Proteomic similarity of the Littorinid snails in the evolutionary context

Arina L. Maltseva[1,*], Marina A. Varfolomeeva[1,*], Arseniy A. Lobov[1,2], Polina Tikanova[1], Marina Panova[1,3], Natalia A. Mikhailova[1,4] and Andrei I. Granovitch[1]

[1] Department of Invertebrate Zoology, St. Petersburg State University, St. Petersburg, Russia
[2] Laboratory of Regenerative Biomedicine, Institute of Cytology Russian Academy of Sciences, St. Petersburg, Russia
[3] Department of Marine Sciences, Tjärnö, University of Gothenburg, Sweden
[4] Centre of Cell Technologies, Institute of Cytology Russian Academy of Sciences, St. Petersburg, Russia
* These authors contributed equally to this work.

ABSTRACT

Background: The introduction of DNA-based molecular markers made a revolution in biological systematics. However, in cases of very recent divergence events, the neutral divergence may be too slow, and the analysis of adaptive part of the genome is more informative to reconstruct the recent evolutionary history of young species. The advantage of proteomics is its ability to reflect the biochemical machinery of life. It may help both to identify rapidly evolving genes and to interpret their functions.
Methods: Here we applied a comparative gel-based proteomic analysis to several species from the gastropod family Littorinidae. Proteomes were clustered to assess differences related to species, geographic location, sex and body part, using data on presence/absence of proteins in samples and data on protein occurrence frequency in samples of different species. Cluster support was assessed using multiscale bootstrap resampling and the stability of clustering—using cluster-wise index of cluster stability. Taxon-specific protein markers were derived using IndVal method. Proteomic trees were compared to consensus phylogenetic tree (based on neutral genetic markers) using estimates of the Robinson–Foulds distance, the Fowlkes–Mallows index and cophenetic correlation.
Results: Overall, the DNA-based phylogenetic tree and the proteomic similarity tree had consistent topologies. Further, we observed some interesting deviations of the proteomic littorinid tree from the neutral expectations. (1) There were signs of molecular parallelism in two *Littoraria* species that phylogenetically are quite distant, but live in similar habitats. (2) Proteome divergence was unexpectedly high between very closely related *Littorina fabalis* and *L. obtusata*, possibly reflecting their ecology-driven divergence. (3) Conservative house-keeping proteins were usually identified as markers for cryptic species groups ("saxatilis" and "obtusata" groups in the *Littorina* genus) and for genera (*Littoraria* and *Echinolittorina* species pairs), while metabolic enzymes and stress-related proteins (both potentially adaptively important) were often identified as markers supporting species branches. (4) In all five *Littorina* species British populations were separated from the European mainland populations, possibly reflecting their recent phylogeographic history. Altogether our study shows that proteomic data, when interpreted in the context of DNA-based phylogeny, can bring additional information on the evolutionary history of species.

INTRODUCTION

Phylogenetic analysis strives to reconstruct the evolutionary history based on variation in heritable traits. At least in some cases, the divergence is considered to be driven by adaptation (*Schluter, 2009*; *Nosil, 2012*). Similar phenotypes (at the molecular levels also) might be formed independently under similar conditions, which is classified as either parallelism or convergence, depending on the degree of genetic relatedness (*Elmer & Meyer, 2011*; *Martin & Orgogozo, 2013*). With respect to this, phenotypic traits alone obviously might be deceptive in phylogenetic studies.

The methodological breakthrough in DNA analysis introduced the molecular markers into phylogenetic studies and led to the development of several quantitative approaches to reconstruct evolutionary processes from DNA variation (*Zuckerkandl & Pauling, 1962*; *Kimura, 1968*; *Felsenstein, 1988*; *Maddison, 1997*). Throughout the "era of DNA-markers" adaptive neutrality has been the central prerequisite (*Edwards, 2009*; *Schlötterer, 2004*; *Patwardhan, Ray & Roy, 2014*). Phylogenetic distance (in terms of divergence time) ought to be measured by random neutral (not adaptive, or affected by genetic hitchhiking, or epistasis, etc.) genetic differentiation. Phylogeny was expected to reflect whole-genome, not locus-specific history since the latter might undergo very strong and fast changes (*Schlötterer, 2004*; *Grover & Sharma, 2016*; *Degnan & Rosenberg, 2009*).

Following the rapid development of sequencing techniques, the next step was taken towards phylogenomics and multi-locus phylogenies (*Edwards, 2009*; *Degnan & Rosenberg, 2009*; *Brito & Edwards, 2009*; *Lemmon & Lemmon, 2013*). The growing number of loci in analyses often provides more confident phylogenetic inferences, but also inevitably includes some loci that are under different types of selection, linked to selected loci or even originated by introgression from other species (*Edwards, 2009*; *Kuhner, Yamato & Felsenstein, 1998*; *Beerli & Felsenstein, 1999*). It is now evident that evolutionary changes affect genome unevenly: there are outlier loci, sometimes forming "islands of speciation", important for fitness and adaptation and consequently displaying outlying pattern of variation (*Luikart et al., 2003*; *Noor & Feder, 2006*; *Michel et al., 2010*; *Feder, Egan & Nosil, 2012*; *Ravinet et al., 2017*). From traditional neutral phylogenies we are moving to the whole-genome view of divergence, incorporating contributions of adaptation, introgression and hybridization.

Yet, even though the DNA-markers approach is able to answer many important evolutionary questions, it is not able to answer all of them. The idea of distinctiveness of species trees from gene trees has a long history since early 1960s (*Cavalli-Sforza, 1964*) until now (*Rosenberg, 2013*). For example, *Degnan & Rosenberg (2006)* showed that, if five or more species are regarded, there always exist branch lengths for a species tree topology, for which gene tree mismatch is more common than concordance between gene and species trees. Even though some efforts were put to minimize misleading effects (*Jewett & Rosenberg, 2012*), "the diversity recovered in our surveys of

DNA sequence evolution within and between species is ultimately an indirect and incomplete window into the history of species" (e.g., due to the different level of complexity of such entities as genes or genomes vs. organisms, population and species) (*Edwards, 2009*). Other reasoning can also be added. Firstly, genomic tools are "too far" from phenotype with its functional aspects, which calls for additional, more physiologically/ biochemically relevant approaches to fill this gap and make feasible functional interpretation of observed genetic differences (rev. e.g., in *Edwards & Batley (2004)* and *Patti, Yanes & Siuzdak (2012)*). Secondly, while phylogeny is a reconstruction of the evolutionary history from the variation in the inherited traits, it is now generally accepted that heritable information in living systems is represented not only by genetic information (*Mameli, 2004*; *Bonduriansky & Day, 2009*; *Danchin et al., 2011*). The incorporation of some non-genetic inheritance phenomena (such as niche construction, developmental bias, epigenetic regulation, etc.) into evolutionary thinking gave birth to the so called "extended synthesis" (*Pigliucci, 2009*; *Pigliucci & Müller, 2010*; *Laland et al., 2015*). Although the need for a revision of the evolutionary theory is being debated (*Laland et al., 2014*), the role of non-genetic inheritance in phenotype evolution can no longer be neglected. One way to incorporate non-genetic inheritance is, again, to include phenotypic data in evolutionary analyses. Thirdly, when analyzing events at shallow evolutionary time scale, the resolution of traditional neutral markers might be insufficient (rev. in *Edwards (2009)*). In the case of ecological speciation (i.e., adaptation-driven sympatric divergence), the neutral whole-genome divergence might be too slow to tell something about a short evolutionary history of a young species. Here again, the neutral genetic/ genomic data need to be supplemented by information on fast evolving parts of a living entity (even though the opposite examples are known as well (*Salzburger, 2018*; *Fang et al., 2018*)). There are multiple examples of using fast-evolving outlier loci to reconstruct recent ecological speciation (*Schluter, 2009*; *Luikart et al., 2003*; *Noor & Feder, 2006*; *Michel et al., 2010*; *Feder, Egan & Nosil, 2012*; *Via, 2009*). However, outliers in a broad sense are not just non-neutrally evolving loci revealed by genome-scans; they underlie proteomic, metabolomics (or morphological, ecological, behavioral, etc.) traits. Analysis of variability in the traits instead of focusing solely on underlying DNA sequences gives us a chance for causative functional explanation of the observed differences, and can bring essentially new information on the evolutionary history of a species.

"The fabric of life is protein-based" (*Karr, 2008*), and the proteome as a molecular phenotype trait is a potentially useful tool for evolutionary studies for several reasons. (1) Protein qualitative/quantitative expression patterns obviously cannot be deduced from genomic data; as the proteome mediates interaction of a genome with the environment, it brings non-redundant information about the physiology of organisms. (2) Protein expression pattern cannot be inferred from transcriptomic data either, as mismatches in quantities of proteins and their corresponding transcripts have been repeatedly detected (*Gygi et al., 1999*; *Greenbaum et al., 2003*; *Mack et al., 2006*; *Waters, Pounds & Thrall, 2006*; *Maier, Güell & Serrano, 2009*). (3) Non-genetic inheritance potentially important for speciation and maintenance of species integrity will affect the expression of proteins (*Danchin et al., 2011*). The earliest taxonomic applications of proteomics date back to the

beginning of 1980s (*Aquadro & Avise, 1981*; *Ohnishi, Kawanishi & Watanabe, 1983*) before they were replaced by DNA sequence analysis. Since then, proteomics studies have been used in a wide range of evolutionary areas, such as ecology, population biology, taxonomy, and evolutionary physiology (*Karr, 2008*; *Dworzanski & Snyder, 2005*; *López, 2005*; *López, 2007*; *Biron et al., 2006*; *Kim et al., 2008*; *Diz, Martínez-Fernández & Rolán-Alvarez, 2012*; *Baer & Millar, 2016*) but still proteomics is far behind genomics in its popularity.

We applied this rationale to an example of the gastropods family Littorinidae (periwinkles), using the proteome as a molecular phenotype trait. These marine intertidal mollusks have been model objects for various studies in evolutionary biology, ecology and adaptation (*Sokolova & Berger, 2000*; *Sokolova & Pörtner, 2001a*; *Sokolova & Pörtner, 2001b*; *Conde-Padín, Caballero & Rolán-Alvarez, 2009*; *Granovitch et al., 2009*; *Johannesson et al., 2010*; *Martínez-Fernández, De la Cadena & Rolán-Alvarez, 2010*; *Panova et al., 2010*; *Ng et al., 2011*; *Panova et al., 2011*; *Panova et al., 2014*; *Canbäck et al., 2012*; *Storey et al., 2013*; *Rolán-Alvarez, Austin & Boulding, 2015*; *García-Souto et al., 2018*; *Maltseva et al., 2016*; *Muraeva et al., 2016*; *Lobov et al., 2018*; *Lobov et al., 2015*). Littorinids provide an opportunity to analyze both closely and more distantly related species from several genera inhabiting contrasting biotopes: *Echinolittorina* and *Littoraria* in tropic and subtropic regions; *Littorina* in moderate and subarctic regions. The set of Northern Atlantics species belongs to the latter genus, forming two subgenera *Littorina* subgen. *Littorina* Férussac, 1822 and *Littorina* subgen. *Neritrema* Récluz, 1869 (*Reid, 1996*). The subgenus *Neritrema* includes two groups of closely related species—the "obtusata"-group (*L. obtusata* (Linnaeus, 1758) and *L. fabalis* (Turton, 1825)) and "saxatilis"-group (*L. saxatilis* (Olivi, 1792), *L. arcana* Hannaford Ellis, 1978 and *L. compressa* Jeffreys, 1865). The sister subgenus to *Littorina Neritrema* is *Littorina Littorina*, containing, among others, *L. littorea* (Linnaeus, 1758). The relationships among the littorinid genera, the subgenera and "obtusata" and "saxatilis" clades are well resolved based both on morphology and several neutral DNA loci (*Reid, 1996*, *1989*; *Reid, Dyal & Williams, 2012*). The hierarchy within the "saxatilis" species group still remains rather questionable, and these three species grouped in different combinations depending on the molecular approach used (*Panova et al., 2014*; *Maltseva et al., 2016*; *Reid, 1996*; *Reid, Dyal & Williams, 2012*; *Crossland et al., 1996*; *Wilding, Grahame & Mill, 2000a*, *2000b*; *Small & Gosling, 2000*). Altogether the littorinids with their well-established molecular phylogeny, some distantly related species living in similar biotopes and other closely related species inhabiting contrasting biotopes, represent a good model for comparative proteomic studies.

The present study aims to assess the utility of proteome variation to infer the evolutionary history of the species under investigation. We do not attempt to infer the true species tree by this quantitative-proteomic approach. Instead, we compare adaptively neutral (represented by neutral DNA-loci) and adaptively non-neutral (evaluated by proteomics) parts of the molecular machinery. Firstly, we compare overall patterns of proteome variation in the littorinids to a well-established phylogeny based on neutral DNA markers. Secondly, we focus on discrepancies between the DNA and proteome trees and

Table 1 Location of sampling sites, collection seasons, and sample composition.

Location	Geographic coordinates	Season	Collected species
Cancale, English Channel, France	48° 70′ N, −1° 84′ W	May, 2014	*L. saxatilis, L. arcana, L. compressa, L. obtusata, L. littorea*
Tromsø, Barents Sea, Norway	69° 43′ N, 18° 60′ E	May, 2014	*L. saxatilis, L. arcana, L. compressa, L. obtusata, L. littorea*
Chupa Bay, White Sea, Russia	66° 29′ N, 33° 68′ W	June, 2014	*L. saxatilis, L. obtusata*
Sheung Pak Nai, New Territories, Hong Kong, China	22° 27′ N, 113° 58′ E	August, 2014	*L. ardouiniana, L. melanostoma*
Oban, Scotland, UK	56° 27′ N, 5° 27′ W	April, 2015	*L. saxatilis, L. arcana, L. compressa, L. obtusata, L. faballis*
Eilat, Red Sea, Israel	29° 30′ N, 34° 54′ E	August, 2015	*E. millegrana, E. marisrubri*

attempt to interpret them in the light of the existing knowledge on evolution and ecology of the species and on the function of the involved proteins.

MATERIALS AND METHODS

Ethics statement

Periwinkles were collected from wild populations (Table 1). Periwinkles are not endangered or protected species in the study regions, so no special permission for their collection was required. Administrations of the Swire Institute of Marine Science, Hong Kong University (Hong Kong, China), the Arctic University of Norway University of Tromsø (Norway), the Educational and research station "Belomorskaia" of St. Petersburg State University (Russia), the Scotland Association of Marine Science (Oban, Scotland, UK) and Interuniversity Institute for Marine Sciences (Eilat, Israel) were informed about the snails sampling. Collection and animal numbers were approved by these authorities. No special approval could be obtained for sample collection at the site near Cancale (France) because it is not an area of any national park or a private territory, so there was no appropriate authority to apply for permission.

Sample collection and preparation

Adult individuals of littorinid snails *Littorina littorea* (Linnaeus, 1758), *Littorina fabalis* (Turton, 1825), *Littorina obtusata* (Linnaeus, 1758), *Littorina arcana* Hannaford Ellis, 1978, *Littorina compressa* Jeffreys, 1865, *Littorina saxatilis* (Olivi, 1792), *Littoraria ardouiniana* (Heude, 1885), *Littoraria* "melanostoma E Asia" (is unnamed, see (*Reid, Dyal & Williams, 2010*)), *Echinolittorina millegrana* (Philippi, 1848) and *Echinolittorina marisrubri* Reid, 2007, were collected from wild populations in six geographic locations (Table 1).

After collection animals were transported to the laboratory. The *Littorina* snails from the Northern Atlantic were kept in aerated moist containers at 8 °C and rinsed with salt water once a day. Snails were acclimated to these standard conditions for no less than 2 days and no longer than 1 week before dissection. The *Echinolittorina* and *Littoraria* snails from the Red and East-China Sea coasts, respectively, were kept in dry containers at room temperature and rinsed with salt water once a day.

For sample preparation snails were dissected under MBS-10 binocular microscope at 79-849 magnifications to identify species, sex and possible trematode or other heavy parasitic infection and to separate body parts (head, foot and penis). The mollusks identification was conducted according to the shell form, sculpture and pigmentation, as well as male and female reproductive system anatomy (*Reid, 1986*, *1996*, *2007*; *Granovitch et al., 2008*). The correctness of the tropical species identification was confirmed by Dr. David G. Reid personally. The identification of the species within the *Littorina* "saxatilis"- and "obtusata"-groups (the most challenging part) was performed as described previously (*Reid, 1996*; *Granovitch et al., 2008*). More details can be found in the File S1. Only the reliably identified mature individuals with well-developed reproductive system and free of trematode or other obvious infection were used in further analysis.

Tissues of up to 20 animals (collected from different parts of intertidal area within the same location) were pooled separately for different body parts (foot, head, penis) and sexes; and homogenized in lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 25 mm Tris, pH 8.2) using Mixer Mill MM 400 (Retsch). Particles were sedimented by centrifugation 12,000g, 15 min, 4 °C and supernatants were frozen at −80 °C until use.

Proteomic analysis strategy

The "biological noise of the system" due to individual variations and microniche adaptation was leveled by pooling, the short acclimation procedure and including biological replicates into the analysis (rev. in *Karp et al. (2005)*). Samples of somatic tissues of males and females were used as biological replicates, as a previous study showed no significant sex effect at the qualitative level (*Maltseva et al., 2016*). Head and foot tissues used for analysis in this study are less sensitive to transient environmental impacts than for example, gill as was demonstrated earlier in a study on salinity stress (*Muraeva et al., 2016*). Technical replicates (at least two for every sample) were averaged to reduce "technical noise" (*Karp et al., 2005*). Altogether this strategy ensures the analysis of the most stable and reproducible proteome, reducing accidental impacts.

In general, the qualitative strategy of a proteomic study aims for the characterization of the set of proteins present in a sample. This was successfully applied to characterize proteomes of different compartments (organelles, tissues, organs) and certain classes of proteins etc., being an integrative part of structural biology. The quantitative approach estimating proteins abundances is a part of functional biology, and is more informative for studies of physiological changes, adaptive mechanisms, etc. (rev. in *Messana et al. (2013)*). We chose the qualitative approach to compare basic proteomes between species since it is less amenable to transient physiological and environmental impacts. Our approach principally resembles the blind whole-genome assaying methods like AFLP, RFLP, RAPD, etc. in the sense of that the identity of analyzed proteomic markers (protein spots) is unknown. The main difference is taking into account the "working part of the living molecular machinery", that is, the proteome, while abovementioned methods engage DNA-regions, both coding and noncoding. We separated proteins in two-dimensional gel electrophoresis, and ultimately, estimated the number of common and unique (new or altered due to amino acid substitutions, modifications, splicing patterns, etc.) protein

signals in single species or species groups in comparison (more details on our approach and rationale of proteomic vs. genetic data comparison can be found in the File S1).

We followed a gel-based methodology because it has, among others, an advantage of whole-protein analysis and consequently saves combinations of different post-translational modifications as individual signals (*Arentz et al., 2015*). We used differential 2D-electrophoresis (DIGE (*Ünlü, Morgan & Minden, 1997*)) because it allows the simultaneous analysis of up to three samples within the same gel (allowing the reliable detection of even minor structural changes) and provides high sensitivity of protein detection (down to 150–500 pg of a single protein) (*Lilley & Friedman, 2004*). Qualitative gel analysis and spot matching were carried out using PDQuest Advanced 8.0.1 software (BioRad, Hercules, CA, USA) by making whole experiment master gel, allowing for the warping of individual gel scans (*Berth et al., 2007*). Spots were considered reliably detected if they were detected in at least two technical replicates of the same sample, or in one technical replicate in at least two different samples. We have taken into account potential post-translational modifications and/or splice variants by identifying independent signals, because every particular modification/splice form possesses its own features and functions, and appearance of a new form/modification represents an important evolutionary event, comparable to amino acid(s) substitution or even the emergence of a new protein. In total, 796 spots were used for inter-species comparison of proteomes. Protein spots differing between species and available for excision after Coomassie gel staining were subjected to tandem mass spectrometry MS/MS-analysis for identification following standard "bottom up" protocol. Basically, in this method proteins are digested into peptides prior to MS/MS-analysis, molecular masses of resulting peptides and their fragments (after fragmentation in the collision cell of mass-spectrometer during MS/MS-analysis) are determined and compared with MS-spectra predicted from a database by software (see File S1 for more details). The success or failure in protein identification did not influence the results of inter-specific comparisons, and aided only in their interpretation (more details on experimental procedures can be found in File S1).

Statistical analysis

All analyses were run using R (*R Core Team, 2015*).

Clustering

Proteomic data were analyzed at two levels. First, data on presence/absence of proteins in samples were used to assess proteomic differences related to geographic location, sex and body part. Second, "consensus" species proteomes were constructed to assess the overall similarity of species proteomes. A protein present in at least one sample of a given species was considered as present in that species. These two datasets were analyzed in the same way: Jaccard dissimilarity index was computed, dendrograms were derived using neighbor joining (NJ) and unweighted pair group method with arithmetic mean (UPGMA) and plotted with dendextend package (*Galili, 2015*). Suitability of NJ vs. UPGMA clustering methods was checked using plots of pairwise distances on a tree vs. original pairwise distances. The both methods yielded comparable results for sample

clustering (File S2); NJ performed slightly better for clustering of consensus species proteomes (File S2).

Stability of clustering

Branch support was assessed by approximately unbiased (AU) *p*-values using multiscale bootstrap resampling (*Shimodaira, 2004*) with 10,000 iterations in the pvclust package (*Suzuki & Shimodaira, 2006*). This number of iterations ensured accurate estimation of AU *p*-values (their standard errors were less than 0.01). Branch support for midpoint-rooted NJ-based trees was assessed using bootstrap with 1,000 iterations (*Efron, Halloran & Holmes, 1996*) in the ape package (*Paradis & Schliep, 2018*). Stability of UPGMA-based sample clustering was additionally assessed using nonparametric bootstrap with 10,000 iterations in the fpc package (*Hennig, 2007*, *2008*, *2015*). After the original tree was split into a specified number of clusters, an index of cluster stability was computed cluster-wise as a mean Jaccard similarity (J) of the actual cluster to the most similar cluster in a bootstrapped data. $J > 0.75$ corresponds to stable and successfully recovered clusters, while $J < 0.5$ marks "dissolved" clusters with uncertain grouping of objects. In addition, the bootstrap samples allowed to estimate the probability of recovering a particular cluster when the underlying data change.

Cluster markers

The qualitative differences of the periwinkle proteomes were assessed using indicator species analysis (IndVal) with the indicspecies package (*De Cáceres & Legendre, 2009*). The IndVal method measures the strength of association of specific features with particular groups of samples. It was originally developed in community ecology and has been successfully used for detection of molecular markers (*Feng, Bootsma & McLellan, 2018*). We have run several analyses to detect taxon-specific protein markers using the data on presence/absence of proteins in the samples (for each of the 10 species individually, *Littoraria* (*L. ardouiniana* + *L. melanostoma*), *Echinolittorina* (*E. millegrana* + *E. marisrubri*), the pair *L. arcana* + *L. saxatilis*, the "saxatilis"-species group (*L. arcana* + *L. saxatilis* + *L. compressa*) of the *Littorina* (subgen. *Neritrema*), the "obtusata"-species group (*L. fabalis* + *L. obtusata*) of the *Littorina* (subgen. *Neritrema*), the subgenus *Neritrema* (genus *Littorina*), and genus *Littorina* (subgenera *Neritrema* + *Littorina*)). Group-equalized version of IndVal was used to avoid biases due to unbalanced group sizes (*De Cáceres & Legendre, 2009*). The *p*-values of the IndVal statistic were corrected for false discovery rate in multiple tests (FDR) using Benjamini–Hochberg correction (*Benjamini & Hochberg, 1995*). The proteins with a significant indicator value (FDR-adjusted $p \leq 0.05$) were used to compute specificity (the probability that the sample belongs to the target category given the particular protein has been found), and sensitivity (the probability of finding the protein given the sample belongs to the target category) of cluster markers.

Phylogenetic analysis

We used sequence data from the previous extensive study on littorinid phylogeny (*Reid, Dyal & Williams, 2012*) to produce a tree only for those species that are included in the present study. Fragments of 28S rRNA, 12S rRNA and cytochrome oxidase subunit I

(COI) were obtained from NCBI; information about sequences used is available in File S1. Sequences were concatenated with different genes allowed to evolve independently. Phylogenetic analysis was performed using Bayesian inference, MrBayes 3.2.6 (*Ronquist & Huelsenbeck, 2003*) with the same parameters as in the original study (*Reid, Dyal & Williams, 2012*): fragments of the three genes were used as a concatenated sequence, where individual genes were unlinked to evolve independently; nucleotide substitution model was GTR + G + I; analysis was performed as two independent runs, five chains in each (four heated and one cold; the first 25% samples from the cold chain were discarded) for 25,000,000 generations with a sample frequency of 1,000, print frequency of 1,000 and diagnostics calculated every 1,000 generations. The convergence between two runs was tested by comparison of statistical parameters in the Tracer Software (http://tree.bio.ed.ac.uk/software/tracer/).

Comparison of phylogenetic and proteomic trees

Prior to comparison, phylogenetic and NJ-based proteomic trees were made ultrametric using non-negative least squares with the phangorn package (*Schliep, 2011*). The Robinson–Foulds distance (both raw and normalized) between unrooted trees was computed to assess the total number of partitions which are implied by one of the trees, but not the other (*Robinson & Foulds, 1981*). Compositional similarity of proteomic and phylogenetic trees was evaluated using the Fowlkes–Mallows index (*Fowlkes & Mallows, 1983*). Correspondence between inter-species distances on the trees was measured using cophenetic correlation (*Sokal & Rohlf, 1962*). In addition, cophenetic distances on the proteomic and phylogenetic trees were expressed as fractions of maximal tree distance to allow direct comparisons (the matrices of transformed distances are presented in Fig. 1).

RESULTS

General trees' topologies

We applied two algorithms for proteomes clustering (NJ and UPGMA) which both gave similar results (Figs. 1 and 2; File S2). We deduced the overall similarity of species proteomes from their clustering based on presence/absence of peptides in consensus proteomes (Fig. 1A; File S2). These results were compared with a phylogenetic tree obtained by Bayesian inference based on three molecular markers used earlier by *Reid, Dyal & Williams (2012)* (Fig. 1B; File S2). The general trees' topologies were similar: the first split was between genera—*Littoraria*, *Echinolittorina* and *Littorina* (however, the genera were split in different order at the proteomic and genetic-based trees); then within *Littorina* subgenera (*L.* (*Littorina*) *littorea* and *L. Neritrema* spp. separated), and finally the "saxatilis"-group split from the "obtusata"-group. This was confirmed by the low Robinson–Foulds distance (raw RF = 2, normalized RF = 0.43), high cophenetic correlation coefficient (0.80), and by the distribution of the Fowlkes–Mallows index, which indicated significant compositional similarity for all numbers of clusters, except nine (Fig. 1C). The discrepancy between the trees with nine clusters came from the "saxatilis"—group species cluster, where three species grouped differently (Fig. 1): *L. compressa* was a sister-group to the pair *L. saxatilis*/*L. arcana* in the proteomic tree and *L. saxatilis*
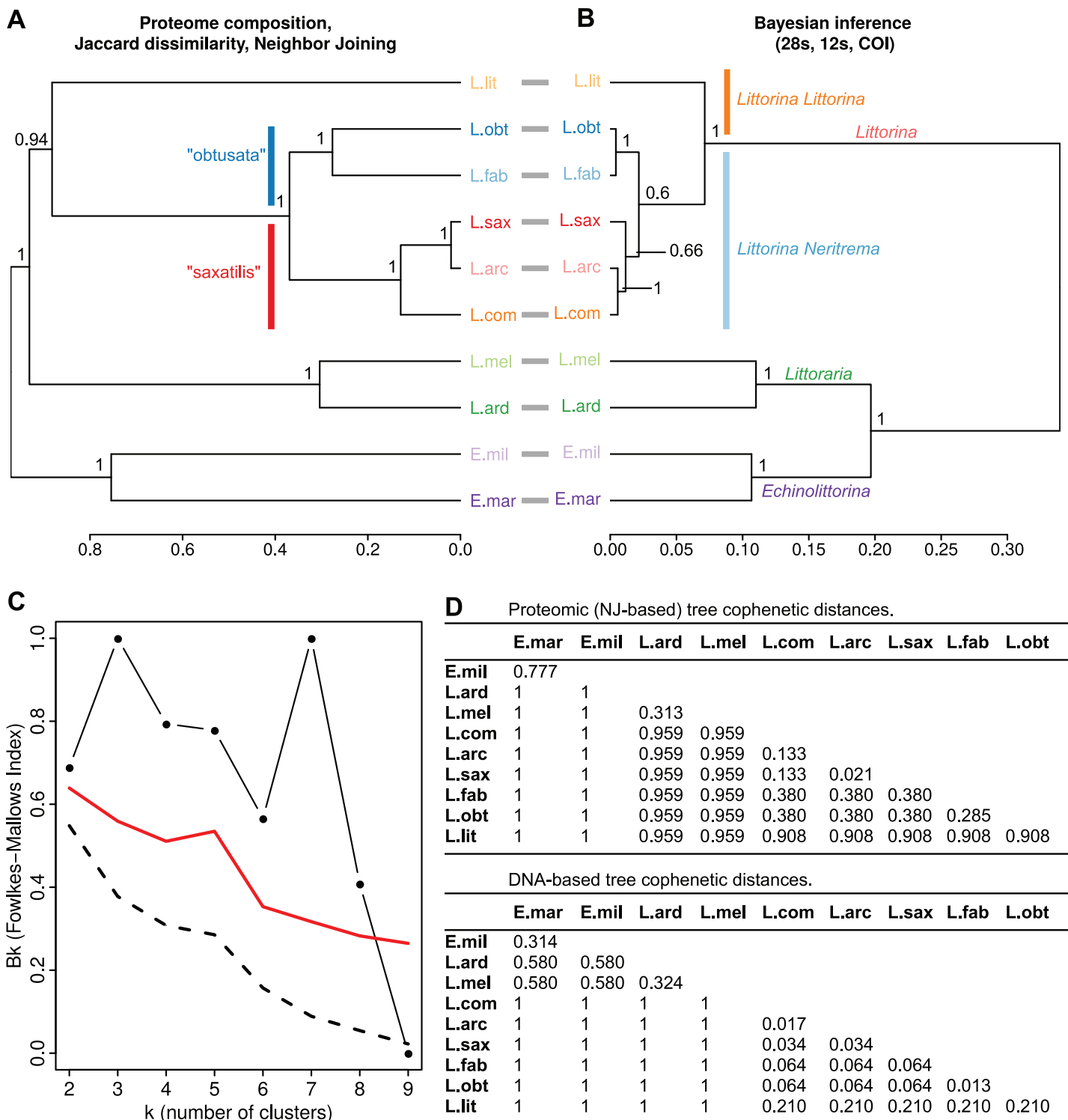
Figure 1 Interspecies relations within the family *Littorinidae*. (A) Dendrogram of consensus species proteomes obtained via neighbor joining based on Jaccard dissimilarities of protein occurrence frequency in samples of different species. The bootstrap support values are shown. (B) The molecular phylogeny tree obtained via Bayesian inference using concatenated partial gene sequences from 28S rRNA, 12S rRNA and cytochrome oxidase C subunit I (COI). Support values are posterior probabilities. Prior to comparison, the both trees (A) and (B) were made ultrametric using non-negative least squares. Robinson–Folds distance between unrooted trees was RF = 2 (normalized RF = 0.143). The cophenetic correlation

was a sister taxon to the pair *L. compressa/L. arcana* in the genetic one. Noteworthy, clustering in the proteomic tree was supported by high bootstrap values, unlike low support (0.66) of the "saxatilis"—cluster in the DNA-tree.

Interspecies cophenetic distances

On the whole, proteomic distances in general exceeded genetic distances (the full matrices of distances, both genetic and proteomic, are available in Fig. 1D). The most robust differences were as follows. Relationships in the "saxatilis" group species cluster: *L. arcana* and *L. compressa* were considered phylogenetically close, but rather distant at the proteomic level (the proteomic and genetic distances between *L. arcana* and *L. compressa* were 0.133 and 0.017, respectively); on the contrary, *L. arcana* was proteomically very similar to *L. saxatilis* in contrast to the genetic distance (the proteomic and genetic distances between *L. arcana* and *L. saxatilis* were 0.021 and 0.034, respectively); the differences were even more profound if UPGMA clusterization algorithm was applied (see File S2). Distances in the "obtusata" group species cluster: *L. fabalis* and *L. obtusata* were identified as phylogenetically very close but strongly diverged at the proteomic level (the proteomic and genetic distances between *L. fabalis* and *L. obtusata* were 0.285 and 0.013, respectively). Their proteomic cophenetic distance is comparable with one between genetically distant *Littoraria ardouiniana* and *L. melanostoma* (Fig. 1A, *Littoraria* species formed common cluster just one step earlier than *L. obtusata/L. fabalis*, Table 2; Table S2). The *Littoraria* genus case. Similarly, proteomic distances were larger than genetic ones for two *Echinolittorina* species and two *Littorina* subgenera (*Littorina* and *Neritrema*). At the same time, two *Littoraria* species showed an unusual outcome: being phylogenetically quite distant, they appeared a bit closer to each other at the proteomic tree (the proteomic and genetic distances between *L. ardouiniana* and *L. melanostoma* were 0.313 and 0.324, respectively), which might be interpreted as a sign of physiological and molecular parallelism.

Effect of geography, sex and body part

To assess the effect of geography, sex and body part on the similarity of proteome composition, we calculated the extended proteomic tree of samples (Fig. 2; File S2). The general topology of this tree was similar to the consensus one. Almost all species formed clear species-clusters, which supports the potential taxonomic use of proteomics. Nevertheless, there was an exception: the pair *L. saxatilis/L. arcana* formed a mixed cluster, where the factor "body part" (and sometimes "geographic location") affected proteome
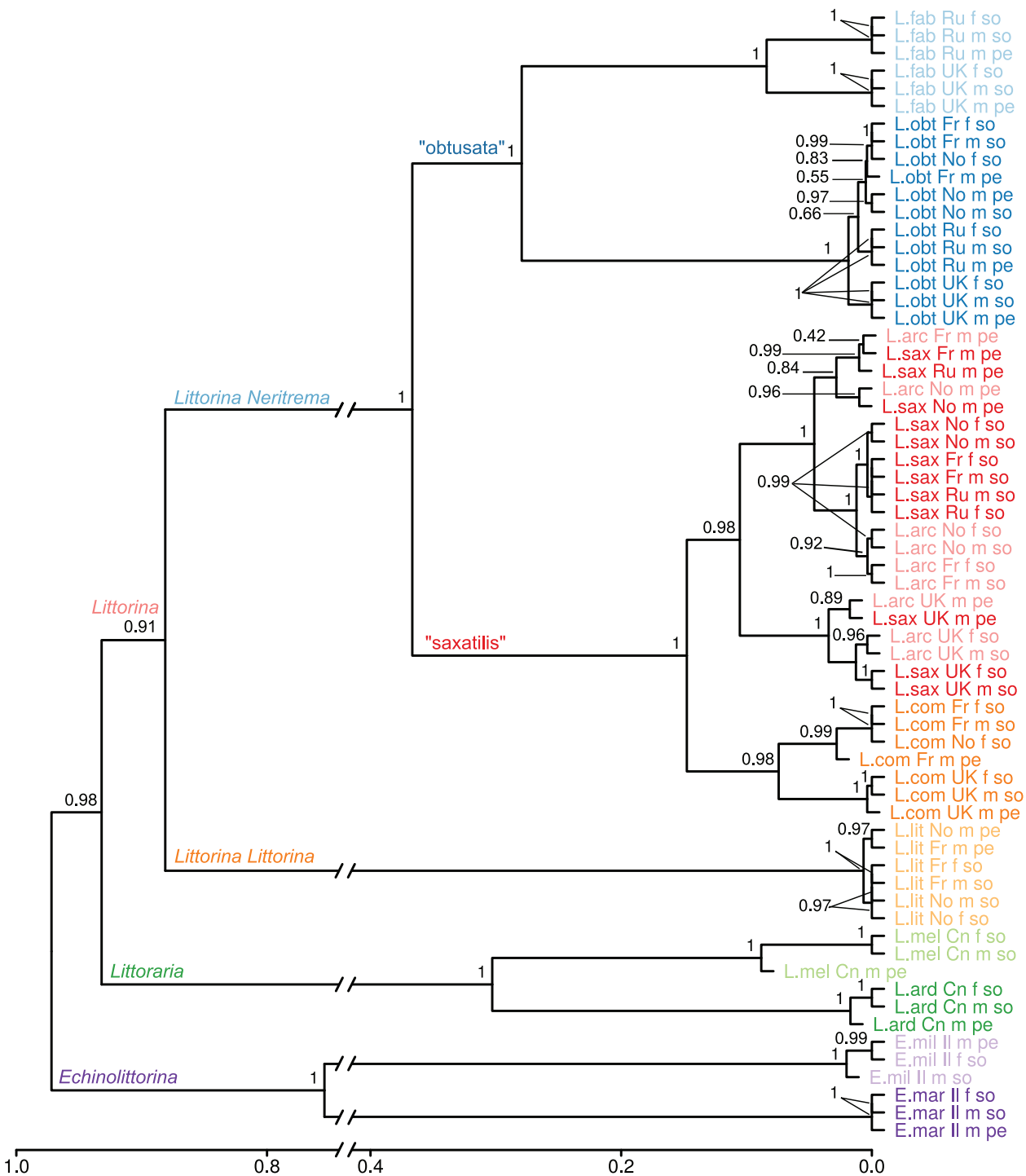
Figure 2 **Dendrogram of proteome UPGMA clustering from samples of the 10 Littorinidae species.** Clustering was produced using unweighted pair group method with arithmetic mean (UPGMA) algorithm based on Jaccard dissimilarity coefficients for the data on presence/absence of proteins in the samples. Sample labels indicate species (L.arc: *Littorina (Neritrema) arcana*; L.comp: *Littorina (Neritrema) compressa*; L.sax: *Littorina (Neritrema) saxatilis*; L.obt: *Littorina (Neritrema) obtusata*; L.fab: *Littorina (Neritrema) fabalis*; L.lit: *Littorina (Littorina) littorea*; L.ard: *Littoraria ardouiniana*; L.mel: *L. melanostoma*; E.mar: *Echinolittorina marisrubri*; E.mil: *E. millegrana*), location (Ru: White Sea, Russia; Fr: English Channel, France; UK: Atlantic coast, Scotland; No: Barents Sea, Norway; Cn: East-China Sea, Hong Kong; Il: Israel), sex (f: female; m: male) and body part (so: foot + head parts; pe: penis). The approximately unbiased bootstrap support values are shown. (neighbor joining-based clustering is presented in File S2).
Full-size 🖼 DOI: 10.7717/peerj.8546/fig-2

Table 2 Composition, protein markers and supporting values of sample clusters. Selected clusters obtained from analysis of Jaccard dissimilarities for the full set of samples (the full list is presented in Table S1). The first column indicates the composition of a cluster; the second is the number of protein markers, identifying the cluster with maximal specificity and sensitivity; the third is the threshold number of clusters when the cluster of this particular composition appears (from a minimal number of clusters to maximal); the fourth is the recovery probability of the particular composition cluster when it appears for the first time (e.g., the "saxatilis"-group cluster appeared at first during partition into 6 clusters; the recovery probability after this partition was 1; while this cluster persisted through 2 more partitions, its probability dropped to 0.99 and 0.90, respectively).

Composition of clusters (CC)	Number of markers (NC)	NC for the CC first appearance	Recovery probability
Littoraria ardouiniana	12	8	0.98
Littoraria melanostoma	24	8	0.98
Echinolittorina marisrubri	50	5	0.95
Echinolittorina millegrana	58	5	0.94
Littorina Littorina littorea	100	4	1
Littorina Neritrema fabalis	28	7	0.92
Littorina Neritrema obtusata	28	7	0.96
Littorina Neritrema arcana	2	–	–
Littorina Neritrema compressa	11	9	0.99
Littoraria (L. ardouiniana /+ L. melanostoma)	74	7	0.82
Echinolittorina (E.millegrana + E.marisrubri)	28	4	1
Littorina arcana + Littorina saxatilis	15	9	0.89
"saxatilis"-group (L. arcana + L. saxatilis + L. compressa)	33	6	1
"obtusata" group (L. fabalis + L. obtusata)	12	6	0.89
Littorina Neritrema	153	2	1
Littorina (Neritrema + Littorina)	17	–	–

stronger than "species". This cannot be due to male species-misidentification, as somatic samples of those males grouped together with unmistakably identified females of these two species. The first within-cluster division splits the UK samples of both species from the mainland ones, suggesting that the geographic variation surpassed the interspecific variability, and a parallel pattern is observed in the proteomes of these two species. Remarkably, distinct UK-clades were also present within the species-clusters of *L. obtusata*, *L. fabalis* and *L. compressa*.

Cluster stability of samples and cluster markers

To further assess the proteome species-specificity we analyzed the stability of sample clustering (*Hennig, 2007*). Splitting the bootstrapped tree of samples into 10 clusters (the number of species analyzed) resulted mainly in same-species clusters with high recovery probability (Table 2; Table S2), with the exception of the *L. arcana*/*L. saxatilis* pair which formed two multi-species clusters, corresponding to European mainland and UK locations (both with maximal recover probability). All species branches were supported by the presence of highly reliable (i.e., maximum specificity and sensitivity) protein markers, except for *L. saxatilis* (Table 2), which appeared poorly distinguishable from *L. arcana* at the proteomic level. For *L. arcana* there were only two specific markers. Moreover, these two markers represent two forms of the same enzyme: the peptidyl-prolyl cis-trans isomerase, which additionally reflects the physiological similarity of these two species.

Markers analysis

We succeeded in the MS/MS-identification of several species (cluster) markers (of 796 protein spots analyzed 128 were successfully annotated, Tables S1 and S3). Expectedly, the most efficient identification was based on a *L. saxatilis* database (LSD) search (*Canbäck et al., 2012*) and gave positive results mostly in the case of *L. saxatilis* and *L. arcana* (~50% efficacy) in comparison to other species (~4–20% efficacy, Table S3). A variety of house-keeping proteins were usually revealed as markers of cryptic groups ("saxatilis" and "obtusata") and genera (*Littoraria* and *Echinolittorina* species pairs), such as protein disulfide isomerase, calreticulin, tropomyosin, troponin T, radular globin, etc. Metabolic enzymes and stress related proteins (both potentially adaptively important), like arginine kinase, fructose-bisphosphate aldolase, malate dehydrogenase, peptidyl-prolyl cis-trans isomerase (PPIase), glutathione-S-transferase, etc., were often identified as species-branch supporting markers (the latter was also the case for *L. saxatilis/L. arcana* pair vs. *L. compressa*). The unusual pattern was found for the *Littoraria* species pair, where species with quite long independent evolutionary history matched by core metabolic enzymes (unlike, e.g., "saxatilis" group), but differed by globins (which were common within either the "saxatilis" or "obtusata" group and within the *Echinolittorina* species pair).

DISCUSSION

In our analysis, the topology of the tree based on the Jaccard dissimilarities of consensus proteomes (Fig. 1) proved to largely agree with the classic morphological (*Reid, 1996*) and the consensus molecular phylogenies (*Reid, Dyal & Williams, 2012*). Moreover, in the tree inferred for location, sex and body part specific samples the strongest differences of proteomes were observed between species: most samples clustered by species (Fig. 2, except for *L. arcana* and *L. saxatilis*, discussed below). This supports the taxonomic utility of proteomic studies, which also was demonstrated in other organisms ranging from bacteria and protists to metazoans (*Dworzanski & Snyder, 2005*; *López, 2005*; *Kim et al., 2008*).

There are however some disagreements in branch length and topology of specific clades between the DNA and proteome trees. Being a kind of molecular phenotype, the proteome embodies both slowly evolving neutral and fast evolving adaptive evolutionary changes. In contrast, classical phylogenetic analyses target neutral DNA loci. Thus, the comparison of these two types of analyses (proteomic vs. genetic molecular markers) can potentially bring some new information about species evolutionary history, as discussed below.

Similar proteomes in deeply diverged species: a possible case of the whole-proteome convergence/parallelism

In general, the proteomic distances in the trees exceeded the genetic ones. This was rather expectable based on the inclusion of fast evolving proteins; for example, arginine kinases (present as several forms in periwinkles proteomes (*Maltseva et al., 2016*)) were among annotated proteins, and these enzymes are well known for their non-conservative nature (*Uda et al., 2006*). However, there was one remarkable exception from the "large proteomic

distances"-rule. Two deeply diverged *Littoraria* species appeared in the proteomic tree closer to each other than in the genetic one. This result may be interpreted as a molecular parallelism. Although there are many known examples of molecular convergence/ parallelism (reviewed in *Storz (2016)*), usually this phenomenon is analyzed at the level of individual proteins of particular families, not the whole proteome. There are strong theoretical arguments for molecular homoplasy resulting from neutral and not adaptation-driven phenomena (*Zou & Zhang, 2015*), for example, a set of structural constraints, limiting protein changes to certain trajectories (reviewed in *Storz (2016)*; *Starr & Thornton (2016)*). Interestingly, in the present study another tropical species pair *E. marisrubri/E. millegrana* (their estimated divergence-time is slightly shorter than that of the *Littoraria* species, ~35 Ma and ~48 Ma, respectively) (*Reid, Dyal & Williams, 2012*) showed no sign of proteomic parallelism, though both species pairs live under extreme conditions of thermal stress, hypoxia and desiccation. Nevertheless, adaptation to a rather special biotope in mangroves, sympatrically inhabited by *L. ardouiniana/ L. melanostoma*, could be another possible explanation, as both species are ecologically very similar in terms of host-tree species distribution, diet and activity (*Lee, Williams & Hyde, 2001*; *Lee & Williams, 2002a*). Altogether, the example of the two *Littoraria* species shows that similar proteomes can emerge from divergent genomes.

Interestingly, globin proteins deviate from the observed pattern of whole-proteomic similarity, showing differences between the two *Littoraria* species. This may appear quite puzzling; for comparison, no globin-related difference was observed between the strongly diverged *Echinolittorina* species. A possible explanation for the globin divergence in *Littoraria* species may be differences in their vertical distribution and expected different aerial-aquatic respiration VO_2-ratio (*McMahon, 1988*; *Lee & Williams, 2002b*). Similarly, the "obtusata"- and "saxatilis"-groups differ in their preferences to littoral zone level (*Granovitch et al., 2008*, *2004*). Accordingly, they demonstrate different VO_2-ratio (*McMahon, 1988*) and different globins, though these species are much more phylogenetically close compared to the *Echinolittorina* species pair.

Distant proteomes in phylogenetically close species with different habitat preferences

The species pair of the "obtusata" group—*Littorina obtusata* and *L. fabalis* shows a pattern that is opposite to the one described above for *Littoraria*. In the *Littorina* "obtusata" group we observed strongly diverged proteomes, apparently originating from very akin genomes in morphologically similar species. These two species are phylogenetically very close (closer than any other in the Littorinidae family; estimated divergence time is less 1 Ma) (*Reid, Dyal & Williams, 2012*) and were shown to share some mitochondrial haplotypes (*Kemppainen et al., 2009*) (and there are signs of limited gene flow between these two species in Portugal (*Costa et al., in press*)). Here, the comparative proteomic analysis revealed robust differences between *L. obtusata* and *L. fabalis*. The distance between them highly exceeds those within "saxatilis"-group. The discrepancy between genetic and proteomic similarities can be an indirect evidence of ecological speciation. *L. obtusata* and *L. fabalis* demonstrate different preferences in their microhabitat

distribution: *L. fabalis* occupies lower intertidal and subtidal levels, while *L. obtusata* prefers the middle littoral zone; at those levels, they are associated with different fucoids (*Reid, 1996*; *Granovitch et al., 2008*, *2004*). One can hypothesize that partitioning of ecological niches drove divergence at the physiological level, finally reflected in very different proteome patterns in these two species. The observed striking proteomic differences were achieved by these two species in relatively short evolutionary time, which implies a high evolvability of the system. However, the exact nature of proteomic differences is unknown (whether it is related to genes duplication, amino acid substitution, gene expression pattern shift, alternative splicing or some posttranslational modification), and at least part of the differences can be due to phenotypic plasticity. Thus, these species present an interesting system to further investigate mechanisms of rapid proteome divergence.

Distant proteomes in phylogenetically close species with similar habitat preferences

The third example, from the *Littorina* "saxatilis" group, suggests that different proteomes can be generated by similar genomes even if corresponding species populate very similar biotope. The three species in this group (*L. arcana*, *L. compressa* and *L. saxatilis*) have a high degree of whole genome similarity, as was demonstrated by comparative genomic hybridization (*Panova et al., 2014*). Earlier molecular phylogenies were inconclusive regarding these species relationships (*Canbäck et al., 2012*; *Reid, 1996*; *Reid, Dyal & Williams, 2012*) and different molecular markers clustered the three species of the "saxatilis" group in all possible combinations (*Panova et al., 2014*; *Reid, Dyal & Williams, 2012*; *Wilding, Grahame & Mill, 2000a*, *2000b*; *Small & Gosling, 2000*; *Knight & Ward, 1991*), which is one line of evidence for their close phylogenetic proximity and a recent evolutionary divergence (divergence time is estimated as ~1–1.5 Ma) (*Reid, Dyal & Williams, 2012*).

In our proteomic tree, the first divergence event splits *L. compressa* from a tight alliance of *L. saxatilis*/*L. arcana* (Fig. 1A). We found 11 species-specific markers for *L. compressa* and a similar number of markers specific for the pair *L. saxatilis*/*L. arcana*, but only two markers specific for *L. arcana* and none for *L. saxatilis* (Table 2). These results are congruent with the fact that both *L. saxatilis* and *L. arcana* demonstrate parallel proteomic shifts along the vertical shore gradient while *L. compressa* does not (*Maltseva et al., 2016*). The separation of *L. compressa* at the proteomic level partially resembles the foregoing case of *L. fabalis* and *L. obtusata* in a sense that it occurs between very closely related species. However, unlike the latter pair, the proteome divergence is not easily explained by ecology in this case. Both *L. arcana* and *L. compressa* are very similar to *L. saxatilis*: their ranges and habitat distribution overlap (*Reid, 1996*; *Granovitch et al., 2013*). Although *L. compressa* was not recorded from fucoid-free substrates, unlike two other species, it invariably lives in sympatry with *L. saxatilis*. All three species are often found together on the same fucoids and rocks within an intertidal area, and there are so far no data suggesting stronger differences in ecological preferences between *L. compressa* and *L. saxatilis* than between *L. saxatilis* and *L. arcana*. It is still possible that *L. compressa*

occupies very specific microhabitats to be elucidated in future studies. Alternatively, some hypothetical switch could have occurred in the molecular machinery functioning of *L. compressa*. The last explanation suggests that speciation events could be related not only to environmental shifts or niche expansion but also to changes in an organism's patterns of interaction with the same (or very similar) environment. Among protein markers specific to *L. compressa* we identified arginine kinase, fructose-bisphosphate aldolase and glutathione-s-transferase. These enzymes are potentially related to stress adaptation (e.g., isolation-related hypoxia during low tide; in details discussed in *Maltseva et al. (2016)*); and such changes are compatible with any of the two proposed explanations.

Strinkingly similar proteomes in phylogenetically close species

In contrast to the *L. compressa* case discussed above, the pair *L. saxatilis*/*L. arcana* showed high proteome similarity. In general, this similarity is expected pattern given the short divergence time and overlapping habitats. What is surprising is that both "geographic location" and "body part" showed stronger impacts on the proteome than "species"; that is, the first split occurs between UK and other regions (further discussed below), then between "foot + head" and "penis" samples, and then between the species (Fig. 2). The misidentification of morphologically similar males of *L. saxatilis* and *L. arcana* could not explain this pattern, because female "foot + head" samples cluster together with male "foot + head" samples; and females of the two species are unambiguously distinguishable.

L. saxatilis and *L. arcana* have strikingly similar proteomes (at least in adults and the analyzed body parts), even though they have different reproductive strategies (*L. arcana* is oviparous while *L. saxatilis* is ovoviviparous) and demonstrate some differences in distribution, morphology and genetics (*Panova et al., 2014*; *Reid, Dyal & Williams, 2012*; *Wilding, Grahame & Mill, 2000a*; *Granovitch et al., 2008*; *Knight & Ward, 1991*; *Mikhailova et al., 2009*). Given their genomic, morphological, ecological and now even proteomic similarity, this pair of species raises intriguing evolutionary questions. (1) What mechanisms maintain this similarity across their range? (2) What kind of barrier separates them from each other given indirect evidence of gene flow between them (*Granovitch et al., 2013*; *Mikhailova et al., 2009*)? (3) What evolutionary forces drove their divergence in the first place?

Only two species-specific protein markers were detected in our analysis for *L. arcana* (and none for *L. saxatilis*). Both *L. arcana*-specific markers (and one marker of *L. compressa*) proved to belong to PPIase class of enzymes, the cyclophilins (Cyp) family. These ubiquitously distributed enzymes are known to facilitate folding of certain proteins and thus regulate diverse cellular processes (rev. in *Galat & Metcalfe (1995)*; *Fischer & Aumüller (2003)*; *Lu et al. (2007)*). In mollusks, Cyps are expressed in hemocytes and gonads and involved in immune response (*Song et al., 2009*; *Ji et al., 2013*). They also participate in the processes of biomineralization during shell and radula formation (*Jackson et al., 2009*; *Marie et al., 2013*; *Nemoto et al., 2012*). These enzymes are believed to be secreted to spatially organize matrix proteins, and thus to control crystallization. Hypothetically, structural differences in Cyps between closely related species could cause tiny changes in this process resulting in interspecies radula variation, which was

documented for *L. arcana* and *L. compressa* (*Reid, 1996*). In our analyses we used head tissues including odontophore and radular invagination, where radula synthesis occurs. Moreover, Cyp of *L. saxatilis* is expected to be catalytically active as there are no substitutions in catalytically important amino acids (based on Cyp-transcript sequence present in LSD transcriptome (*Canbäck et al., 2012*)). Further studies are needed to examine the precise compartment of expression and functioning of the identified Cyps.

Legacy of post-glacial history in proteome divergence

Overall, samples from different localities grouped by species implying a tight link between genome and proteome and a minor influence of physical factors, locality or season. Nevertheless, the same clustering pattern repetitively appeared within every *Neritrema* species: UK samples formed a distinct clade (Fig. 2), while continental samples clustered more often according to the tissue type. In the pair *L. arcana/L. saxatilis* separation between UK and continental samples even precedes the separation of the species (Fig. 2; Table S2). *L. saxatilis* UK populations have been previously shown genetically distinct from the mainland using mitochondrial markers (*Panova et al., 2011*; *Doellman et al., 2011*) and it has been suggested that UK *L. saxatilis* persisted in the UK within some local refugium during the last glacial maximum with following postglacial expansion on the British Islands, but not to the mainland. European mainland populations of *L. saxatilis* were presumably recolonized from other refugia than the UK (*Panova et al., 2011*; *Doellman et al., 2011*). The existence of a UK refugium has been also suggested for *L. compressa* and *L. arcana* (*Doellman et al., 2011*). Our findings are well in line with such scenario. Moreover, both "obtusata" species have possibly also survived in a hypothetical UK refugium, as a separate UK-clade was also detected in these two species. Recently, a new phylogeographic study suggested the UK as a possible glacial refugium for *L. fabalis* and *L. obtusata* (*Sotelo et al., in press*). Overall, we observed a pattern of geographic isolation at the proteome level, which is very similar to effect of the recent post-glacial history observed in mtDNA. It would be interesting to test this effect further by proteomic analyses of other populations known to have survived in separate glacial refugia, such as *L. saxatilis*, *L. obtusata* and *L. fabalis* in the Iberian peninsula, or *L. saxatilis* and *L. obtusata* in North America.

Qualitative proteomics is not reliable for intergenera analysis

When qualitative data of different genera are put into comparative analysis, proteomics cannot resolve their relationship (unlike the interspecies analysis). This occurs because too many proteins are already changed and too few are conserved, yielding a low number of common "anchoring" proteins between strongly diverged genera and poor grouping of the genera with comparably high binary dissimilarity. This is well illustrated by short genera branches relative to species ones (Figs. 1 and 2). Thus, although in our study genera clustered differently than in DNA-markers-based tree, this result is not sufficient to claim a conflict between the two approaches in this case.

CONCLUSIONS

Here we evaluated the usefulness of the proteomic approach in comparative evolutionary studies. Proteomics has an advantage of the ability to reflect the biochemical machinery of life in comparison to morphological traits. In addition, it records the features undetected by genomic approaches, such as stable patterns of molecular functioning (including those maintained by non-genetic mechanisms, but crucial for typical species physiology). This allows to easily reveal potentially adaptive differences (like fast evolving proteins) in diverging taxa. Owing to incorporation of both slowly and fast evolving proteins, proteomes may be a basis for higher resolution reconstruction than neutral DNA-markers at short evolutionary scales. For the same reason proteome-based analysis is less informative for genera and higher rank taxa, because they might have accumulated too many changed proteins. Our data demonstrate that, although there is a clear connection between the proteome and genome, proteomic and neutral DNA-markers-based evolutionary reconstructions can either agree or contradict each other, depending on whether adaptive processes have been involved in the recent history of the taxa in comparison. Any of the particular outcomes can be informative for the interpretation of species relationships, their evolutionary histories, and the causal and functional background of their evolution.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Arina L. Maltseva conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, funding aquisition, and approved the final draft.
- Marina A. Varfolomeeva conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Arseniy A. Lobov performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Polina Tikanova performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Marina Panova analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Natalia A. Mikhailova conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Andrei I. Granovitch conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, administration of the project, funding aquisition, and approved the final draft.

Field Study Permissions

The following information was supplied relating to field study approvals (i.e., approving body and any Reference Numbers):

Periwinkles are not the endangered or protected species, so no special permission for their collection was required. Administrations of the Swire Institute of Marine Science, Hong Kong University (Hong Kong, China), the Arctic University of Norway University of Tromso (Norway), the White Sea Marine Biological Station of St. Petersburg State University (Russia), the Scotland Association of Marine Science (Oban, Scotland, UK) and Interuniversity Institute for Marine Sciences (Eilat, Israel) were informed about the snails sampling. Collection and animal numbers were approved by those authorities.

No special approval could be obtained for sample collection at the site near Cancale (France) because it is not an area of any national park or a private territory, so there was no appropriate authority to apply for any permission.

Data Availability

The following information was supplied regarding data availability:

A table containing protein spots presence/absence in samples of 10 Littorinid species is available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8546#supplemental-information.

REFERENCES

Aquadro CF, Avise JC. 1981. Genetic divergence between rodent species assessed by using two-dimensional electrophoresis. *Proceedings of the National Academy of Sciences* **78(6)**:3784–3788 DOI 10.1073/pnas.78.6.3784.

Arentz G, Weiland F, Oehler MK, Hoffmann P. 2015. State of the art of 2D DIGE. *PROTEOMICS: Clinical Applications* **9(3–4)**:277–288 DOI 10.1002/prca.201400119.

Baer B, Millar AH. 2016. Proteomics in evolutionary ecology. *Journal of Proteomics* **135**:4–11 DOI 10.1016/j.jprot.2015.09.031.

Beerli P, Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152(2)**:763–773.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57(1)**:289–300 DOI 10.1111/j.2517-6161.1995.tb02031.x.

Berth M, Moser FM, Kolbe M, Bernhardt J. 2007. The state of the art in the analysis of two-dimensional gel electrophoresis images. *Applied Microbiology and Biotechnology* **76(6)**:1223–1243 DOI 10.1007/s00253-007-1128-0.

Biron DG, Loxdale HD, Ponton F, Moura H, Marché L, Brugidou C, Thomas F. 2006. Population proteomics: an emerging discipline to study metapopulation ecology. *Proteomics* **6(6)**:1712–1715 DOI 10.1002/pmic.200500423.

Bonduriansky R, Day T. 2009. Nongenetic inheritance and its evolutionary implications. *Annual Review of Ecology, Evolution, and Systematics* **40(1)**:103–125 DOI 10.1146/annurev.ecolsys.39.110707.173441.

Brito PH, Edwards SV. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135(3)**:439–455 DOI 10.1007/s10709-008-9293-3.

Canbäck B, André C, Galindo J, Johannesson K, Johansson T, Panova M, Tunlid A, Butlin R. 2012. The *Littorina* sequence database (LSD): an online resource for genomic data. *Molecular Ecology Resources* **12(1)**:142–148 DOI 10.1111/j.1755-0998.2011.03042.x.

Cavalli-Sforza LL. 1964. Population structure and human evolution. *Proceedings of the Royal Society* **164**:362–379.

Costa D, Sotelo G, Kaliontzopoulou A, Carvalho J, Butlin R, Hollander J, Faria R. (in press). Hybridization patterns between two marine snails, *Littorina fabalis* and *L. obtusata*. *Ecology and Evolution* DOI 10.1002/ece3.5943.

Conde-Padín P, Caballero A, Rolán-Alvarez E. 2009. Relative role of genetic determination and plastic response during ontogeny for shell-shape traits subjected to diversifying selection. *Evolution: An International Journal of Organic Evolution* **63(5)**:1356–1363 DOI 10.1111/j.1558-5646.2009.00636.x.

Crossland S, Coates D, Grahame J, Mill PJ. 1996. *The Littorina saxatilis species complex—interpretation using random amplified polymorphic DNAs: origin and evolutionary radiation of the Mollusca*. Oxford: Oxford University Press, 205–209.

Danchin É, Charmantier A, Champagne FA, Mesoudi A, Pujol B, Blanchet S. 2011. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews Genetics* **12(7)**:475–486 DOI 10.1038/nrg3028.

De Cáceres M, Legendre P. 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90(12)**:3566–3574 DOI 10.1890/08-1823.1.

Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene trees. *PLOS Genetics* **2(5)**:e68 DOI 10.1371/journal.pgen.0020068.

Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* **24(6)**:332–340 DOI 10.1016/j.tree.2009.01.009.

Diz AP, Martínez-Fernández M, Rolán-Alvarez E. 2012. Proteomics in evolutionary ecology: linking the genotype with the phenotype. *Molecular Ecology* **21(5)**:1060–1080 DOI 10.1111/j.1365-294X.2011.05426.x.

Doellman MM, Trussell GC, Grahame JW, Vollmer SV. 2011. Phylogeographic analysis reveals a deep lineage split within North Atlantic *Littorina saxatilis*. *Proceedings of the Royal Society B: Biological Sciences* **278(1722)**:3175–3183 DOI 10.1098/rspb.2011.0346.

Dworzanski JP, Snyder AP. 2005. Classification and identification of bacteria using mass spectrometry-based proteomics. *Expert Review of Proteomics* **2(6)**:863–878 DOI 10.1586/14789450.2.6.863.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63(1)**:1–19 DOI 10.1111/j.1558-5646.2008.00549.x.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution: International Journal of Organic Evolution* **63(1)**:1–19 DOI 10.1111/j.1558-5646.2008.00549.x.

Edwards D, Batley J. 2004. Plant bioinformatics: from genome to phenome. *Trends in Biotechnology* **22(5)**:232–237 DOI 10.1016/j.tibtech.2004.03.002.

Efron B, Halloran E, Holmes S. 1996. Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences USA* **93(23)**:13429 DOI 10.1073/pnas.93.23.13429.

Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology & Evolution* **26(6)**:298–306 DOI 10.1016/j.tree.2011.02.008.

Fang B, Merilä J, Ribeiro F, Alexandre CM, Momigliano P. 2018. Worldwide phylogeny of three-spined sticklebacks. *Molecular Phylogenetics and Evolution* **127**:613–625 DOI 10.1016/j.ympev.2018.06.008.

Feder JL, Egan SP, Nosil P. 2012. The genomics of speciation-with-gene-flow. *Trends in Genetics* **28(7)**:342–350 DOI 10.1016/j.tig.2012.03.009.

Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* **22(1)**:521–565 DOI 10.1146/annurev.ge.22.120188.002513.

Feng S, Bootsma M, McLellan SL. 2018. Human-associated Lachnospiraceae genetic markers improve detection of fecal pollution sources in urban waters. *Applied and Environmental Microbiology* **84(14)**:e00309-18 DOI 10.1128/AEM.00309-18.

Fischer G, Aumüller T. 2003. *Regulation of peptide bond cis/trans isomerization by enzyme catalysis and its implication in physiological processes: reviews of physiology, biochemistry and pharmacology*. Berlin: Springer, 105–150.

Fowlkes EB, Mallows CL. 1983. A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association* **78(383)**:553–569 DOI 10.1080/01621459.1983.10478008.

Galat A, Metcalfe SM. 1995. Peptidylproline cis/trans isomerases. *Progress in Biophysics and Molecular Biology* **63(1)**:67–118 DOI 10.1016/0079-6107(94)00009-X.

Galili T. 2015. Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31(22)**:3718–3720 DOI 10.1093/bioinformatics/btv428.

García-Souto D, Alonso-Rubido S, Costa D, Eirín-López J, Rolán-Álvarez E, Faria R, Galindo J, Pasantes J. 2018. Karyotype characterization of nine periwinkle species (Gastropoda, Littorinidae). *Genes* **9(11)**:517 DOI 10.3390/genes9110517.

Granovitch AI, Loskutova ZI, Gracheva YA, Mikhailova NA. 2008. Morphometric comparison of the copulatory organ in mollusks of saxatilis species complex (Caenogastropoda: Littorinidae): problems of identification of species and species status. *Zoologichesky Zhurnal* **87(12)**:1425–1436.

Granovitch AI, Maximovich AN, Avanesyan AV, Starunova ZI, Mikhailova NA. 2013. Micro-spatial distribution of two sibling periwinkle species across the intertidal indicates hybrdization. *Genetica* **141(7–9)**:293–301 DOI 10.1007/s10709-013-9728-3.

Granovitch AI, Mikhailova NA, Znamenskaya O, Petrova YA. 2004. Species complex of mollusks of the genus *Littorina* (Gastropoda: Prosobranchia) from the eastern Murman coast. *Zoologichesky Zhurnal* **83(11)**:1305–1316 DOI 10.1007/s10709-009-9397-4.

Granovitch AI, Yagunova EB, Maximovich AN, Sokolova IM. 2009. Elevated female fecundity as a possible compensatory mechanism in response to trematode infestation in populations of *Littorina saxatilis* (Olivi). *International Journal for Parasitology* **39(9)**:1011–1019 DOI 10.1016/j.ijpara.2009.02.014.

Greenbaum D, Colangelo C, Williams K, Gerstein M. 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biology* **4(9)**:117 DOI 10.1186/gb-2003-4-9-117.

Grover A, Sharma PC. 2016. Development and use of molecular markers: past and present. *Critical Reviews in Biotechnology* **36(2)**:290–302 DOI 10.3109/07388551.2014.959891.

Gygi SP, Rochon Y, Franza BR, Aebersold R. 1999. Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* **19(3)**:1720–1730 DOI 10.1128/MCB.19.3.1720.

Hennig C. 2007. Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis* **52(1)**:258–271 DOI 10.1016/j.csda.2006.11.025.

Hennig C. 2008. Dissolution point and isolation robustness: robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis* **99(6)**:1154–1176 DOI 10.1016/j.jmva.2007.07.002.

Hennig C. 2015. *FPC: Flexible Procedures for Clustering*. Version 2.1-10. *Available at https://cran.r-project.org/web/packages/fpc/index.html*.

Jackson DJ, McDougall C, Woodcroft B, Moase P, Rose RA, Kube M, Reinhardt R, Rokhsar DS, Montagnani C, Joubert C, Piquemal D, Degnan BM. 2009. Parallel evolution of nacre building gene sets in molluscs. *Molecular Biology and Evolution* **27(3)**:591–608 DOI 10.1093/molbev/msp278.

Jewett EM, Rosenberg NA. 2012. iGLASS: an improvement to the GLASS method for estimating species trees from gene trees. *Journal of Computational Biology* **19(3)**:293–315 DOI 10.1089/cmb.2011.0231.

Ji C, Wu H, Wei L, Zhao J, Wang Q, Lu H. 2013. Responses of *Mytilus galloprovincialis* to bacterial challenges by metabolomics and proteomics. *Fish & Shellfish Immunology* **35(2)**:489–498 DOI 10.1016/j.fsi.2013.05.009.

Johannesson K, Panova M, Kemppainen P, André C, Rolán-Alvarez E, Butlin RK. 2010. Repeated evolution of reproductive isolation in a marine snail: unveiling mechanisms of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365(1547)**:1735–1747 DOI 10.1098/rstb.2009.0256.

Karp NA, Spencer M, Lindsay H, O'Dell K, Lilley KS. 2005. Impact of replicate types on proteomic expression analysis. *Journal of Proteome Research* **4(5)**:1867–1871 DOI 10.1021/pr050084g.

Karr TL. 2008. Application of proteomics to ecology and population biology. *Heredity* **100(2)**:200–206 DOI 10.1038/sj.hdy.6801008.

Kemppainen P, Panova M, Hollander J, Johannesson K. 2009. Complete lack of mitochondrial divergence between two species of NE Atlantic marine intertidal gastropods. *Journal of Evolutionary Biology* **22(10)**:2000–2011 DOI 10.1111/j.1420-9101.2009.01810.x.

Kim GH, Shim JB, Klochkova TA, West JA, Zuccarello GC. 2008. The utility of proteomics in algal taxonomy: *Bostrychia radicans/B. moritziana* (Rhodomelaceae, Rhodophyta) as a model study. *Journal of Phycology* **44(6)**:1519–1528 DOI 10.1111/j.1529-8817.2008.00592.x.

Kimura M. 1968. Evolutionary rate at the molecular level. *Nature* **217(5129)**:624–626 DOI 10.1038/217624a0.

Knight AJ, Ward RD. 1991. The genetic relationships of three taxa in the *Littorina saxatilis* species complex (Prosobranchia: Littorinidae). *Journal of Molluscan Studies* **57(1)**:81–91 DOI 10.1093/mollus/57.1.81.

Kuhner MK, Yamato J, Felsenstein J. 1998. Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149(1)**:429–434.

Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, Jablonka E, Odling-Smee J. 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences* **282(1813)**:20151019 DOI 10.1098/rspb.2015.1019.

Laland K, Uller T, Feldman M, Sterelny K, Müller GB, Moczek A, Jablonka E, Odling-Smee J, Wray GA, Hoekstra HE, Futuyma DJ, Lenski RE, Mackay TFC, Schluter D, Strassmann JE. 2014. Does evolutionary theory need a rethink? *Nature* **514(7521)**:161–164 DOI 10.1038/514161a.

Lee OHK, Williams GA. 2002a. Spatial distribution patterns of *Littoraria* species in Hong Kong mangroves. *Hydrobiologia* **481(1–3)**:137–145 DOI 10.1023/A:1021241810526.

Lee OHK, Williams GA. 2002b. Locomotor activity patterns of the mangrove littorinids, *Littoraria ardouiniana* and *L. melanostoma*, in Hong Kong. *Journal of Molluscan Studies* **68(3)**:235–241 DOI 10.1093/mollus/68.3.235.

Lee OH, Williams GA, Hyde KD. 2001. The diets of *Littoraria ardouiniana* and *L. melanostoma* in Hong Kong mangroves. *Journal of the Marine Biological Association of the United Kingdom* **81(6)**:967–973 DOI 10.1017/S002531540100491X.

Lemmon EM, Lemmon AR. 2013. High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* **44(1)**:99–121 DOI 10.1146/annurev-ecolsys-110512-135822.

Lilley KS, Friedman DB. 2004. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Review of Proteomics* **1(4)**:401–409 DOI 10.1586/14789450.1.4.401.

Lobov AA, Maltseva AL, Mikhailova NA, Granovitch AI. 2015. LOSP: a newly identified sperm protein from *Littorina obtusata*. *Journal of Molluscan Studies* **81(4)**:512–515 DOI 10.1093/mollus/eyv010.

Lobov AA, Maltseva AL, Starunov VV, Babkina IY, Ivanov VA, Mikhailova NA, Granovitch AI. 2018. LOSP: a putative marker of parasperm lineage in male reproductive system of the prosobranch mollusk *Littorina obtusata*. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **330(4)**:193–201 DOI 10.1002/jez.b.22803.

Lu KP, Finn G, Lee TH, Nicholson LK. 2007. Prolyl cis-trans isomerization as a molecular timer. *Nature Chemical Biology* **3(10)**:619–629 DOI 10.1038/nchembio.2007.35.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4(12)**:981–994 DOI 10.1038/nrg1226.

López JL. 2005. Role of proteomics in taxonomy: the Mytilus complex as a model of study. *Journal of Chromatography B* **815(1–2)**:261–274 DOI 10.1016/j.jchromb.2004.10.064.

López JL. 2007. Applications of proteomics in marine ecology. *Marine Ecology Progress Series* **332**:275–279 DOI 10.3354/meps332275.

Mack PD, Kapelnikov A, Heifetz Y, Bender M. 2006. Mating-responsive genes in reproductive tissues of female *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences* **103(27)**:10358–10363 DOI 10.1073/pnas.0604046103.

Maddison WP. 1997. Gene trees in species trees. *Systematic biology* **46(3)**:523–536 DOI 10.1093/sysbio/46.3.523.

Maier T, Güell M, Serrano L. 2009. Correlation of mRNA and protein in complex biological samples. *FEBS Letters* **583(24)**:3966–3973 DOI 10.1016/j.febslet.2009.10.036.

Maltseva AL, Varfolomeeva MA, Lobov AA, Mikhailova NA, Renaud PE, Grishankov AV, Volovik KY, Granovitch AI. 2016. Measuring physiological similarity of closely related littorinid species: a proteomic insight. *Marine Ecology Progress Series* **552**:177–193 DOI 10.3354/meps11770.

Mameli M. 2004. Non-genetic selection and nongenetic inheritance. *British Journal for the Philosophy of Science* **55(1)**:35–71 DOI 10.1093/bjps/55.1.35.

Marie B, Jackson DJ, Ramos-Silva P, Zanella-Cléon I, Guichard N, Marin F. 2013. The shell-forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-specific novelties. *FEBS Journal* **280(1)**:214–232 DOI 10.1111/febs.12062.

Martin A, Orgogozo V. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67(5)**:1235–1250 DOI 10.1111/evo.12081.

Martínez-Fernández M, De la Cadena M, Rolán-Alvarez E. 2010. The role of phenotypic plasticity on the proteome differences between two sympatric marine snail ecotypes adapted to distinct micro-habitats. *BMC Evolutionary Biology* **10(1)**:65 DOI 10.1186/1471-2148-10-65.

McMahon RF. 1988. Respiratory response to periodic emergence in intertidal molluscs. *American Zoologist* **28(1)**:97–114 DOI 10.1093/icb/28.1.97.

Messana I, Cabras T, Iavarone F, Vincenzoni F, Urbani A, Castagnola M. 2013. Unraveling the different proteomic platforms. *Journal of Separation Science* **36(1)**:128–139 DOI 10.1002/jssc.201200830.

Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL. 2010. Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences* **107(21)**:9724–9729 DOI 10.1073/pnas.1000939107.

Mikhailova NA, Gracheva YA, Backeljau T, Granovitch AI. 2009. A potential species-specific molecular marker suggests interspecific hybridization between sibling species *Littorina arcana* and *L. saxatilis* (Mollusca, Caenogastropoda) in natural populations. *Genetica* **137(3)**:333–340 DOI 10.1007/s10709-009-9397-4.

Muraeva OA, Maltseva AL, Mikhailova NA, Granovitch AI. 2016. Mechanisms of adaption to salinity stress in marine gastropods *Littorina saxatilis*: a proteomic analysis. *Cell and Tissue Biology* **10(2)**:160–169 DOI 10.1134/S1990519X16020085.

Nemoto M, Wang Q, Li D, Pan S, Matsunaga T, Kisailus D. 2012. Proteomic analysis from the mineralized radular teeth of the giant Pacific chiton, *Cryptochiton stelleri* (Mollusca). *Proteomics* **12(18)**:2890–2894 DOI 10.1002/pmic.201100473.

Ng TPT, Davies MS, Stafford R, Williams GA. 2011. Mucus trail following as a mate-searching strategy in mangrove littorinid snails. *Animal Behaviour* **82(3)**:459–465 DOI 10.1016/j.anbehav.2011.05.017.

Noor MAF, Feder JL. 2006. Speciation genetics: evolving approaches. *Nature Reviews Genetics* **7(11)**:851–861 DOI 10.1038/nrg1968.

Nosil P. 2012. *Ecological speciation*. Oxford: Oxford University Press, 304.

Ohnishi S, Kawanishi M, Watanabe TK. 1983. Biochemical phylogenies of *Drosophila*: protein differences detected by two-dimensional electrophoresis. *Genetica* **61(1)**:55–63 DOI 10.1007/BF00563232.

Panova M, Blakeslee AMH, Miller AW, Mäkinen T, Ruiz GM, Johannesson K, André C. 2011. Glacial history of the North Atlantic marine snail, *Littorina saxatilis*, inferred from distribution of mitochondrial DNA lineages. *PLOS ONE* **6(3)**:e17511 DOI 10.1371/journal.pone.0017511.

Panova M, Boström J, Hofving T, Areskoug T, Eriksson A, Mehlig B, Mäkinen T, André C, Johannesson K. 2010. Extreme female promiscuity in a non-social invertebrate species. *PLOS ONE* **5(3)**:e9640 DOI 10.1371/journal.pone.0009640.

Panova M, Johansson T, Canbäck B, Bentzer J, Rosenblad MA, Johannesson K, Tunlid A, André C. 2014. Species and gene divergence in *Littorina* snails detected by array comparative genomic hybridization. *BMC Genomics* **15(1)**:687 DOI 10.1186/1471-2164-15-687.

Paradis E, Schliep K. 2018. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35(3)**:526–528 DOI 10.1093/bioinformatics/bty633.

Patti GJ, Yanes O, Siuzdak G. 2012. Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* **13(4)**:263–269 DOI 10.1038/nrm3314.

Patwardhan A, Ray S, Roy A. 2014. Molecular markers in phylogenetic studies: a review. *Journal of Phylogenetics & Evolutionary Biology* **2(2)**:131 DOI 10.4172/2329-9002.1000131.

Pigliucci M. 2009. An extended synthesis for evolutionary biology. *Annals of the New York Academy of Sciences* **1168(1)**:218–228 DOI 10.1111/j.1749-6632.2009.04578.x.

Pigliucci M, Müller G. 2010. *Evolution-the extended synthesis*. Cambridge: MIT Press, 495.

R Core Team. 2015. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. *Available at* https://www.R-project.org/.

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, Noor MAF, Mehlig B, Westram AM. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30(8)**:1450–1477 DOI 10.1111/jeb.13047.

Reid DG. 1986. *Mainwaringia* Nevill, 1885, a littorinid genus from Asiatic mangrove forests, and a case of protandrous hermaphroditism. *Journal of Molluscan Studies* **52(3)**:225–242 DOI 10.1093/mollus/52.3.225.

Reid DG. 1989. The comparative morphology, phylogeny and evolution of the gastropod family Littorinidae. *Philosophical Transactions of the Royal Society B: Biological Sciences* **324(1220)**:1–110 DOI 10.1098/rstb.1989.0040.

Reid DG. 1996. *Systematics and Evolution of Littorina*. London: Ray Society.

Reid DG. 2007. *The genus Echinolittorina Habe, 1956 (Gastropoda: Littorinidae) in the Indo-West Pacific Ocean. Zootaxa 1420*. Auckland: Magnolia Press, 161.

Reid DG, Dyal P, Williams ST. 2010. Global diversification of mangrove fauna: a molecular phylogeny of *Littoraria* (Gastropoda: Littorinidae). *Molecular Phylogenetics and Evolution* **55(1)**:185–201 DOI 10.1016/j.ympev.2009.09.036.

Reid DG, Dyal P, Williams ST. 2012. A global molecular phylogeny of 147 periwinkle species (Gastropoda, Littorininae). *Zoologica Scripta* **41(2)**:125–136 DOI 10.1111/j.1463-6409.2011.00505.x.

Robinson DF, Foulds LR. 1981. Comparison of phylogenetic trees. *Mathematical Biosciences* **53(1–2)**:131–147 DOI 10.1016/0025-5564(81)90043-2.

Rolán-Alvarez E, Austin C, Boulding EG. 2015. The contribution of the genus *Littorina* to the field of evolutionary ecology. *Oceanography and Marine Biology* **53**:157–214.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19(12)**:1572–1574 DOI 10.1093/bioinformatics/btg180.

Rosenberg NA. 2013. Discordance of species trees with their most likely gene trees: a unifying principle. *Molecular Biology and Evolution* **30(12)**:2709–2713 DOI 10.1093/molbev/mst160.

Salzburger W. 2018. Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics* **19(11)**:705–717 DOI 10.1038/s41576-018-0043-9.

Schliep KP. 2011. Phangorn: phylogenetic analysis in R. *Bioinformatics* **27(4)**:592–593 DOI 10.1093/bioinformatics/btq706.

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science* **323(5915)**:737–741 DOI 10.1126/science.1160006.

Schlötterer C. 2004. The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics* **5(1)**:63–69 DOI 10.1038/nrg1249.

Shimodaira H. 2004. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statistics* **32(6)**:2616–2641 DOI 10.1214/009053604000000823.

Small MP, Gosling EM. 2000. Genetic structure and relationships in the snail species complex *Littorina arcana* Hannaford Ellis, *L. compressa* Jeffreys and *L. saxatilis* (Olivi) in the British Isles using SSCPs of cytochrome-b fragments. *Heredity* **84**:692–701 DOI 10.1046/j.1365-2540.2000.00717.x.

Sokal RR, Rohlf FJ. 1962. The comparison of dendrograms by objective methods. *Taxon* **11(2)**:33–40 DOI 10.2307/1217208.

Sokolova IM, Berger VJ. 2000. Physiological variation related to shell colour polymorphism in White Sea *Littorina saxatilis*. *Journal of Experimental Marine Biology and Ecology* **245(1)**:1–23 DOI 10.1016/S0022-0981(99)00132-X.

Sokolova IM, Pörtner HO. 2001a. Physiological adaptations to high intertidal life involve improved water conservation abilities and metabolic rate depression in *Littorina saxatilis*. *Marine Ecology Progress Series* **224**:171–186 DOI 10.3354/meps224171.

Sokolova I, Pörtner HO. 2001b. Temperature effects on key metabolic enzymes in *Littorina saxatilis* and *L. obtusata* from different latitudes and shore levels. *Marine Biology* **139(1)**:113–126 DOI 10.1007/s002270100557.

Song X, Wang L, Song L, Zhao J, Zhang H, Zheng P, Qiu L, Liu X, Wu L. 2009. A cyclophilin A inducible expressed in gonad of zhikong scallop *Chlamys farreri*. *Molecular Biology Reports* **36(6)**:1637–1645 DOI 10.1007/s11033-008-9363-8.

Sotelo G, Duvetorp M, Costa D, Panova M, Johannesson K, Faria R. Phylogeographic history of flat periwinkles, *Littorina fabalis* and *L. obtusata*. *BMC Evolutionary Biology* (in press).

Starr TN, Thornton JW. 2016. Epistasis in protein evolution. *Protein Science* **25(7)**:1204–1218 DOI 10.1002/pro.2897.

Storey KB, Lant B, Anozie OO, Storey JM. 2013. Metabolic mechanisms for anoxia tolerance and freezing survival in the intertidal gastropod, *Littorina littorea*. *Comparative Biochemistry and*

Physiology Part A: Molecular & Integrative Physiology **165(4)**:448–459
DOI 10.1016/j.cbpa.2013.03.009.

Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nature Reviews Genetics* **17(4)**:239–250 DOI 10.1038/nrg.2016.11.

Suzuki R, Shimodaira H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22(12)**:1540–1542 DOI 10.1093/bioinformatics/btl117.

Uda K, Fujimoto N, Akiyama Y, Mizuta K, Tanaka K, Ellington WR, Suzuki T. 2006. Evolution of the arginine kinase gene family. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **1(2)**:209–218 DOI 10.1016/j.cbd.2005.10.007.

Ünlü M, Morgan ME, Minden JS. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18(11)**:2071–2077 DOI 10.1002/elps.1150181133.

Via S. 2009. Natural selection in action during speciation. *Proceedings of the National Academy of Sciences* **106(Suppl. 1)**:9939–9946 DOI 10.1073/pnas.0901397106.

Waters KM, Pounds JG, Thrall BD. 2006. Data merging for integrated microarray and proteomic analysis. *Briefings in Functional Genomics and Proteomics* **5(4)**:261–272 DOI 10.1093/bfgp/ell019.

Wilding CS, Grahame J, Mill PJ. 2000a. Mitochondrial DNA CoI haplotype variation in sibling species of rough periwinkles. *Heredity* **85(1)**:62–74 DOI 10.1046/j.1365-2540.2000.00718.x.

Wilding CS, Grahame J, Mill PJ. 2000b. Nuclear DNA restriction site polymorphisms and the phylogeny and population C structure of an intertidal snail species complex (*Littorina*). *Hereditas* **133(1)**:9–18 DOI 10.1111/j.1601-5223.2000.00009.x.

Zou Z, Zhang J. 2015. Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? *Molecular Biology and Evolution* **32(8)**:2085–2096 DOI 10.1093/molbev/msv091.

Zuckerkandl E, Pauling L. 1962. Molecular disease, evolution and genic heterogeneity. In: Kasha M, Pullman B, eds. *Horizons in Biochemistry*. New York: Academic Press, 189–225.