respect to its size, gene content and gene arrangement; at the same time, at the level of nucleotide sequence, it evolves five to ten times faster than typical single-copy nuclear DNA (AVISE 1986, AVISE et al. 1987, MORITZ et al. 1987). Animal mtDNA is generally a closed-circular genome containing 37 genes. The common mtDNA composition includes two *rRNA* genes (*srRNA* and *lrRNA*), 13 protein-coding genes for subunits of the enzymes of respiratory chain complexes (*nad1-nad6, nad4L, cox1-cox3, cytb, atp6* and *atp8*) and 22 *tRNA* genes necessary for the translation of the proteins encoded in the mtDNA. Besides, it is haploid and basically contains no introns, repetitive DNA, pseudogenes and mobile elements. All these features have made this small molecule (first analyses were done with the method of restriction fragment length polymorphisms, RFLP) (AVISE et al. 1979) and its single genes (mainly *lrRNA* and *cox1*) popular objects of population genetics and phylogenetic studies for 30 years. Recently, sequencing of whole mitochon-

drial genomes has become increasingly popular; they provide information not only about the level of variation of their component genes, but also on their arrangement and rearrangement in the molecule, and they include also non-coding sequences which constitute ca. 6–10% of the genome.

Because of the very large number of possible gene arrangements, the arrangement of mitochondrial genes provides a very complex character set for phylogenetic inference. Assuming complete positional independence of all genes and two transcriptional orientations, these 37 genes could potentially be arranged in more than 2×10^{52} different ways; thus, the probability of achieving the same arrangement by convergence is negligible. Identical gene arrangements are likely to be shared only as a result of common ancestry, making homoplasy unlikely (BOORE & BROWN 1994b). Although mtDNA evolves rapidly at the level of nucleotide sequence, rearrangements in the gene order appear to be rare (BOORE & BROWN 1994b).

The first mitochondrial genomes were studied in the 1980s; they were vertebrate genomes, of man (ANDERSON et al. 1981), mouse, cattle, rat, chicken and clawed frog *Xenopus laevis* (for review see AVISE 1986 and WENNE 1993). The first studied invertebrates included a fruit fly *Drosophila jakuba* and a sea urchin *Paracentrotus lividus* (for review see WENNE 1993), and – in the 1990s – two nematode species: *Caenorhabditis elegans* and *Ascaris suum* (OKIMOTO et al. 1992) and a bivalve *Mytilus edulis* (HOFFMANN et al. 1992). The number of known complete animal mitochondrial sequences is increasing; at present the international database (Refseq Genbank) includes over 1,700 sequences (PRUITT et al. 2009). Complete mitochondrial gene arrangements have been published for 58 chordate species; 44 of them have identical arrangement of all 37 genes, the remaining 14 (24%) depart from the common vertebrate arrangement mainly in the location of tRNA genes and non-coding regions (BOORE 1999). In some invertebrate taxa the gene order and gene content are less conserved and can vary across phyla, and sometimes even within phyla (BOORE 1999). The observed variation of mtDNA most often involves the number and location of tRNA genes and gene *atp8* (BOORE & BROWN 1994a, GISSI et al. 2008, XU et al. 2010). For example, the gene content has been shown to vary in nematodes which lack *atp8* (OKIMOTO et al. 1992), in most bivalve species which both lack *atp8* and have an extra tRNA for methionine (HOFFMANN et al. 1992, MIZI et al. 2005, ZBAWICKA et al. 2007, ZHENG et al. 2010), and in cnidarians, which have lost nearly all tRNA genes except tRNA for methionine and gained one or two additional genes not found so far in the other mtDNAs, for example a homologue to the bacterial mismatch repair gene *mutS* (BEAGLEY et al. 1998, BOORE 1999; for review see GISSI et al. 2008). Two rRNA genes are always located in mtDNA and their duplications are very rare, but observed in parasitic nematodes *Thaumamermis cosgrovei* and *Strelkovimermis spiculatus*, in the bivalve *Crassostrea gigas* and in the chigger mite *Leptotrombidium pallidum* (for review see GISSI et al. 2008).

Among molluscs, complete mitochondrial genomes are known for 64 species of five of the eight classes within the phylum: Bivalvia (28), Gastropoda (20), Polyplacophora (1), Cephalopoda (13) and Scaphopoda (2), and they show an extraordinary amount of variation in their gene arrangement (see SERB & LYDEARD 2003 and GISSI et al. 2008 for review, BOORE et al. 2004, MIZI et al. 2005, ZBAWICKA et al. 2007, 2010, THEOLOGIDIS et al. 2008). Class Bivalvia is the best studied, with the data pertaining to 28 species. Because of doubly uniparental inheritance (DUI, SOROKA 2010) in some species two (female and male) mitochondrial genomes are present, and nowadays 40 complete and two incomplete mitochondrial genomes are known in bivalves (Table 1). Recently, the studies on this class have been intensified, and 1/2 of the mitochondrial genomes included in GenBank at present were studied in 2008–2010. More than half of the known mitochondrial genomes are those of marine bivalves where, besides DUI (Doubly Uniparental Inheritance of mtDNA), also gene translocation in all gene classes (protein-coding, tRNA, and rRNA) has been observed, as well as gene duplications and losses and even recombination between both genomes, or role-reversal. For example, gene *srRNA* has been duplicated in *Crassostrea* (YU et al. 2008, REN et al. 2009, GenBank sequences), a second tandem copy of *cox2* of a different length is present in *Venerupis* (BRETON et al. 2009), and tRNA genes have been duplicated in three species of *Mytilus* (*trnM*) (HOFFMANN et al. 1992, MIZI et al. 2005, ZBAWICKA et al. 2007), in *Venerupis* (*trnM* and *trnV*) (BRETON et al. 2009) and in *Crassostrea* (*trnQ*) (MILBURY & GAFFNEY 2005). The protein-coding gene *atp8* is absent in 10 bivalve mt genomes, for example in three species of *Mylitus*, two species of *Crassostrea*, in *Venerupis* and *Sinonovacula* (HOFFMANN et al. 1992, SERB & LYDEARD 2003, MIZI et al. 2005, ZBAWICKA et al. 2007, GISSI et al. 2008, ZHENG et al. 2010). Also, the two congeneric oyster species *Crassostrea virginica* and *C. gigas* show significant differences in the gene content and arrangement, with extensive rearrangements of most tRNA genes, including their duplications (MILBURY & GAFFNEY 2005). The first reported case of a split ribosomal RNA in a metazoan has been described in the two above-mentioned species of *Crassostrea*; another unique feature is an inferred translational frameshift within the cytochrome b gene in *C. virginica* (MILBURY & GAFFNEY 2005).

Compared to marine bivalves, their freshwater relatives are much less well studied. Among freshwater Unionidae, comprising more than 600 species (GRAF & CUMMINGS 2007, BOGAN 2008, BOGAN &

Table 1. Complete and incomplete (*) mitochondrial genomes of bivalves available in GenBank

Family	Species	Accession	Length (bp)	Source
Mytilidae	*Mytilus edulis*	AY484747	16740	BOORE et al. 2004
	Mytilus galloprovincialis – F type	AY497292	16744	MIZI et al. 2005
	Mytilus galloprovincialis – M type	AY363687	16626	MIZI et al. 2005
	Mytilus trossulus	NC_007687	18652	BRETON et al. 2006
	Mytilus trossulus – F type	DQ198231	16781	ZBAWICKA et al. 2007
	Mytilus trossulus – M type	DQ198225	16634	ZBAWICKA et al. 2007
	Mytilus trossulus – F type	GU936625	18653	ZBAWICKA et al. 2010
	Mytilus trossulus – M type	GU936626	16578	ZBAWICKA et al. 2010
	Mytilus trossulus – M type	GU936627	17538	ZBAWICKA et al. 2010
Ostreidae	*Crassostrea angulata*	NC_012648	18225	RefSeq
	Crassostrea ariakensis	NC_012650	18414	RefSeq
	Crassostrea gigas	NC_001276	18224	RefSeq
	Crassostrea hongkongensis	NC_011518	16475	YU et al. 2008, REN et al. 2009
	Crassostrea sikanea	NC_012649	18243	RefSeq
	Crassostrea virginica	NC_007175	17244	MILBURY & GAFFNEY 2005
Veneridae	*Venerupis philippinarum* – F type	NC_003354	22676	RefSeq
	Venerupis philippinarum – M type	AB065374	21441	RefSeq
Solecurtidae	*Sinonovacula constricta*	NC_011075	17225	ZHENG et al. 2010
Cardiidae	*Acanthocardia tuberculata*	NC_008452	16104	DREYER & STEINER 2006
Hiatellidae	*Hiatella arctica*	NC_008451	18244	DREYER & STEINER 2006
Pectinidae	*Argopecten irradians*	NC_009687	16221	RefSeq
	Chlamys farreri	NC_012138	21695	WU et al. 2009
	Mimachlamys nobilis	NC_011608	17963	WU et al. 2009
	Mizuhopecten yessoensis	NC_009081	20414	SATO & NAGASHIMA 2001
	Placopecten magellanicus	NC_007234	32115	LA ROCHE et al. 1990
Unionidae	*Anodonta woodiana* – F type*	AB055627	4582	GenBank
	Anodonta woodiana – M type*	AB055626	4585	GenBank
	Cristaria plicata – F type	NC_012716	15712	RefSeq
	Hyriopsis cumingii – F type	NC_011763	15954	RefSeq
	Inversidens japanensis – F type	AB055625	16826	GenBank
	Inversidens japanensis – M type	AB055624	16967	GenBank
	Lampsilis ornata – F type	AY365193	16060	SERB & LYDEARD 2003
	Pyganodon grandis – F type	FJ809754	15848	BRETON et al. 2009
	Pyganodon grandis – M type	FJ809755	17071	BRETON et al. 2009
	Quadrula quadrula – F type	FJ809750	16033	BRETON et al. 2009
	Quadrula quadrula – M type	FJ809751	16970	BRETON et al. 2009
	Venustaconcha ellipsiformis – F type	FJ809753	15976	BRETON et al. 2009
	Venustaconcha ellipsiformis – M type	FJ809752	17174	BRETON et al. 2009
	Unio pictorum – F type	HM014130-3	15760-62	SOROKA & BURZYŃSKI 2010
	***Unio pictorum* – F type**	HM014134	**15761**	Current studies
	***Sinanodonta woodiana* – F type**	HQ283345	**16243**	Current studies

ROE 2008), complete mitochondrial genomes have been described for only eight species, F and M types in *Inversidens japanensis* (GenBank: AB055625 and AB055624), *Pyganodon grandis*, *Quadrula quadrula* and *Venustaconcha ellipsiformis* (BRETON et al. 2009), and only F-types in *Lampsilis ornata* (SERB & LYDEARD 2003), *Hyriopsis cumingii* (GenBank: NC_011763), *Cristaria plicata* (GenBank: FJ986302) and *Unio pictorum* (SOROKA & BURZYŃSKI 2010) (Table 1). F and M genomes in *Anodonta woodiana* (unpublished data, GenBank: AB055627 and AB055626, respectively) are partly known, 4500 bp of each. All known mt genomes in unionids (SERB & LYDEARD 2003, BRETON et al. 2009), except *U. pictorum* (SOROKA & BURZYŃSKI 2010), are based on analyses of single individuals. In the last case studies on complete mtDNA genomes in four specimens of *U. pictorum* from Poland have made it possible for the first time to assess the intra-specific variation which is 0.1% for that species (SOROKA & BURZYŃSKI 2010).

The analyses of eight female mitochondrial genomes in unionids have revealed the presence of 37 genes typical of animals, located on both DNA strands (11 and 26 genes on the heavy and the light strands, respectively), with identical or nearly identical arrangement (BRETON et al. 2009, SOROKA & BURZYŃSKI 2010). A different arrangement of nine genes in the region between *cox2* and *lrRNA* has been observed in *I. japanensis* and *H. cumingii* (Unionidae: Ambleminae: Gonideini) (BRETON et al. 2009). Most recent analyses of unionid F and M genomes indicate, with a high probability, the presence of an additional gender-specific mitochondrial gene (one F-specific and one M-specific open reading frames), with a different location in each genome (BRETON et al. 2009). It is conjectured that these F-specific and M-specific genes play reproductive, transmission and/or gender-specific adaptive roles in the M and F mtDNA genomes in unionids (BRETON et al. 2009). The arrangement of genes in unionid F-genomes is radically different from that in marine bivalves in which transcription of all genes takes place on one DNA strand (BOORE et al. 2004, MILBURY & GAFFNEY 2005, MIZI et al. 2005, ZBAWICKA et al. 2007, BRETON et al. 2009).

Confirmation or rejection of the suspected wide variation of gene arrangement in bivalves necessitates more data on whole genomes in further, different species of the group. So far gene duplication and/or loss has been observed in almost every taxon for which a complete mt genome is available. In all the studied genomes of freshwater bivalves gene *atp8* is present in F-type; it is not found in M-type of *I. japanensis* and *P. grandis* (BRETON et al. 2009). In *L. ornata* its complete copy codes for a protein which is longer than the "typical" *atp8* by extra 13 amino acids on the C-terminus, while in *I. japanensis* this copy is non-functional due to the presence of several stop codons after the first 31 amino acids (SERB & LYDEARD 2003). In *U. pictorum* the gene is functional and codes for 64 amino-acids (SOROKA & BURZYŃSKI 2010). Studying more unionid mtDNA genomes will make it possible to ascertain if the variation in the gene arrangement observed till now in two species of Gonideini is characteristic of this group of bivalves only, or perhaps more widespread, and if the patterns of mtDNA organisation in unionids are more numerous.

Studies on whole mitochondrial genomes include mainly comparative analyses of the number and arrangement of genes, as well as the number and location of non-coding sequences, based on single representatives of species. The first studies on variation of mitochondrial genomes of freshwater bivalves were done on *Pyganodon grandis* using RFLP technique. They revealed a great variation in F and M genomes, of 0.5% and 12% respectively (LIU et al. 1996a, b). However, the only studies on variation in mitochondrial genomes in marine and freshwater bivalves using complete sequencing pertain to *M. trossulus* (ZBAWICKA et al. 2010) and *U. pictorum* (SOROKA & BURZYŃSKI 2010).

The objective of the present study was to sequence and characterise the female mitochondrial genome in another species, *Sinanodonta woodiana* and in more specimens of *Unio pictorum* from various regions of Poland. The results were analysed with respect to the number and arrangement of genes and non-coding regions in the studied genomes. The material included several specimens of each species which made it possible to determine the level of intra- and inter-specific variation of the whole molecule of mtDNA. The obtained data were compared with data on other genomes of marine and freshwater bivalves.

MATERIAL AND METHODS

The material included four specimens of two species: *Sinanodonta woodiana* and *Unio pictorum*. Females of *S. woodiana* (=*Anodonta woodiana*) came from the Odra River near Szczecin (specimen no. 22) and Lake Ślesińskie near Konin (specimen no. 124). Females of *U. pictorum* were also collected in two localities in Poland: Odra River near Szczecin (specimen no. 147) and Krutynia River in the Mazurian Lakeland (specimen no. 310). The material was collected in 2005–2009.

Bivalve sex can be determined exclusively based on observation of mature gonads (PIECHOCKI & DYDUCH-FALNIOWSKA 1993), hence the material was collected only in the spring-summer period when such observations were possible to perform in a reliable way. Sex was determined based on microscope images

of gonad sections. In females large oocytes were visible; in males – fine, shiny spermatozoa.

Total DNA was isolated from the gonads of the females using phenol-chloroform method. In order to obtain full-length mtDNA ca. 1.6 kb long, long-range PCR reactions (LR-PCR) (ZBAWICKA et al. 2007) were performed for each specimen, using DUF1 and UR1 primers which are universal for F type (Table 2). Sequential analyses of mtDNA genomes for one specimen of each species were done with primer walking method (specimens 124 and 147), for the second specimen – with re-PCR reaction (ZBAWICKA et al. 2007). The detailed list of the primers used, conditions of PCR and re-PCR reactions, and sequencing for *U. pictorum* were presented in SOROKA & BURZYŃSKI (2010); the same procedures were applied for *S. woodiana* and the primers used are presented in Table 2 and Fig. 1.

For long-range PCR Phusion High-Fidelity DNA (Finnzymes) polymerase was used according to the manufacturer protocol. All re-amplifications were carried out in 10 µl reaction volumes containing 1:800 dilution of long range PCR as the template DNA, 0.7 µM of each primer, 200 µM nucleotides, 2.0 mM $MgCl_2$, 0.1 unit of DyNAzyme™ EXT DNA polymerase (Finnzymes) and appropriate reaction buffer. After an initial 2-min denaturation at 94°C, 25 cycles were used with denaturation at 94°C for 40 sec, annealing for 30 sec at 51 or 55°C, and extension at 72°C for 55 sec with a final 5 min extension at 72°C. All PCR reactions were performed in Master gradient cycler from Eppendorf. Prior to sequencing PCR products were purified using Exonuclease I and shrimp Alkaline phosphatase (WERLE et al. 1994). The purified PCR product was sequenced using BigDye Terminator chemistry.

Both strands of DNA were sequenced and assembled (consensus) using DNAMAN 5.2.9 software (Lynnon Corporation, Canada).

All bioinformatic analyses were done in MEGA4 (TAMURA et al. 2007). Estimates of the genetic divergence (*K*) were calculated with the Kimura two-parameter model (KIMURA 1980). The divergences in the protein-coding genes both in synonymous (K_s) and non-synonymous (K_a) sites were calculated by the modified Nei-Gojobori method (NEI & GOJOBORI 1986) with Jukes-Cantor correction. Codon usage statistics was also calculated with this programme. The following sequences from GenBank were used for comparative analysis: FJ986302 (*Cristaria plicata*), AY365193 (*Lampsilis ornata*), FJ809754 (*Pyganodon*

Table 2. Primers used in LR-PCR, re-PCR and sequencing of the complete mitochondrial genome of *Sinanodonta woodiana*

Primer	Primer Sequence (5' to 3')
LR-PCR	
DUF1	GGTTTGCGACCTCGATGTTGG
UR1	TTCCTAGTCTTGCCATTCACTGGC
re-PCR	
AW18U	CGATAACGTGGTAAGCTCCCTC
AW23L	GTCAATATCCCAGCAGCTAAAGAC
AW.nn.05F	GGTTACCAAAGGCTCA
AW.nn.05R	CCCAAACCATCTCCTA
AW19U	GGGGGTTTAGTAAAGTAGGTGTGG
AW20L	AAGTATACTCTTCTATCACAACTACACC
AW.nn.01F	TTAGGTTTAGTGAAGAGTTG
AW.nn.01R	CGAACCAAAGATACTAATAC
UT.co1.07	GATATTGTTTTACATGATACTTA
UP.at6.01	CCTAAAGTAATTGGACGAA
AW.nn.03F	CTGGATTTGGGATAATTTCT
AW.nn.03R	CGGACTAAACTCTACTAAATGAA
AW02U	TATTAGCATTATGGTCAGGTTTGATTG
AW01L	CACATAATAGGAGGCTCAGTTTTATTAAC
CO9	CGAGTGCCTTTTTCTTTACG
AW2	GATGCATATACAAAATAGGGTCTCC
AW.nn.02F	AGGTAGTTAAGTCTGTTGTAG
AW.nn.02R	GTAAAGCTACATAACCAGTA
AW11U	CTTATATATGTTGAGCCATACAGTAGTCG
AW12L	GCTACTATCAAAACTCAACTATTGAGC
AW04U	CTTTAGATAGCTTGAGGATTGGC
AW16L	ACTAACCTATATCTTAATTAATGCCTCG
UP.lrn.04	CACGCTCACGCTAACG
UP.srn.01	TGGTGCCAGCAGTCG
AW24U	CATTAGCGTGAATAGTGCGTATTAG
AW25L	AAGCAGCCACCCAATAACG
UP.trn.02	TTTAACATTTTCAGTGTTATG
AW.nn.04R	AGCTAATAAACCAGCAA
AW.nn.04F	TTTAGTTAATTGAGGCG
AW13L	AGTTTGGGACTTTTTCTTATTCTCAC
AW22U	GTAATAATGGTAGAGATTAGGTGTTGG
AW17L	ATAAAGCAACGCTCCTAAGCAC

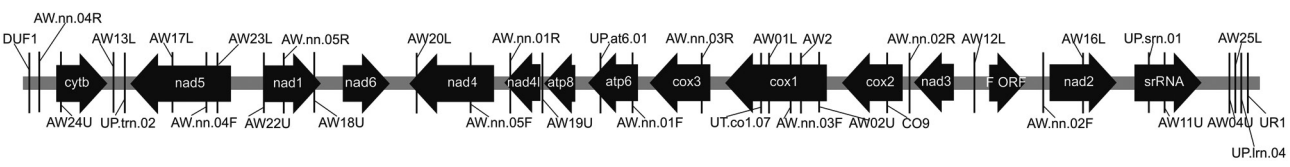

Fig. 1. Sequencing strategy scheme. Primers below the figure are forward primers, above the figure – reverse primers

grandis), FJ809750 (*Quadrula quadrula*), FJ809753 (*Venustaconcha ellipsiformis*) and FJ529186 (*Hyriopsis cumingii*), AB055625 (*Inversidens japanensis*). Dotplot comparisons were carried out using the DNAMAN 5.2.9 programme (Lynnon BioSoft, Canada).

The four complete mtDNA sequences of *S. woodiana* (HQ283345 and HQ283346) and *U. pictorum* (HM014130 and HM014134) obtained in this study were submitted to GenBank.

RESULTS

SINANODONTA WOODIANA

The two F type sequences of mt genomes of *S. woodiana* were 16,243 (Odra River) and 16,242 bp (Lake Ślesińskie) long, had 66% A+T bases and intraspecies nucleotide diversity was estimated at 0.1% due to 20 polymorphic sites (Tables 3, 4). They included 17 nucleotide substitutions and three deletions/insertions of single nucleotides. Eight of these changes affected non-coding regions, eleven occurred in protein-coding genes, and single polymorphic sites were located in gene *lrRNA*. All the nucleotide substitutions were of the transition type (most often C/T), except for one transversion (A/T) in 617 position within the non-coding sequence. The total of 11 polymorphic sites in protein-coding genes were located in five genes among the 14 contained in mtDNA. Single polymorphic sites were observed in genes *cox3* and *nad2*, double in *nad6* and *nad4* and as many as five transitions were found in *nad5* (Table 4). Most (99.8%) mutations in protein-coding genes were synonymous and involved the third position in the codon; single cases of mutational changes in the first and second positions of the codon resulted from amino-acid substitutions isoleucine/valine (gene *nad4*) and threonine/methionine (gene *nad6*), respectively (Table 5). The variation of the five polymorphic genes ranged from 0.1% in genes *nad4*, *cox3* and

Table 3. Average length, base composition of genes and non-coding regions of two *S. woodiana* haplotypes: NCRα – non-coding region alpha, NCRβ – non-coding region beta, NCRγ – non-coding region gamma, H – genes encoded on the heavy strand (anticlockwise), L – genes encoded on the light strand (clockwise)

Gene/Region	Coding strand		Length (bp)	Base Composition (%), L strand			
				T	C	A	G
Non-coding regions	NCRα	–	346	37.2	16.9	36.4	9.5
	NCRβ	–	41	31.7	17.1	46.3	4.9
	NCRγ	–	176.5	48.4	10.2	39.1	2.3
All non-coding regions			1261.5	37.6	16.6	38.6	7.2
All tRNA genes		20L 2H	1443	31.0	17.5	34.7	16.8
rRNA genes	*srRNA*	L	847	27.0	18.9	37.6	16.5
	lrRNA	L	1284	29.6	18.1	37.1	15.2
Both rRNA genes		L	2131	28.6	18.4	37.3	15.7
Protein-coding genes	*nad1*	L	900	33.8	22.2	30.0	14.0
	nad6	L	492	37.5	20.4	30.8	11.3
	nad4	H	1347	42.7	9.5	22.4	25.4
	nad4L	H	297	45.1	6.1	17.8	31.0
	atp8	H	189	38.1	7.4	24.9	29.6
	atp6	H	702	43.2	9.4	22.5	24.9
	cox3	H	780	42.7	11.0	20.6	25.7
	cox1	H	1542	42.3	14.1	20.8	22.8
	cox2	H	681	38.5	10.7	23.3	27.5
	nad3	H	357	42.6	10.9	20.2	26.3
	F ORF	L	261	31.0	24.9	36.8	7.3
	nad2	L	964	35.1	21.2	33.1	10.6
	cytb	L	1149	36.2	22.4	28.1	13.3
	nad5	H	1737	41.0	9.6	24.3	25.0
All protein-coding genes		5L 9H	11457	39.6	14.4	25.1	20.9
Complete genome			**16243**	**28.0**	**22.3**	**37.8**	**11.8**

Table 4. Variable sites in two mitochondrial genomes of *Sinanodonta woodiana*

Individual number	Length (bp)	Variable sites																		
		NCR α 8	NC no. 4 268	NC no. 4 614	NC no. 4 617	*nad6* 1780	*nad6* 1826	*nad4* 2307	*nad4* 2510	NC no. 9 4902	NC no. 9 4910	*cox3* 5334	NCR γ 8505	*nad2* 10104	NC no. 20 10307	*lrRNA* 11449	*nad5* 14534	*nad5* 15000	*nad5* 15129	*nad5* 15555	*nad5* 15897
22	16243	C	T	C	A	C	G	C	T	G	T	T	–	T	G	A	A	C	C	C	T
124	16242	T	–	–	T	T	A	T	C	A	C	C	T	C	A	G	G	T	T	T	C

Table 5. Length of the genes (nt/aa) and sequence divergence within and between two unionid mitochondrial genomes

Gene		*Sinanodonta woodiana*				*Unio pictorum*				*S. woodiana/U. pictorum*		
			K	K_s	K_a		K	K_s	K_a	K	K_s	K_a
rRNA genes	*srRNA*	847	0.000	NA[b]	NA[b]	859	0.001	NA[b]	NA[b]	0.154	NA[b]	NA[b]
	lrRNA	1284	0.001	NA[b]	NA[b]	1302	0.000	NA[b]	NA[b]	0.146	NA[b]	NA[b]
Protein-coding genes[a]	*nad1*	900/300	0.000	NA[b]	NA[b]	900/300	0.000	NA[b]	NA[b]	0.319	n/c[c]	0.104
	nad6	492/164	0.004	0.007	0.003	501/167	0.000	NA[b]	NA[b]	0.344	1.163	0.188
	nad4	1347/449	0.001	0.002	0.001	1347/449	0.000	NA[b]	NA[b]	0.268	1.060	0.118
	nad4L	297/99	0.000	NA[b]	NA[b]	297/99	0.003	0.011	0.000	0.206	0.602	0.112
	atp8	189/63	0.000	NA[b]	NA[b]	195/65	0.000	NA[b]	NA[b]	0.423	0.985	0.305
	atp6	702/234	0.000	NA[b]	NA[b]	702/234	0.000	NA[b]	NA[b]	0.244	0.993	0.106
	cox3	780/260	0.001	0.004	0.000	780/260	0.000	NA[b]	NA[b]	0.147	0.741	0.030
	cox1	1542/514	0.000	NA[b]	NA[b]	1554/518	0.000	NA[b]	NA[b]	0.169	1.041	0.018
	cox2	681/227	0.000	NA[b]	NA[b]	681/227	0.000	NA[b]	NA[b]	0.119	0.353	0.040
	nad3	357/119	0.000	NA[b]	NA[b]	357/119	0.000	NA[b]	NA[b]	0.246	0.992	0.110
	F ORF	261/87	0.000	NA[b]	NA[b]	252/84	0.000	NA[b]	NA[b]	0.814	2.679	0.611
	nad2	964/322	0.001	0.004	0.000	966/322	0.001	0.000	0.001	0.304	1.279	0.147
	cytb	1149/383	0.000	NA[b]	NA[b]	1149/383	0.000	NA[b]	NA[b]	0.260	1.486	0.082
	nad5	1737/579	0.003	0.009	0.000	1734/578	0.000	NA[b]	NA[b]	0.299	1.211	0.142
Total		16243				15761						

[a]Length of protein-coding genes includes the initiation and stop codons
[b]Not applicable
[c]Not possible to estimate evolutionary distance

nad2 to 0.4% in *nad6*. Length polymorphism involving deletion or insertion of single nucleotides in three sites in non-coding sequences no. 4 and NCRγ was observed in the two compared mtDNA molecules (Table 4).

All the codons appeared in *S. woodiana* and the total of 3,799 amino-acids were encoded by 14 genes (Table 6). The most frequently used codons with the frequency of 6–7% were TTA (leucine), TTT (phenylalanine) and ATT (isoleucine). The least frequently used codons were CGC (arginine, 0.11%), followed by TGC (cysteine) and CGG (arginine) with the frequency 0.24% each. Stop codon TAA was more frequently used than TAG. One incomplete stop codon was observed in gene *nad2*; full TAA stop codon is completed by the addition of 3' A residues to the mRNA. In mtDNA of *Sinanodonta* codons terminating with A or T were nearly three times more frequent (74%) than those terminating with G or C (Table 6).

Twenty five non-coding regions of lengths from 2 to 399 bp, and total length 1,261-2 bp, constituting 7.8% of the genome were identified in mtDNA of *S. woodiana* (Tables 3, 7). Significant non-coding regions were called NCRα (346 bp between *nad5* and *trnQ*), NCRβ (41 bp between *trnF* and *nad5*) and NCRγ (176 bp between *nad3* and *trnA* without *trnH*), and their nucleotide content is presented in Table 3. A large non-coding region no. 4 (399 bp) was located between two tRNA genes for isoleucine and valine, with A+T content of 72.4%. The remaining non-coding regions were small and contained from 2 to 45 bp. The analysed specimens had the same number of non-coding sequences with identical length, except for single deletions/insertions in NC no. 4 and NCRγ. In all, four transitions and one transversions were observed within the non-coding sequences (Table 4).

UNIO PICTORUM

Comparison of the two mt genomes of *U. pictorum* is presented in Table 8, and the detailed composition of the genome – in Table 9. The mt genomes were 15760-1 bp long, and the content of A+T bases was 65%. The nucleotide variation between the two specimens of *U. pictorum* from the Odra River near Szczecin (specimen no. 147) and the Krutynia River in the Mazurian Lakeland (specimen no. 310) was 0.03%, and was a result of six polymorphic sites in their mtDNA (Table 8), of which two were new, compared with the earlier-analysed four specimens of the species from Western Pomerania and central Poland

Table 6. Codon usage in 14 mitochondrial protein-coding genes in *S. woodiana* (3,799 codons) and *U. pictorum* (3,805 codons): N – total number of particular codon in all proteins, % – percentage of total codon usage, RSCU – relative synonymous codon usage

Amino Acid	Codon	*Sinanodonta woodiana*			*Unio pictorum*		
		N	%	RSCU	N	%	RSCU
F (Phe)	TTT	253	6.66	1.71	259	6.81	1.74
	TTA	43	1.13	0.29	38	1.00	0.26
L (Leu)	TTA	283.5	7.46	2.94	243	6.39	2.46
	TTG	110	2.90	1.14	161	4.23	1.63
	CTT	90	2.37	0.93	87	2.29	0.88
	CTC	26	0.68	0.27	23	0.60	0.23
	CTA	57	1.50	0.59	68	1.79	0.69
	CTG	12.5	0.33	0.13	10	0.26	0.10
I (Ile)	ATT	224	5.90	1.65	233	6.12	1.62
	ATC	47.5	1.25	0.35	55	1.45	0.38
M (Met)	ATA	149.5	3.94	1.43	119.5	3.14	1.28
	ATG	59	1.55	0.57	66.5	1.75	0.72
V (Val)	GTT	161.5	4.25	1.81	184	4.84	2.10
	GTC	25	0.66	0.28	19	0.50	0.22
	GTA	99	2.61	1.11	90	2.37	1.03
	GTG	71	1.87	0.80	58	1.52	0.66
S (Ser)	TCT	114	3.00	2.45	109	2.86	2.30
	TCC	24	0.63	0.51	24	0.63	0.51
	TCA	51	1.34	1.09	48	1.26	1.01
	TCG	10	0.26	0.21	18	0.47	0.38

Table 6 continued

Amino Acid	Codon	*Sinanodonta woodiana*			*Unio pictorum*		
		N	%	RSCU	N	%	RSCU
P (Pro)	CCT	80	2.11	2.29	67	1.76	1.99
	CCC	11	0.29	0.31	12	0.32	0.36
	CCA	38.5	1.01	1.10	48	1.26	1.42
	CCG	10.5	0.28	0.30	8	0.21	0.24
T (Thr)	ACT	71	1.86	1.56	76	2.00	1.80
	ACC	42	1.11	0.93	35	0.92	0.83
	ACA	53	1.40	1.17	47	1.24	1.11
	ACG	15.5	0.41	0.34	11	0.29	0.26
A (Ala)	GCT	121	3.19	2.36	125	3.29	2.33
	GCC	24	0.63	0.47	30	0.79	0.56
	GCA	50.5	1.33	0.99	47	1.24	0.87
	GCG	9.5	0.25	0.19	13	0.34	0.24
Y (Tyr)	TAT	99	2.61	1.40	95	2.50	1.41
	TAC	42	1.11	0.60	40	1.05	0.59
TERM	TAA	7	0.18	1.00	9	0.24	1.00
	TAG	6	0.16	1.00	5	0.13	1.00
H (His)	CAT	57	1.50	1.56	57	1.50	1.48
	CAC	16	0.42	0.44	20	0.53	0.52
Q (Gln)	CAA	43	1.13	1.34	51	1.34	1.48
	CAG	21	0.55	0.66	18	0.47	0.52
N (Asn)	AAT	75	1.97	1.42	77	2.02	1.38
	AAC	31	0.82	0.58	35	0.92	0.63
K (Lys)	AAA	69.5	1.83	1.42	63	1.66	1.31
	AAG	28.5	0.75	0.58	33	0.87	0.69
D (Asp)	GAT	50	1.32	1.56	48.5	1.27	1.44
	GAC	14	0.37	0.44	19	0.50	0.56
E (Glu)	GAA	46	1.21	0.98	53.5	1.41	1.18
	GAG	48	1.26	1.02	37	0.97	0.82
C (Cys)	TGT	47	1.24	1.68	42	1.10	1.56
	TGC	9	0.24	0.32	12	0.32	0.44
W (Try)	TGA	65	1.71	1.17	63	1.66	1.14
	TGG	46	1.21	0.83	48	1.26	0.86
R (Arg)	CGT	15	0.39	1.02	14	0.37	0.93
	CGC	4	0.11	0.27	4	0.11	0.27
	CGA	31	0.82	2.10	30	0.79	2.00
	CGG	9	0.24	0.61	12	0.32	0.80
S (Ser)	AGT	48	1.26	1.03	61	1.60	1.29
	AGC	12	0.32	0.26	14	0.37	0.30
	AGA	53	1.40	1.14	59	1.55	1.25
	AGG	61	1.61	1.31	46	1.21	0.97
G (Gly)	GGT	92	2.42	1.19	111	2.92	1.45
	GGC	19	0.50	0.25	20	0.53	0.26
	GGA	65	1.71	0.84	72	1.89	0.94
	GGG	133	3.50	1.72	104	2.73	1.36

Table 7. Non-coding regions of maternally inherited mitochondrial genomes in nine unionid species. F ORF fragment is regarded as coding which is not considered in GenBank, except *S. woodiana* and *U. pictorum*. GenBank entries for the remaining species contain a large non-coding region located between *trnE* and *nad2*, and for *H. cumingii* and *I. japanensis* – between *trnE* and *trnW*, which includes F ORF (BRETON et al. 2009)

	Sinanodonta woodiana	*Unio pictorum*	*Cristaria plicata*	*Lamprsilis ornata*	*Pyganodon grandis*	*Quadrula quadrula*	*Venustaconcha ellipsiformis*	*Hyriopsis cumingii*	*Inversidens japanensis*
GenBank accession no	HQ283345	HM014134	NC_012716	AY365193	FJ809754	FJ809750	FJ809753	NC_011763	AB055625
% in genome	7.8	4.9	5.4	6.5	5.8	7.0	6.5	6.8	11.8
Total NC (bp)	1262	771	845	1040	923	1126	1043	1082	1979
NC regions (N)	25	27	23	30	24	29	27	28	27
NCRα (bp)	346	259	288	248	444	348	308	201	1196
NCRα (A+T)	73.6	73.0	71.5	64.1	71.2	64.9	67.8	67.1	57.5

(SOROKA & BURZYŃSKI 2010). In all cases single polymorphic sites were located within coding regions, except genes *nad2* (positions 9328 and 9447) and *srRNA* (positions 10305 and 10441), which had two variable sites each (current studies and SOROKA & BURZYŃSKI 2010). The new polymorphic sites identified in these genes were characteristic of specimen no. 310 from the Mazurian Lakeland. The observed nucleotide substitutions in four cases were of the transition type (3× C/T i G/A), in one case of transversion type (T/A), and two cases involved tRNA genes (*trnA* and *trnL*), protein-coding genes (*nad4L* and *nad2*) and RNA gene for the small ribosome subunit (Table 8). The species showed also size polymorphism of mtDNA molecule. The two female mitochondrial genomes differed in one base in gene *trnH* (Table 8). The observed thymine deletion occurred only in one specimen (147) of *U. pictorum*.

Intraspecies nucleotide diversity (π) was estimated at 0.03% in this study and for other Polish specimens of *U. pictorum* (SOROKA & BURZYŃSKI 2010), but each specimen represented a different haplotype. Even though the number of sequences was small (total of five Polish individuals), and therefore the results must be interpreted cautiously, the observed level of polymorphism was unexpectedly low. The 0.3% (KÄLLERESJO et al. 2005) and 0.4% (SOROKA 2010) variation reported in conserved *cox1* gene of this species shows that substitutions accumulate in mitochondrial sequences of *U. pictorum*. Assuming the neutral evolution model, the observed small nucleotide variation suggests a very low effective population size of the Polish *U. pictorum*, possibly caused by a recent population bottleneck (SOROKA & BURZYŃSKI 2010). The phenomenon is of potential importance for environmental management policies.

Table 5 presents gene lengths and numbers of amino-acids encoded by protein-coding genes, as well as the level of their variation in *U. pictorum* and between *U. pictorum* and *S. woodiana*. The variation in gene *nad4L* (length 297 bp), estimated at K=0.003, involved the third position in the codon and did not cause amino-acid substitution (K_s=0.011, K_a=0.0). The variation in gene *nad2* (length 966 bp), though it involved the third position in the codon, resulted in substitution of glutamate for aspartate (K=0.001, K_s=0.0, K_a=0.001). The parameters for the two specimens of *U. pictorum*, presented in Table 9, differed slightly in the variable regions from corresponding values for the four specimens of the species presented in SOROKA & BURZYŃSKI (2010).

A total of 3,805 amino-acids were encoded by the mt genome of *U. pictorum*, and all codons were present in this species (Table 6). The most frequent codons were TTT (phenyloalanine), TTA (leucine) and ATT (isoleucine), with the frequency exceeding 6%. Excluding termination codons, CGC (arginine) and CCG (proline) were used less than 10 times. Stop

Table 8. Variable sites in two mitochondrial genomes of *Unio pictorum*

Individual number	Length (bp)	Variable sites					
		nad4L	*trnH*	*trnA*	*nad2*	*srRNA*	*trnLUAG*
		3472	8092	8236	9447	10441	12317
147	15760	T	—	A	A	C	T
310	15761	C	T	G	T	T	C

codon TAA was nearly twice as frequent as TAG. Codons ending in A or T were more frequent (73%) than those ending in G or C (Table 6).

The analysed specimens had the same number of non-coding sequences, with identical length, and no variation was observed in these regions. Their total number was 27 and they constituted nearly 5% of mitochondrial genome (Tables 7, 9). The regions NCRα, NCRβ and NCRγ were located in the same places in the genome as in *S. woodiana* and had 259, 36 and 137 bp, respectively. Variation in non-coding sequences in *U. pictorum* was observed only in one specimen from the Szczecin Lagoon and involved two transitions T/A and A/T in region NC no. 6 (SOROKA & BURZYŃSKI 2010).

Analysis of additional specimens of *U. pictorum,* though they originated from geographically very remote localities (ca. 550 km in a straight line), did not increase the originally estimated level of intraspecies variation in mtDNA, which was 0.03% (SOROKA & BURZYŃSKI 2010). In this study the variation in the two examined genomes was also 0.03%, and the observed differences pertained to the third decimal. The level of variation of mitochondrial DNA of the species can be thus reliably estimated based on analysis of two specimens, and examining more specimens

Table 9. Average length, base composition of genes and non-coding regions of two *U. pictorum* haplotypes: NCRα – non-coding region alpha, NCRβ – non-coding region beta, NCRγ – non-coding region gamma, H – genes encoded on the heavy strand (anticlockwise), L – genes encoded on the light strand (clockwise)

Gene/Region		Coding strand	Length (bp)	Base Composition (%), L strand			
				T	C	A	G
Non-coding regions	NCRα	–	259	37.1	23.9	35.9	3.1
	NCRβ	–	36	30.6	11.1	55.6	2.8
	NCRγ	–	137	34.3	18.2	44.5	2.9
All non-coding regions			771	36.0	21.2	39.6	3.2
All tRNA genes		20L 2H	1427	30.6	17.3	34.7	17.4
rRNA genes	*srRNA*	L	859	23.6	21.4	39.1	15.9
	lrRNA	L	1302	28.7	19.5	36.9	14.9
Both rRNA genes		L	2161	26.7	20.2	37.8	15.3
Protein-coding genes	*nad1*	L	900	34.5	21.6	30.3	13.6
	nad6	L	501	35.1	22.8	32.9	9.2
	nad4	H	1347	43.1	10.5	20.3	26.1
	nad4L	H	297	45.1	7.4	17.9	29.6
	atp8	H	195	37.4	5.1	26.2	31.3
	atp6	H	702	44.7	9.1	20.7	25.5
	cox3	H	780	44.3	11.2	17.8	26.7
	cox1	H	1554	42.8	14.0	20.3	22.9
	cox2	H	681	41.1	10.5	22.3	26.1
	nad3	H	357	44.6	8.1	20.4	26.9
	F ORF	L	249	26.5	27.4	37.4	8.7
	nad2	L	966	32.3	21.7	34.9	11.1
	cytb	L	1149	36.5	21.8	27.8	13.9
	nad5	H	1734	42.5	9.7	23.6	24.2
All protein-coding genes		5L 9H	11382	40.1	14.4	24.5	21.0
Complete genome			**15761**	**26.5**	**23.1**	**38.6**	**11.8**

does not contribute much to the general characteristics of the mitochondrial genome. Analyses of whole mitochondrial genomes of a few specimens of *U. pictorum* showed a highly conservative character of non-coding regions within the species, variation of four out of the 14 protein-coding genes, both ribosomal RNA genes and three out of the 22 transfer RNA genes (this study and SOROKA & BURZYŃSKI 2010).

DISCUSSION

MITOCHONDRIAL GENOME ORGANISATION

The studied mitochondrial genomes of *Sinanodonta woodiana* and *Unio pictorum* contain the same genes in identical arrangement, which is also typical of freshwater American Ambleminae (Fig. 2) (BRETON et al. 2009). The genes are arranged in the following order (bolded genes are transcribed from the opposite heavy strand): *trnQ, trnC, trnI, trnV, trnL2(UAA), nad1, trnG, nad6,* ***nad4, nad4L, atp8, trnD, atp6, cox3, cox1, cox2, nad3, trnH,*** *trnA, trnS1(UGA), trnS2(UCU), trnE,* F ORF, *nad2, trnM, trnW, trnR, srRNA, trnK, trnT, trnY, lrRNA, trnL1(UAG), trnN, trnP, cytb, trnF* and ***nad5***. The gene order detected only in Gonideini (*Hyriopsis cumingii* and *Inversidens japanensis*) is slightly different. The fragment between ***cox2*** and *trnW* is organised in the following way: ***trnH,*** *trnS1, nad2, trnM,* ***nad3****, trnA, trnS2, trnE,* F ORF (BRETON et al. 2009).

The female genomes of *S. woodiana* and *U. pictorum* differ in the size of mtDNA, which results mainly from the number of non-coding sequences, since eight protein genes show equal lengths in the two species, and the remaining protein-coding genes, ribosomal RNA genes and transfer RNA genes are very similar in length, except for *nad2* (Tables 3, 5, 9). *U. pictorum* has a very small (15761 bp) and compact genome, and only that of *Cristaria plicata* is by 50 bp smaller, with a comparable number of non-coding sequences (Tables 1, 7). The genome of *S. woodiana* is among the largest female genomes among unionids, and is by 483 bp larger than that of *U. pictorum*; only that of *Inversidens japanensis* is by 584 bp larger. Compared to the marine Mytilidae, the difference in length is more apparent (500-3,000 bp), no doubt due to much longer non-coding sequences in the latter family (BOORE et al. 2004, MIZI et al. 2005, ZBAWICKA et al. 2007, 2010, CAO et al. 2009).

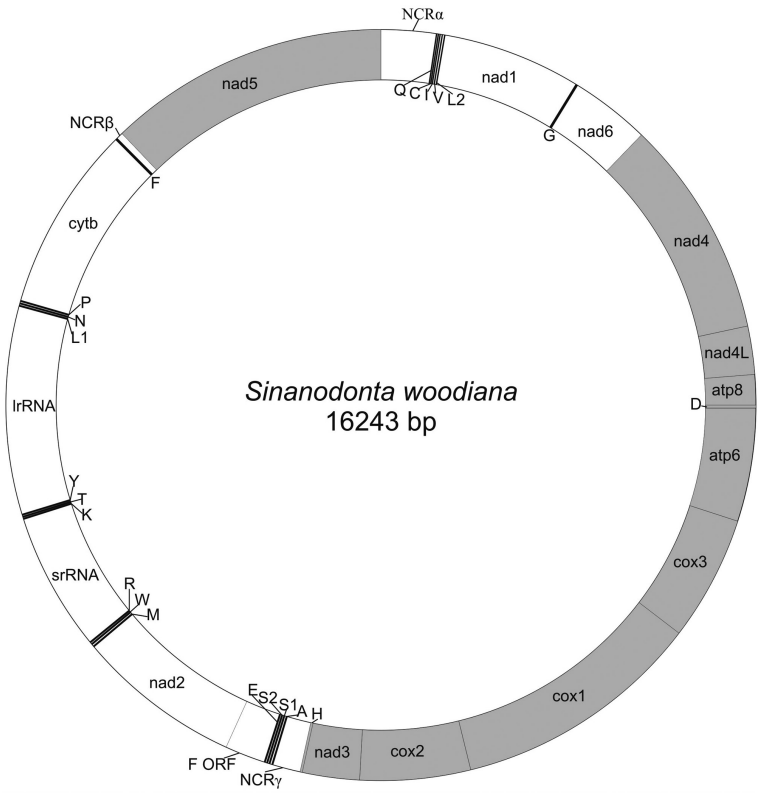

Fig. 2. Genetic map of the mitochondrial F genome of *Sinanodonta woodiana*: the grey colour shows the genes encoded on the heavy strand, the white colour shows the genes encoded on the light strand. All tRNA genes are located on the light strand except *trnH* and *trnD*. Transfer RNA genes are designated by the one-letter amino acid code for the corresponding amino acid, with two *trnL* and the two *trnS* differentiated on the basis of their anticodon

Female mitochondrial genomes of *S. woodiana* and *U. pictorum* contain 65.8 and 65.1% A+T bases, respectively, and the values are comparable to those found in other species of unionoids: 62.4% in *L. ornata* (SERB & LYDEARD 2003), 62% in *Q. quadrula* and *V. ellipsiformis*, 64% in *P. grandis* (BRETON et al. 2009) and 67% in *S. constricta* (ZHENG et al. 2010). The A+T content in marine bivalves is somewhat smaller: in four species of *Mytilus* and two of *Crassostrea* it is ca. 62% (BOORE et al. 2004, MILBURY & GAFFNEY 2005, MIZI et al. 2005, ZBAWICKA et al. 2007). A higher A+T content has been observed in other molluscs, such as black chiton *Katharina tunicata*, 69.0% (BOORE & BROWN 1994a); land snail *Albinaria coerulea*, 70.7% (HATZOGLOU et al. 1995), or scaphopod *Graptacme eborea*, 74.1% (BOORE et al. 2004). This variation in A+T content is among the highest observed within a phylum and reflects high heterogeneity of molluscan mtDNA (BOORE 1999, BOORE et al. 2004, MILBURY & GAFFNEY 2005, MIZI et al. 2005, ZBAWICKA et al. 2007)

Nucleotide variation of whole mitochondrial genomes in *S. woodiana* and *U. pictorum* is 21% and thus comparable to the level of variation of other female genomes in Unionidae: 25% for *P. grandis* (FJ809754), *Q. quadrula* (FJ809750) and *V. ellipsiformis* (FJ809753). Variation between female and male mt genomes in Unionidae is extremely high and amounts to 50–52% (DOUCET-BEAUPRÉ et al. 2010), while in marine bivalves of the genus *Mytilus* it is 20–26% (MIZI et al. 2005, ZBAWICKA et al. 2007, 2010).

Intra-specific variation in *S. woodiana* is more than three times greater than that observed in *U. pictorum*. This low nucleotide diversity in the native *U. pictorum* indicates a small effective population size of this species in Poland (SOROKA & BURZYŃSKI 2010). The wide variation in *S. woodiana*, which was introduced in Poland in the 1980s, suggests that the source populations were abundant and genetically diverse. Earlier studies on single mitochondrial genes (*cox1*, *cytb*) in numerous specimens of the species showed the absence of variation in such genes (SOROKA 2005, 2008, 2010), like in this study.

NON-CODING REGIONS

The studied unionid species have an almost identical location of non-coding regions though they differ in their length and in the percentage they constitute in the genome. The very compact genome of *U. pictorum* has 27 shorter non-coding regions, while one of the largest unionid genomes, that of *S. woodiana*, contains 25 longer non-coding regions, and the two species have 5 and 8% non-coding sequences in their mtDNA, respectively. Excluding the regions NCRα, NCRβ and NCRγ, the mean size of non-coding region is 14 bp in *U. pictorum* and 32 bp in *S. woodiana*. The remaining known female mitochondrial genomes in unionids have 23 to 30 non-coding regions, which constitute from 5 to 12% of the genome (Table 7).

A typical main control region (CR, a non-coding region of DNA involved in replication of the mt genome) in animal mitochondrial genomes is (i) the longest non-coding region, (ii) a region with relatively high A+T content, (iii) a region in which repetitive elements and secondary structures frequently occur, and/or (iv) a region associated with abrupt changes in base composition bias (BOORE 1999, SERB & LYDEARD 2003, CAO et al. 2004, KUHN et al. 2006, BRUGLER & FRANCE 2008). The NCRα, NCRβ and NCRγ in *S. woodiana* and *U. pictorum* correspond to the respective regions shared by other known F and M genomes in Unionidae (BRETON et al. 2009, DOUCET-BEAUPRÉ et al. 2010), but differ in their length and nucleotide composition (Table 7). Interestingly, the change of transcription orientation occurs within all of them (SERB & LYDEARD 2003, BRETON et al. 2009), separating parts of genomes with different compositional biases (SOROKA & BURZYŃSKI 2010), which makes them candidates for CR mtDNA.

NCRα in all known unionid species except *S. woodiana* is the longest non-coding sequence in the genome, with A+T content in four species of ca. 72%, or ca. 66% in another four, and only exceptionally 58% (*I. japanensis*). The next longest NCRγ has a higher A+T content of 88 and 79% in *S. woodiana* and *U. pictorum*, respectively. *S. woodiana* has a different, longest non-coding region with a high A+T content (over 72%), with no counterpart in other genomes. In *L. ornata* NCRβ (136 bp, between *nad5-trnF*) is exceptionally long and also has a high A+T content (76.8%) but the control region has not been localised in this species (SERB & LYDEARD 2003). The main control region has not been identified in the veneroid clam *Venerupis philippinarum* (BRETON et al. 2009) and in the scaphopod *Graptacme eborea* (BOORE et al. 2004), but has been characterised in mytilid species (CAO et al. 2004, 2009, MIZI et al. 2005). The CR in species of *Mytilus* is the largest non-coding region (1,150 bp), it is capable of producing characteristic secondary structures but contains only 60–61% A+T, so the above-average A+T content is not mandatory in the CR (CAO et al. 2004, 2009, MIZI et al. 2005). In addition, the *Mytilus* CR contains conserved motives thought to play an important role in replication and transcription, and it is broadly similar to the mammalian control region (CAO et al. 2004).

Comprehensive studies on the control regions and origins of replication mtDNA, considering many criteria: length, repeats/secondary structure, A+T content and AT-skew, have not provided definitive solutions, and the locations of control regions could be variable among unionid bivalve mitochondrial genomes (BRETON et al. 2009). Two of the three shared major non-coding regions (NCRα and NCRγ) are potential heavy-strand control regions (O_H) for regulat-

ing replication and/or transcription, but multiple and potentially bidirectional light-strand origins of replication (O_L) are present in unionid mitochondrial genomes (BRETON et al. 2009). Localisations of O_H and O_L in *I. japanensis* and *H. cumingii* are the most reliable, since they meet the criteria of length, A+T content, repeat/secondary structure and also AT-skew values (negative or positive). In these species O_H would be located between *nad5-trnQ* (=NCRα) and O_L would be located between *nad3* and *cox2* containing a non-coding region which on the heavy strand is capable of forming a stem-loop structure (BRETON et al. 2009). This non-coding region is 104 bp long and has 67% A+T content; NCRγ located nearby (between *nad3* and *trnA*) is marginally longer (111 bp) and has a significantly lower A+T content (only 59%). Besides, BRETON et al. (2009) do not exclude that the role of O_L in unionids may be played by tRNA genes, and more specifically the heavy strand containing *trnD* in *P. grandis, Q. quadrula* and *V. ellipsiformis*. There are reports that heavy strand harbouring tRNA genes can sometimes function as an O_L by forming alternative secondary structures other than their classical cloverleaf structures, so they could have an alternative O_L function (SELIGMANN et al. 2006, SELIGMANN 2008). Besides, this alternative O_L can frequently evolve and disappear in mtDNA sequences, which is very likely in F and M genomes of some unionid species (SELIGMANN et al. 2006, BRETON et al. 2009). Localisation and organisation of the control region in the scaphopod *G. eborea* whose mtDNA (14,492 bp) is very unusual in lacking any large non-coding regions, is even more interesting. This species has fewer non-coding nucleotides than any other mtDNA studied to date, with the largest non-coding region only 24 bp long (BOORE et al. 2004). Whatever regulatory elements may be present in *G. eborea*, they are apparently short, dispersed, and/or rapidly changing. Thus among molluscs the control region of mtDNA is not conservative with respect to its localisation, nucleotide similarity or potential secondary structures, which renders its studies difficult and requires an individual approach to each newly studied mitochondrial genome.

PROTEIN-CODING GENES

The number of 13 genes coding for the components of the respiratory chain with their identical localisation (with few exceptions) is a very constant character of vertebrate mitochondrial genomes (DESJARDINS & MORAIS 1990, BOORE 1999, GISSI et al. 2008). On the other hand, a wide variation is observed among bivalves (see Introduction), both with respect to the different arrangement of these 13 protein-coding genes in each family, and to duplication (*cox2* in *V. philippinarum*), elongation (*cox2* in Unionidae M genomes), shortening (*atp8* in *V. philippinarum*), or absence (*atp8* in six species of *Crassostrea*, tree of *Mytilus* and *S. constricta*) of single genes in the mitochondrial genomes. However, unionid mitochondrial genomes have the same number and identical arrangement of the genes, and direction of their transcription (Fig. 3), except *H. cumingii* and *I. japanensis* representing the subfamily Ambleminae and the tribe Gonideini. These two species have a different arrangement of genes between *cox2* and *srRNA*, with the same direction of transcription of all the genes being maintained except the two ribosomal RNA genes in *I. japanensis*.

An exceptional case is the presence of an additional protein-coding gene in all the studied unionid mitochondrial genomes (BRETON et al. 2009), including *S. woodiana* and *U. pictorum* (Tables 3, 9). The gene has not been completely identified and is provisionally termed F ORF and M ORF, for female and male genomes, respectively. The gender-specific ORFs could be responsible for the different mode of transmission of the mtDNA and/or gender-specific adaptive functions of the M and F mtDNA genomes in unionid bivalves (BRETON et al. 2009). The activity of this protein has been demonstrated in female and male gonads of *Venerupis ellipsiformis* (BRETON et al. 2009). In female genomes of *L. ornata* and *V. ellipsiformis* (the most closely related species) ORF regions show 60% amino acid identity and the divergence at the nucleotide level follows the pattern expected for a protein-coding gene under purifying selection, where third and second codon positions exhibit the largest and smallest numbers of substitutions, respectively (BRETON et al. 2009). The length of this gene has been found to vary much, and consequently also the number of encoded amino-acids varies, ranging from 66 to 92 in seven female genomes in Unionidae (BRETON et al. 2009). Also in the two studied species the length of F ORF differs (Table 5); it shows only 33% amino acid identity and the highest nucleotide (K=0.814) and amino-acid variation (K_s and K_a) among all the coding mitochondrial genes.

In the two studied unionid genomes, eight genes have identical lengths in the two species, and three of them: *cox2, cox3,* and *nad3* are identical with respect to the number of encoded amino-acids also in *C. plicata*, *H. cumingii*, *L. ornata, P. grandis, Q. quadrula* and *V. ellipsiformis*. Genes coding for subunits of cytochrome c oxidase (I, II and III) show the smallest level of nucleotide variation, below 20%. The variation of the remaining genes, excluding F ORF, varies from 21% for *nad4L* to 42% for *apt8*. Gene *atp8* has the highest (excluding F ORF) aminoacid variation (K_a=0.305), indicating either a very relaxed selection or the influence of diversifying adaptive evolution. Overall, the genes coding for NADH dehydrogenase subunits are rich in synonymous substitutions (K_s>0.990, all except *nad4L*) whereas those coding for subunits of cytochrome c oxidase and cytochrome b

are very well conserved at amino-acid level (K_a<0.1). Again, similar patterns of polymorphisms have been observed in *M. galloprovincialis* (MIZI et al. 2005) and *M. trossulus* (ZBAWICKA et al. 2007) indicating dominant role of functional constraints.

In this study no variation was observed in the most frequently studied mitochondrial genes: *cox1*, *cytb* and *lrRNA* (except one polymorphic site in *S. woodiana*) in *S. woodiana* and *U. pictorum*. For *U. pictorum* the estimated level of variation of gene *cox1* is 0.4%, and the differences are exclusively synonymous, and variation of gene *cytb* is 0.6%, where 1/4 differences were non-synonymous substitutions (SOROKA 2010). Studies on more numerous specimens of *S. woodiana* confirm the absence of variation in the two genes (SOROKA 2010), in spite of the observed intra-specific variation of the whole mitochondrial genome which is wide, compared to *U. pictorum*.

Genes *nad4* and *nad4L* in mtDNA of *S. woodiana* and *U. pictorum* show an 8 bp overlap. The phenomenon has been observed in all known F genomes in Unionidae, with the same accuracy of overlapping 8 bp in that site (except *I. japanensis*) (BRETON et al. 2009, SOROKA & BURZYŃSKI 2010). The two genomes of *S. woodiana* and *U. pictorum* show a 1 bp overlap between genes *nad2* and *trnM* while the overlap in that site in *I. japanensis* is as long as 9 bp. In the remaining studied genomes either two nucleotides overlap in that site, like in *P. grandis*, or the two genes adjoin (*C. plicata* and *L. ornata*), or, like in *Q. quadrula*, there may be even seven nucleotides forming a non-coding sequence between the two genes.

CODON USAGE OF PROTEIN-CODING GENES

The 14 protein-coding genes of the mitochondrial genome of *S. woodiana* code for 3,819 amino-acids, compared to 3,805 in *U. pictorum*. In both species leucine is the most frequent amino-acid, followed by phenylalanine and isoleucine; arginine is the least frequent. Codon usage in these species is very similar: the same three most frequent codons with small differences in their percentage, the same frequency of TCC (serine) and CGC (arginine) codons, with the last codon being the least frequent. The same codons (TTT, TTA and ATT) are the most frequent also in other bivalves: *L. ornata* (SERB & LYDEARD 2003), *M.*

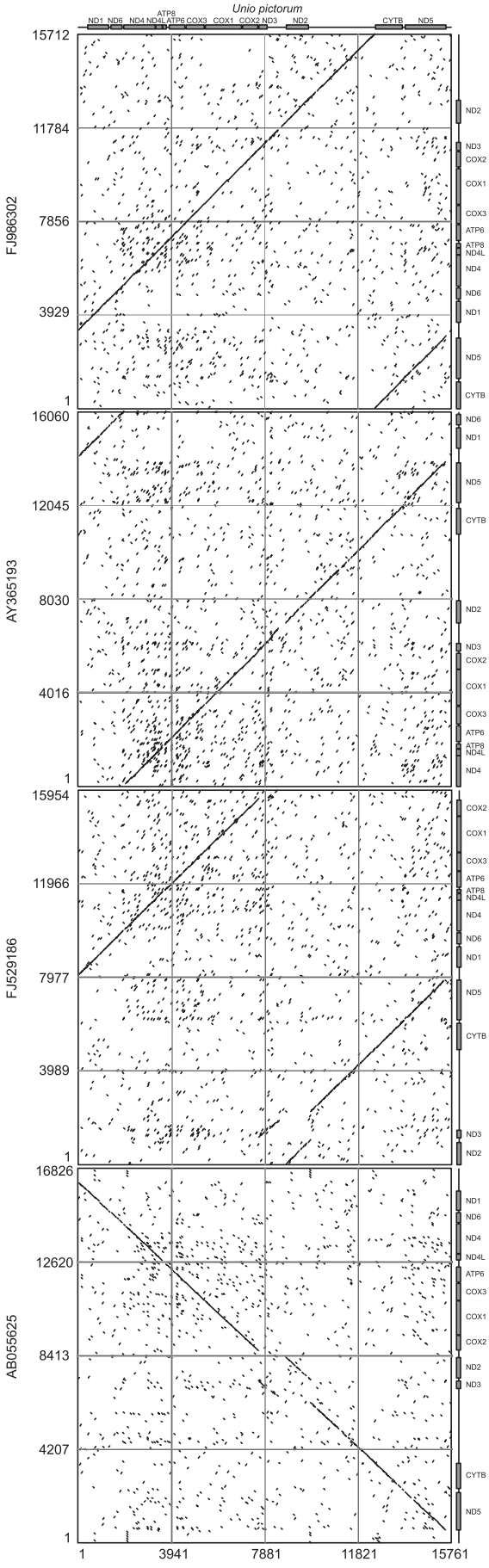

Fig. 3. Dotplot comparison of mt genomes of *Unio pictorum* and other unionid: *Cristaria plicata* (FJ986302), *Lampsilis ornata* (AY365193), *Hyriopsis cumingii* (FJ529186), and *Inversidens japanensis* (AB055625). Protein-coding genes are shown for reference, as annotated in the relevant GenBank records. The sequences deposited in GenBank differ in the origin and orientation but are mostly co-linear – they represent two genome organizations, single gene inversion involving *nad2* and *nad3* genes being the only difference between the two architectures

edulis (BOORE et al. 2004), *M. galloprovincialis* (MIZI et al. 2005), *M. trossulus* (ZBAWICKA et al. 2007), *C. virginica* and *C. gigas* (MILBURY & GAFFNEY 2005), gastropods *Cepaea nemoralis* (TERRETT et al. 1996) and *Pupa strigosa* (KURABAYASHI & UESHIMA 2000) as well as scaphopod *Graptacme eborea* (BOORE et al. 2004). The most frequent codons in most vertebrates are different: CTA in carp *Cyprinus carpio* (CHANG et al. 1994) and man (INGMAN et al. 2000) and ATT in clawed frog *Xenopus laevis* (ROE et al. 1985) and zebrafish *Danio rerio* (BROUGHTON et al. 2001).

Bivalve species differ in their least frequent codons, except codon CGC (SERB & LYDEARD 2003, BOORE et al. 2004, ZBAWICKA et al. 2007). For example codon CCG occurs in F genomes at a frequency of ca. 0.24 in *S. woodiana* and *U. pictorum*, 0.43 in *L. ornata* (SERB & LYDEARD 2003), 0.68 in *M. edulis* (BOORE et al. 2004) and *M. trossulus* (ZBAWICKA et al. 2007) and 0.72 in *M. galloprovincialis* (MIZI et al. 2005), whereas the least frequently used codons in mtDNA of *M. trossulus* are ACC, CCC, CGC and TCC (ZBAWICKA et al. 2007). The four least used codons in *M. galloprovincialis* are TCC, CGC, ACC and ACG (MIZI et al. 2005).

Of the two mitochondrial stop codons, TAA is more frequent in female genomes, and TAG in male genomes (MIZI et al. 2005, ZBAWICKA et al. 2007). Only female genomes of *L. ornata* contain twice as many stop codons TAG (SERB & LYDEARD 2003). In *U. pictorum* all stop codons are complete and TAA is nearly twice more frequent than TAG. The incomplete stop codon in gene *nad2* is observed in *S. woodiana* (T · ·), is also present in that site in *C. plicata*, and TAA stop codon is completed by the addition of 3' A residues to the mRNA in the two species. Other incomplete terminal codons have been found to occur in two genes in *C. virginica* (*cox3* and *nad4L*) and one gene in *C. gigas* (*nad4L*) (MILBURY & GAFFNEY 2005).

Codons ending in A or T are more frequent in *S. woodiana* and *U. pictorum* (ca. 74%) than those ending in G or C. This bias is also evident in the relative synonymous codon usage (RSCU) statistics for 4-fold degenerate site (Table 6). In other bivalves also most codons end with A or T (65% in *C. virginica*, 67% in *L. ornata* and 71% in *C. gigas*), which is consistent with the typical invertebrate bias favouring codons ending in A and T (SERB & LYDEARD 2003, MILBURY & GAFFNEY 2005).

RIBOSOMAL RNA GENES

The two ribosomal RNA genes in *U. pictorum* are longer than in *S. woodiana* by 12 and 18 bp, respectively for *srRNA* and *slRNA* (Table 5). Gene *srRNA* in *U. pictorum* (859 bp) is the longest among Unionidae (SERB & LYDEARD 2003, BRETON et al. 2009), the longest genes *lrRNA*, exceeding 1,300 bp, are found in *L. ornata* (SERB & LYDEARD 2003), *I. japanensis* and *U. pictorum*.

In Unionidae the content of A+T bases is greater in *lrRNA* (ca. 64%) than in *srRNA* (ca. 62%), though in *C. plicata* the values are similar in the two genes (each 64%), and in *I. japanensis* the values are smaller (58 and 56%, respectively). The difference in the two ribosomal RNA genes betwen *S. woodiana* and *U. pictorum* is ca. 15% (Table 5), while considering all nine known unionid genomes the variation is slightly greater in *lrRNA* than in *srRNA* and amounts to somewhat over 19 and 18%, respectively. The smallest differences in the two ribosomal RNA genes are those between *L. ornata* and *V. elipsiformis* (0.063 and 0.073, respectively for *srRNA* and *lrRNA*), the most dissimilar species is *I. japanensis*. In Unionidae the differences between F and M haplotypes of gene *srRNA* are 23–29% and 31–35% for *lrRNA* (SOROKA 2010), while in marine bivalves *Mytilus* they are lower – ca. 15% for both ribosomal RNA genes (MIZI et al. 2005, ZBAWICKA et al. 2007).

The large subunit rRNA gene is split into two fragments in two species of *Crassostrea* (Bivalvia: Ostreoidea), a phenomenon not previously observed in metazoan mitochondrial DNA (MILBURY & GAFFNEY 2005). The two species differ in localisation of *lrRNA*, 5' half in the genome and the level of variation of the two regions of the gene. The *lrRNA* 3' half with identical location in *C. gigas* and *C. virginica* shows high nucleotide identity between these two species (82%), while the 5' half shows only 65% identity (MILBURY & GAFFNEY 2005).

TRANSFER RNA GENES

The tRNA genes are the most evolutionarily mobile elements of the animal mitochondrial genome and variation in mitochondrial tRNA gene organisation has been found in many divergent taxa (BOORE 1999, MUELLER & BOORE 2005, BLEIDORN et al. 2006, PODSIADŁOWSKI 2006, PODSIADŁOWSKI et al. 2006). Mitochondrial genomes of *S. woodiana* and *U. pictorum* include 22 transfer genes each; their localisation in the genome is identical, and the same as in other species of Unionidae (SERB & LYDEARD 2003, DOUCET-BEAUPRÉ et al. 2010). In both studied genomes most tRNA genes are located on the light strand, only t*rnH* (histidine) and *trnD* (aspartate) are on the heavy strand with most protein-coding genes (Fig. 2, Tables 3, 9). The length of tRNA genes ranges from 60 to 71 bp and the mean GC content varies between 19.0% in *S. woodiana* and *trnW* (tryptophane) and 46.8% in *U. pictorum* and *trnY* (tyrosine). In the genomes of the remaining unionids the lengths of tRNA genes are the same, but the GC content is more similar and ranges from 35 to 37% (DOUCET-BEAUPRÉ et al. 2010).

Twenty-two sequences identified in *Sinanodonta* and *Unio* fold into tRNA secondary structures and possess their correct anticodons. Nearly all of the putative tRNAs have a normal cloverleaf structure: a

seven-member amino-acid acceptor arm, a five-member anticodon stem, and a seven-member anticodon loop, DHU stem and loop, and TΨC stem and loop. However, in *S. woodiana* and *U. pictorum*, in 41 and 14% respectively, tRNAs are present with eight-member amino-acid acceptor arm, of which two are identical in the two species, *trnS2* (UCU) (serine) and *trnT* (threonine). Only gene *trnY* in both species has a six-member amino-acid acceptor arm. In both studied species the DHU arm of *trnS1* (UGA) is unpaired. The absence of a DHU arm in this gene is even characteristic of unionid bivalves, since such form occurs in seven other female and male genomes (Fig. 4) except male genome of *I. japanensis* (SERB & LYDEARD 2003, DOUCET-BEAUPRÉ et al. 2010). This is not unusual, since the DHU arm encoded by the mitochondrial *trnS1* is unpaired in other metazoan taxa (HOFFMANN et al. 1992, YAMAZAKI et al. 1997, TOMITA et al. 2002, BLEIDORN et al. 2006, PODSIADLOWSKI 2006, PODSIADLOWSKI et al. 2006). Even more tRNAs with unpaired DHU arms have been observed in *L. ornata* (also *trnS2*) (SERB & LYDEARD 2003), *Q. quadrula* (also lysine *trnK*) and two tRNAs in each known unionid male genome (Table 1), however always in threonine *trnT* in *P. grandis*, *Q. quadrula* and *V. ellipsiformis* (DOUCET-BEAUPRÉ et al. 2010). The unpaired DHU arm has been found also in the marine bivalve genus *Mytilus*, in one or two genes (HOFFMANN et al. 1992, MIZI et al. 2005) and in the scaphopod *Graptacme eborea* in *trnR* (arginine) and *trnS1* (BOORE et al. 2004). *S. woodiana* and *U. pictorum* have typical TΨC arm and loop in all tRNAs, but the TΨC loop is missing in the *L. ornata* tRNA for cysteine (SERB & LYDEARD 2003) and in few gastropod tRNAs, where the arm either is missing or reduced (YAMAZAKI et al. 1997).

Gene *trnS1* (UGA) in freshwater bivalves is characterised not only by coding identical secondary structure with the unpaired DHU arm (Fig. 4) but also by very similar nucleotide composition (Table 10). Its sequence is identical in *S. woodiana* and *U. pictorum*, *L. ornata* and *V. ellipsiformis* and differs in one substitution in *P. grandis* and *H. cumingii* (Fig. 4). For the remaining species the variation ranges from 0.017 to 0.055, except *I. japanensis*, whose *trnS1* shows exceptionally great differences, exceeding 0.500, relative to the other species of Unionidae, and the smallest content of GC pairs. The mean variation of *trnS1* for the nine unionid species is 0.156 and, excluding *I. japanensis*, only 0.041 and involves only substitutions within loops. For other tRNA genes interspecies diversity among unionids is up to 11%, though like in the case of *trnS1* there are no differences in threonine *trnT* between *S. woodiana* and *U. pictorum* (SOROKA 2010).

Mispairing between bases in the stem (T·T and T·G) occurs in each tRNA of the studied species and is consistent across several taxa (HOFFMANN et al. 1992, YAMAZAKI et al. 1997, SERB & LYDEARD 2003, BOORE et al. 2004, MIZI et al. 2005, DOUCET-BEAUPRÉ et al. 2010, SOROKA 2010). One or two mismatched nucleotide pairs occur within tRNA stems in *S. woodiana* and *U. pictorum*. None of the 22 tRNA genes in *S. woodiana* shows variation, while three genes in *U. pictorum* are polymorphic: *trnA*, *trnH* and *trnL* (Table 8). Two substitutions occurred in the amino-acid acceptor stem of *trnA* and *trnL*, the observed deletion was located in DHU loop of *trnH*. The level of intra-specific variation of these genes in *U. pictorum* is up to 1%. Much greater differences between female and male tRNAs, exceeding 20%, have been observed in unionids (SOROKA 2010) and twice smaller (over 10%) in *M. galloprovincialis* (MIZI et al. 2005) and *M. trossulus* (ZBAWICKA et al. 2007). In mytilid species substitutions in tRNAs more often involve the three loops (59%) than the four stems, among which the most variable is the amino-acid acceptor stem (MIZI et al. 2005), which is also variable in *U. pictorum*. Genes tRNA are among the most conservative in representatives of Unionidae and Mytilidae (MIZI et al. 2005, ZBAWICKA et al. 2007). The small variation is probably a result of the small number of sites in these short genes, where mutations would not affect the function. Even mutations located in the loops are not regarded as completely neutral, since the loops participate in the tertiary structure of these molecules (MIZI et al. 2005).

Table 10. Kimura 2-parameter distances between unionid species for Serine *trnS1* (UGA)

	1	2	3	4	5	6	7	8	9
1. *S. woodiana*									
2. *U. pictorum*	0.000						**Overall**	**Mean**	**0.156**
3. *P. grandis*	0.017	0.017				**Excluding *I. japanensis***		**Mean**	**0.041**
4. *H. cumingii*	0.017	0.017	0.000						
5. *Q. quadrula*	0.035	0.035	0.018	0.018					
6. *C. plicata*	0.054	0.054	0.036	0.036	0.055				
7. *L. ornate*	0.054	0.054	0.036	0.036	0.055	0.036			
8. *V. ellipsiformis*	0.054	0.054	0.036	0.036	0.055	0.036	0.000		
9. *I. japanensis*	0.547	0.547	0.555	0.555	0.601	0.555	0.651	0.651	

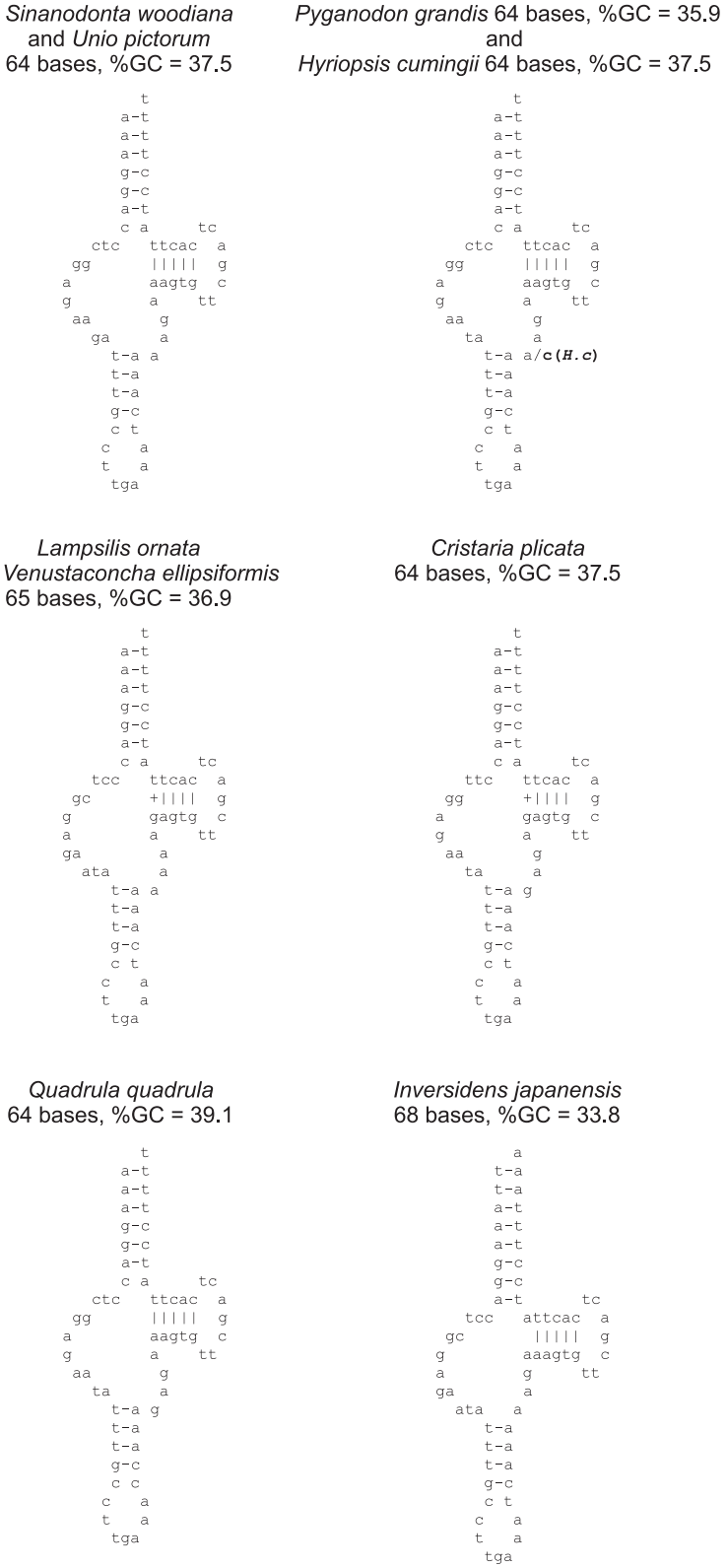

Fig. 4. Putative cloverleaf structure for the tRNA transporting serine (S1) of nine unionid species

ACKNOWLEDGEMENTS

I am grateful to Dr. ARTUR BURZYŃSKI (Polish Academy of Science, Sopot) for his help with selection of primers and gene annotation. I thank Prof. ANDRZEJ LESICKI (Adam Mickiewicz University, Poznań) for his valuable editorial comments. The studies were financed by the Polish Ministry of Science and Higher Education, grant N 303 364 33.

REFERENCES

ANDERSON S., BANKIER A. T., BARRELL B. G., DE BRUIJN M. H., COULSON A. R., DROUIN J., EPERON I. C., NIERLICH D. P., ROE B. A., SANGER F., SCHREIER P. H., SMITH A. J., STADEN R., YOUNG I. G. 1981. Sequence and organization of the human mitochondrial genome. Nature 290: 457–465.

AVISE J. C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. R. Soc. Lond. B 312: 325–342.

AVISE J. C., ARNOLD J., BALL R. M., BERMINGHAM E., LAMB T., NEIGEL J. E., REEB A. C., SAUNDERS N. C. 1987. Intra-specific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Ann. Rev. Ecol. Syst. 18: 489–522.

AVISE J. C., LANSMAN R. A., SHADE R. O. 1979. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. Genetics 92: 279–295.

BEAGLEY C. T., OKIMOTO R., WOLSTENHOLME D. R. 1998. The mitochondrial genome of the sea anemone *Metridium senile* (Cnidaria): introns, a paucity of tRNA genes, and a near-standard genetic code. Genetics 148: 1091–1108.

BLEIDORN C., PODSIADLOWSKI L., BARTOLOMAEUS T. 2006. The complete mitochondrial genome of the orbiniid polychaete *Orbinia latreillii* (Annelida, Orbiniidae) – A novel gene order for Annelida and implications for annelid phylogeny. Gene 370: 96–103.

BOGAN A. E. 2008. Global diversity of freshwater mussels (Mollusca, Bivalvia) in freshwater. Hydrobiologia 595: 139–147.

BOGAN A. E., ROE K. J. 2008. Freshwater bivalve (Unioniformes) diversity, systematics, and evolution: status and future directions. J. N. Am. Benthol. Soc. 27: 349–369.

BOORE J. L. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27: 1767–1780.

BOORE J. L., BROWN W. M. 1994a. Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics 138: 423–443.

BOORE J. L., BROWN W. M. 1994b. Mitochondrial genomes and the phylogeny of Mollusks. The Nautilus Spp. 2: 61–78.

BOORE J. L., MEDINA M., ROSENBERG L. A. 2004. Complete sequences of the highly rearranged molluscan mitochondrial genomes of the scaphopod *Graptacme eborea* and the bivalve *Mytilus edulis*. Mol. Biol. Evol. 21: 1492–1503.

BRETON S., BURGER G., STEWART D. T., BLIER P. U. 2006. Comparative analysis of gender-associated complete mitochondrial genomes in marine mussels (*Mytilus* spp.). Genetics 172: 1107–1119.

BRETON S., BEAUPRÉ H. D., STEWART D. T., PIONTKIVSKA H., KARMAKAR M., BOGAN A. E., BLIER P. U., HOEH W. R. 2009. Comparative mitochondrial genomics of freshwater mussels (Bivalvia: Unionoida) with doubly uniparental inheritance of mtDNA: gender-specific open reading frames and putative origins of replication. Genetics 183: 1575–1589.

BROUGHTON R. E., MILAM J. E., ROE B. A. 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. Genome Res. 11: 1958–1967.

BRUGLER M. R., FRANCE S. C. 2008. The mitochondrial genome of a deep-sea bamboo coral (Cnidaria, Anthozoa, Octocorallia, Isididae): genome structure and putative origins of replication are not conserved among octocorals. J. Mol. Evol. 67: 125–136.

CAO L., KENCHINGTON E., ZOUROS E., RODAKIS G. C. 2004. Evidence that the large noncoding sequence is the main control region of maternally and paternally transmitted mitochondrial genomes of the marine mussel (*Mytilus* spp.). Genetics 167: 835–850.

CAO L., ORT B. S., MIZI A., POGSON G., KENCHINGTON E., ZOUROS E., RODAKIS G. C. 2009. The control region of maternally and paternally inherited mitochondrial genomes of three species of the sea mussel genus *Mytilus*. Genetics 181: 1045–1056.

CHANG Y. S., HUANG F. L., LO T. B. 1994. The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. J. Mol. Evol. 38: 138–155.

DESJARDINS P., MORAIS R. 1990. Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J. Mol. Biol. 212: 599–634.

DOUCET-BEAUPRÉ H., BRETON S., CHAPMAN E., BLIER P., BOGAN A., STEWART D., HOEH W. 2010. Mitochondrial phylogenomics of the Bivalvia (Mollusca): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. BMC Evol. Biol. 10: 50 doi: 10.1186/1471-2148-10-50.

DREYER H., STEINER G. 2006. The complete sequences and gene organisation of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiatella arctica* -and the first record for a putative ATPase subunit 8 gene in marine bivalves. Frontiers in Zoology 3: 13.

GISSI C., IANNELLI F., PESOLE G. 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101: 301–320.

GRAF D. L., CUMMINGS K. S. 2007. Review of the systematics and global diversity of freshwater mussel species (Bivalvia: Unionoida). J. Moll. Stud. 73: 291–314.

HATZOGLOU E., RODAKIS G. C., LECANIDOU R. 1995. Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. Genetics 140: 1353–1366.

HOFFMANN R. J., BOORE J. L., BROWN W. M. 1992. A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. Genetics 131: 397–412.

INGMAN M., KAESSMANN H., PÄÄBO S., GYLLENSTEN U. 2000. Mitochondrial genome variation and the origin of modern humans. Nature 408: 708–713.

KÄLLERSJO M., VON PROSCHWITZ T., LUNDBERG S., ELDENÄS P., ERSÉUS C. 2005. Evaluation of ITS rDNA as a comple-

ment to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using Unionidae from north-western Europe. Zool. Scr. 34: 415–424.

KIMURA M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111–120.

KUHN K., STREIT B., SCHWENK K. 2006. Conservation of structural elements in the mitochondrial control region of *Daphnia*. Gene 420: 101–112.

KURABAYASHI A., UESHIMA R. 2000. Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: systematic implication of the genome organization. Mol. Biol. Evol. 17: 266–277.

LA ROCHE J., SNYDER M., COOK D. I., FULLER K., ZOUROS E. 1990. Molecular characterization of a repeat element causing large-scale size variation in the mitochondrial DNA of the sea scallop *Placopecten magellanicus*. Mol. Biol. Evol. 7: 45–64.

LIU H., MITTON J., HERRMANN J. 1996a. Genetic differentiation in and managment recommendations for the freshwater mussel, *Pyganodon grandis* (Say, 1829). Am. Malacol. Bull. 13: 117–124.

LIU H., MITTON J., WU S. 1996b. Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the fresh-water mussel, *Anodonta grandis grandis*. Evolution 50: 952–957.

MILBURY C. A., GAFFNEY P. M. 2005. Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. Mar. Biotechnol. (NY) 7: 697–712.

MIZI A., ZOUROS E., MOSCHONAS N., RODAKIS G. C. 2005. The complete maternal and paternal mitochondrial genomes of the Mediterranean mussel *Mytilus galloprovincialis*: implications for the doubly uniparental inheritance mode of mtDNA. Mol. Biol. Evol. 22: 952–967.

MORITZ C., DOWLING T., BROWN W. 1987. Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. Annu. Rev. Ecol. Systematics 18: 269–292.

MUELLER R. L., BOORE J. L. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Mol. Biol. Evol. 22: 2104–2112.

NEI M., GOJOBORI T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3: 418–426.

OKIMOTO R., MACFARLANE J. L., CLARY D. O., WOLSTENHOLME D. R. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. Genetics 130: 471–498.

PIECHOCKI A., DYDUCH-FALNIOWSKA A. 1993. Mięczaki (Mollusca). Małże (Bivalvia). Fauna Słodkowodna Polski 7A, PWN, Warszawa.

PODSIADLOWSKI L. 2006. The mitochondrial genome of the bristletail *Petrobius brevistylis* (Archaeognatha: Machilidae). Insect Mol. Biol. 15: 253–258.

PODSIADLOWSKI L., CARAPELLI A., NARDI F., DALLAI R., KOCH M., BOORE J. L., FRATI F. 2006. The mitochondrial genomes of *Campodea fragilis* and *Campodea lubbocki* (Hexapoda: Diplura): High genetic divergence in a morphologically uniform taxon. Gene 381: 49–61.

PRUITT K. D., TATUSOVA T., KLIMKE W., MAGLOTT D. R. 2009. NCBI Reference Sequences: current status, policy and new initiatives. Nucleic Acids Res. 37: D32–36.

REN J., LIU X., ZHANG G., LIU B., GUO X. 2009. "Tandem duplication-random loss" is not a real feature of oyster mitochondrial genomes. BMC Genomics 10: 84.

ROE B. A., MA D. P., WILSON R. K., WONG J. F. 1985. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J. Biol. Chem. 260: 9759–9774.

SATO M., NAGASHIMA K. 2001. Molecular characterization of a mitochondrial DNA segment from the Japanese scallop (*Patinopecten yessoensis*): demonstration of a region showing sequence polymorphism in the population. Mar. Biotechnol. (NY) 3: 370–379.

SELIGMANN H. 2008. Hybridization between mitochondrial heavy strand tDNA and expressed light strand tRNA modulates the function of heavy strand tDNA as light strand replication origin. J. Mol. Biol. 379: 188–199.

SELIGMANN H., KRISHNAN N. M., RAO B. J. 2006. Possible multiple origins of replication in primate mitochondria: Alternative role of tRNA sequences. J. Theor. Biol. 241: 321–332.

SERB J. M., LYDEARD C. 2003. Complete mtDNA sequence of the North American freshwater mussel, *Lampsilis ornata* (Unionidae): an examination of the evolution and phylogenetic utility of mitochondrial genome organization in Bivalvia (Mollusca). Mol. Biol. Evol. 20: 1854–1866.

SOROKA M. 2005. Genetic variability among freshwater mussel *Anodonta woodiana* (Lea, 1834) (Bivalvia: Unionidae) populations recently introduced in Poland. Zool. Sci. 22: 1137–1144.

SOROKA M. 2008. Doubly uniparental inheritance of mitochondrial DNA in the freshwater bivalve *Anodonta woodiana* (Bivalvia: Unionidae). Folia Biol. (Krakow) 56: 91–95.

SOROKA M. 2010. Characteristics of mitochondrial DNA of unionid bivalves (Mollusca: Bivalvia: Unionidae). I. Detection and characteristics of doubly uniparental inheritance (DUI) of unionid mitochondrial DNA. Folia Malacol. 18: 147–188.

SOROKA M., BURZYŃSKI A. 2010. Complete sequences of maternally inherited mitochondrial genomes in mussels *Unio pictorum* (Bivalvia, Unionidae). J. Appl. Genet. 51: 469–476.

TAMURA K., DUDLEY J., NEI M., KUMAR S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596–1599.

TERRETT J. A., MILES S., THOMAS R. H. 1996. Complete DNA sequence of the mitochondrial genome of *Cepaea nemoralis* (Gastropoda: Pulmonata). J. Mol. Evol. 42: 160–168.

THEOLOGIDIS I., FODELIANAKIS S., GASPAR M. B., ZOUROS E. 2008. Doubly uniparental inheritance (DUI) of mitochondrial DNA in *Donax trunculus* (Bivalvia: Donacidae) and the problem of its sporadic detection in Bivalvia. Evolution 62: 959–970.

TOMITA K., YOKOBORI S., OSHIMA T., UEDA T., WATANABE K. 2002. The cephalopod *Loligo bleekeri* mitochondrial genome: Multiplied noncoding regions and transposition of tRNA genes. J. Mol. Evol. 54: 486–500.

WENNE R. 1993. Zróżnicowanie przestrzenne i ewolucja wybranych gatunków małży morskich. Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.

WERLE E., SCHNEIDER C., RENNER M., VOLKER M., FIEHN W. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res. 22: 4354–4355.

WU X., XU X., YU Z., KONG X. 2009. Comparative mitogenomic analyses of three scallops (Bivalvia: Pectinidae) reveal high level variation of genomic organization and a diversity of transfer RNA gene sets. BMC Research Notes 2: 69.

XU K., KANNO M., YU H., LI Q., KIJIMA A. 2010. Complete mitochondrial DNA sequence and phylogenetic analysis of Zhikong scallop *Chlamys farreri* (Bivalvia: Pectinidae). Mol. Biol. Rep. Doi: 10.1007/s11033-010-9974-8.

YAMAZAKI N. R., UESHIMA J. A., TERRETT S. I., YOKOBORI M., KAIFU R., SEGAWA T., KOBAYASHI K. I., NUMACHI T., UEDA K., NISHIKAWA K., WATANABE K., THOMAS R. H. 1997. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary structures. Genetics 145: 749–758.

YU Z., WEI Z., KONG X., SHI W. 2008. Complete mitochondrial DNA sequence of oyster *Crassostrea hongkongensis* – a case of Tandem duplication-random loss for genome rearrangement in *Crassostrea*? BMC Genomics 9: 477.

ZBAWICKA M., BURZYŃSKI A., SKIBINSKI D., WENNE R. 2010. Scottish *Mytilus trossulus* mussels retain ancestral mitochondrial DNA: complete sequences of male and female mtDNA genomes. Gene 456: 45–53.

ZBAWICKA M., BURZYŃSKI A., WENNE R. 2007. Complete sequences of mitochondrial genomes from the Baltic mussel *Mytilus trossulus*. Gene 406: 191–198.

ZHENG R., LI J., NIU D. 2010. The complete DNA sequence of the mitochondrial genome of *Sinonovacula constricta* (Bivalvia: Solecurtidae). Acta Oceanol. Sin. 9: 88–92.

Received: September 21st, 2010
Accepted: November 4th, 2010