Signatures of the evolution of parthenogenesis and cryptobiosis in the genomes of panagrolaimid nematodes

Philipp H. Schiffer[1,2*§], Etienne G.J. Danchin[3], Ann M. Burnell[4], Anne-Marike Schiffer[5], Christopher J. Creevey[6], Simon Wong[7], Ilona Dix[4], Georgina O'Mahony[4], Bridget A. Culleton[4], Corinne Rancurel[3], Gary Stier[1], Elizabeth A. Martínez-Salazar[9], Aleksandra Marconi[10], Urmi Trivedi[11], Michael Kroiher[1], Michael A.S. Thorne[12], Einhard Schierenberg[1], Thomas Wiehe[2], Mark Blaxter[10,11]

1 Zoologisches Institut, Universität zu Köln, 50764 Köln, Germany

2 Institut für Genetik, Universität zu Köln, 50674 Köln, Germany

3 INRA, Université Côte d'Azur, CNRS, ISA, France

4 Maynooth University Department of Biology, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland

5 Department of Experimental Psychology, University of Oxford & Division of Psychology, Department of Life Science, Brunel University London

6 Institute of Biological, Environment and Rural Sciences, Aberystwyth University, Aberystwyth, Ceredigion, UK

7 Irish Centre for High-End Computing, Tower Building, Trinity Technology & Enterprise Campus, Grand Canal Quay, Dublin, D02 HP83, Ireland

8 Megazyme, Bray Business Park, Bray, Co. Wicklow, A98 YV29, Ireland

9 Unidad Académica de Ciencias Biológicas, Laboratorio de Colecciones Biológicas y Sistemática Molecular, Universidad Autónoma de Zacatecas, Zacatecas, México

10 Institute of Evolutionary Biology, The University of Edinburgh, Edinburgh EH9 3FL, UK

11 Edinburgh Genomics, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, UK

12 British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

* Corresponding author: philipp.schiffer@gmail.com

§ Current addresses:

PS Department for Biosciences, University College London, London, UK

ORCiDs: http://orcid.org/
PHS: 0000-0001-6776-0934
EGJD: 0000-0003-4146-5608
CJC: 0000-0001-7183-1555
CR: 0000-0001-9355-5491
EAMS: 0000-0001-5183-6930
AM: 0000-0001-8481-9336
MAST: 0000-0001-7759-612X
ES: 0000-0003-1823-748X
TW: 0000-0002-8932-2772
MB: 0000-0003-2861-949X

Emails:
PHS: philipp.schiffer@gmail.com
ED: etienne.danchin@inra.fr
AB: ann.burnell@nuim.ie
AMS: anne-marike.schiffer@brunel.ac.uk
CJC: chc30@aber.ac.uk
SW: simon.wong@ichec.ie
ID: ilona.dix@nuim.ie
GM: georginaomahony@gmail.com
BAC: bridgetculleton@gmail.com
CR: corinne.rancurel@inra.fr
GS: gary.koeln@googlemail.com
EAMS: eamsuaz@gmail.com
AM: amarconi@ed-alumni.net
UT: urmi.trivedi@ed.ac.uk
MK: m.kroiher@uni-koeln.de
MAST: mior@bas.ac.uk
ES: e.schierenberg@uni-koeln.de
TW: t.wiehe@uni-koeln.de
MB: mark.blaxter@ed.ac.uk

Keywords: Parthenogenesis, Anhydrobiosis, Horizontal Gene Transfer, Nematodes, Evolution, Genomes

Abstract

Most animal species reproduce sexually, but parthenogenesis, asexual reproduction of various forms, has arisen repeatedly. Parthenogenetic lineages are usually short lived, though in some environments parthenogenesis may be advantageous. *Panagrolaimus* nematodes have colonised environments ranging from arid deserts to arctic and antarctic biomes. Many are parthenogens, and most have cryptobiotic (anhydrobiotic and/or cryobiotic) abilities. We compared the genomes and transcriptomes of parthenogenetic and sexual *Panagrolaimus* species to identify systems that contribute to these striking abilities. The parthenogens are triploids of hybrid origin. All *Panagrolaimus* have acquired genes through horizontal gene transfer, some of which are likely to contribute to cryptobiosis.

Introduction

Despite the general prevalence of sexual reproduction in animals, parthenogenesis has evolved in many taxa. Parthenogenesis, or, more strictly the cessation of outcrossing after meiotic recombination, results in the gradual accumulation of deleterious mutation that cannot be purged independently of all linked loci[1]. This condemns non-recombining parthenogens to evolutionary failure through Muller's ratchet[2]. The evolutionary scandals of the phylogenetic persistence, and abundant speciation, of some asexual taxa (such as bdelloid rotifers and darwinulid ostracods) may be explained by the recruitment of recombination-like mechanisms such as gene conversion and inter-individual horizontal gene transfers[3]. Parthenogens have been observed to amplify the number of gene or genome copies through polyploidy or aneuploidy, and this may buffer lineages against mutation accumulation in essential genes[4]. Comparisons of related sexual and asexual taxa are required to understand the origins and persistence of asexuality, and the effects of asexuality on genome evolution.

Asexual reproduction has long been associated with organisms found in extreme or peripheral environments, such as arid deserts where water availability is ephemeral and unpredictable, or polar ecosystems where much of the annual cycle involves temperatures well below zero °C. Parthenogens are able to colonise ephemeral or widely spaced niches as single reproductive propagules[5]. In turn, organisms that inhabit or colonise extreme environments, are often cryptobionts. Cryptobiosis is the ability of an organism to survive with no evidence of active metabolism under conditions inimical to life. Cryobionts survive freezing, avoiding lethal ice-crystal disruption of biological membranes, while anhydrobionts survive loss of effectively all their body water. Cryobionts are frequently also anhydrobiotic, as ice sublimation in very low temperatures can freeze-dry. The mechanistic bases of cryptobiosis have been explored in some species (*Aphelenchus* nematodes[6], the chironomid *Polypedilum vanderplanki*[7], and the tardigrade *Ramazzottius varieornatus* and *Hybsibius dujardini*[8]) and individual molecular players identified. Intrinsically unstructured proteins and trehalose are important molecular components in anhydrobiotic mechanisms. Intrinsically disordered proteins do not adopt a stable secondary structure in solution, and are thought to replace structural water on dehydration. They include the late embryogenesis abundant (LEA)-like proteins[7, 9,10], and

the tardigrade-specific SAHS, MAHS and CAHS family proteins[8]. Trehalose is likely to act in a similar way. While some species are able to undergo cryptobiosis rapidly at any stage of development[8], others require extensive preconditioning, presumably to activate physiological systems that produce anhydro- or cryo-protectants.

The association between asexuality and cryptobiosis may be mechanistically unlinked, as both traits are advantageous to organisms colonising extreme habitats. Alternatively, the mechanisms giving rise to asexuality may have made acquisition of cryptobiosis traits easier. In the bdelloid rotifers it has been suggested that DNA breaks resulting from the stress of cryptobiosis coupled with a gene-conversion mechanism for avoidance of Muller's ratchet may have promoted the acquisition of foreign DNA (horizontal gene transfer), which in turn conferred improved cryptobiotic potential[3].

Many asexual and/or cryptobiotic taxa of nematodes have been described. The Panagrolaimidae includes both sexual (amphimictic) and parthenogenetic taxa. Many, perhaps all, panagrolaims are cryptobionts, with both cryobiotic and anhydrobiotic abilities, and are found in extreme environments[11]. Although parthenogenetic species are widely distributed around the globe, a single origin of parthenogenesis in panagrolaims is likely[12]. The evolutionary age of the asexual panagrolaims is unknown. Panagrolaims are very good cryptobionts: some species can survive very rapid desiccation, and many species can survive desiccation at any life-cycle stage without pre-conditioning[11].

Here we compare the genomes and transcriptomes of parthenogenetic and amphimictic panagrolaims, including a *Propanagrolaimus* species which is not a good cryptobiont. We identify a single, recent origin of asexual *Panagrolaimus* species, possibly involving triploidisation. We identify candidate genes that may be important for cryptobiosis, some of which are likely to have been acquired horizontally. The genome and transcriptome data and analysis presented here constitute a resource for future functional studies into the evolutionary origins of parthenogenesis and cryptobiosis.

Results

New genomes and transcriptomes from panagrolaimid nematodes

We sequenced and assembled *de novo* the genomes or transcriptomes of five panagrolaimid nematodes (**Fig. 1**) using contamination- and heterozygosity-aware methodologies. We reassembled the genome of *Panagrolaimus davidi* DAW1[13]. Genome assembly spans ranged from 49 Mb to 118 Mb.

We independently estimated the genome sizes of strains in culture using Feulgen Image Analysis Densitometry (FIAD) and quantitative PCR (RT-D-PCR). For the amphimictic *Panagrolaimus* isolates diploid genome sizes were estimated to be between 0.14 (*P.* sp. ES5, using FIAD) and 0.18 (*P. superbus*, using RT-D-PCR) pg DNA per nucleus, or ~140-180 Mb. For the parthenogens (*P.* sp. PS1159, PS1579), we obtained FIAD values between 0.23 and 0.27 pg DNA per nucleus (~230-270 Mb), i.e. ~50% larger than the amphimictic strains. The diploid genome size of *Propanagrolaimus* sp. JU765 was estimated at 0.13 pg DNA per nucleus (130 Mb).

The spans of the haploid genome assemblies of *P.* sp. ES5 and *P. superbus* (90 Mb and 76 Mb respectively, **Fig. 1**) are congruent with the sizes estimated by densitometry and PCR. Similarly, the haploid genome span of the hermaphrodite outgroup *Prop.* sp. JU765 was 64 Mb, in line with the FIAD estimate. The correspondence between assembly span and measured nuclear DNA mass was less clear-cut in the parthenogens. The assembly span of *P.* sp. PS1159 was only 85 Mb, similar to that of the haploid genomes of the sexual species but one third of the physical measurement. The assembly span for *P. davidi* was 175 Mb prior to cleaning and removal of heterozygosity and 118 Mb in the final scaffolding.

Evidence for triploidy in parthenogenetic *Panagrolaimus*

We karyotyped 15 *Panagrolaimus* species and found that all the amphimictic species only had a diploid set of 8 chromosomes (2n=8), while all the parthenogenetic species had 12 chromosomes in total (xn=12) (**Fig. 1**, Supplementary Figure 1). This could correspond to a shift in haploid chromosome number from n=4 to n=6, or triploidy (where n=4 is maintained but somatic nuclei are 3n). Observation of polar bodies in the parthenogenetic species shows that meiosis is present (**Fig. 1**). To test for possible

triploidy in the parthenogens we examined read depth coverage of variants in *P.* sp. PS1159 and presence of homeologous loci in *P. davidi*. Read mapping and variant calling on the assembled genomes yielded the expected result for the amphimictic taxa (*P.* sp. ES5 and *P. superbus*), where heterozygous sites had a mode of variants present at 50%:50%. However, in the parthenogenetic taxa we observed a clear peak in minor allele frequencies at ~33%, consistent with the presence of three copies at these sites (**Fig. 2**). Using a two sample Kolmogorov-Smirnov test we confirmed that the allele frequency spectra of the parthenogenetic and amphimictic *Panagrolaimus* species, as well as the parthenogenetic *Panagrolaimus* and androdioecious *Propanagrolaimus*, belong to two different distributions (see Supplement). These data support triploidy in the parthenogenetic *Panagrolaimus* species.

We performed gene-finding on the new and reassembled genomes, and predicted proteins from the transcriptomes of *P.* sp. DL137 and *P.* sp. PS1579. The number of gene models for the panagrolaimid species had poor correlation with their genome spans (**Fig. 1**, Supplement). We retrieved domain annotations for 77% - 88% of the predicted genes in each species. These were clustered using OrthoFinder and analysed with kinfin (D. Laetsch, M. Blaxter, unpublished; http://doi.org/10.5281/zenodo.290589), including proteomes from *Prop.* sp. JU765 (this work), and other Tylenchina (Clade IV) nematodes (*Panagrellus redivivus*[14], *Meloidogyne hapla*[15], *Acrobeloides nanus* [Schiffer, Polsky et al. in final prep.] and *Bursaphelenchus xylophilus*[16]) as outgroups. In over 1,600 Orthofinder groups we found an excess of proteins for each of the parthenogenetic panagrolaims. Using the gene-tree reconciliation approach implemented in GRAMPA[17] and the OrthoFinder derived FastME[18] trees, the most parsimonious placement of the source of a second set of proteins in the parthenogens was basal to a clade containing both parthenogenetic and sexual species (**Fig. 1**). It is thus unlikely that the extra copies arose by whole genome duplication within the parthenogenetic lineage, and this topology supports an allopolyploid origin of the parthenogens.

Most parthenogenetic species are expected to be young. To measure divergence time between the parthenogenetic and amphimictic *Panagrolaimus*, we implemented an approach based on genome wide divergence. As calibration we used available divergence time estimates for *Caenorhabditis* species[19], corrected by generation time measurements

in *Panagrolaimus* and *Propanagrolaimus* (Supplementary). The two amphimictic species *P. superbus* and *P.* sp. ES5 were the closest sisters (Table 1, **Fig. 2**, also compare to SFig. 2). Assuming an average of 8 days per generation, the split between parthenogenetic and gonochoristic *Panagrolaimus* species is estimated to have occurred 1.3 – 1.4 Myr ago. Using a very conservative measure of 50 days per generation this split would have occurred 8 to 8.5 Myr ago.

Response of panagrolaim genomes to triploidy and parthenogenesis

The evolution of parthenogenesis through hybridisation is rapid. Newly formed hybrids may undergo genome and transcriptome shock[14], as co-adapted complexes can be disrupted by the new ploidy, and interacting partners that have accumulated differences through drift are brought into association. Such maladapted interactions need to be eliminated or moderated, and, in the case of parthenogens especially, the machineries of sex determination, fertilisation/oocyte activation and development are likely to undergo accelerated evolution. In parthenogens, the cost of producing males, which play no genetic part in species fitness is particularly high.

No males have been observed by us in any parthenogenetic *Panagrolaimus* species, not even under stress conditions (e.g. prolonged culturing at elevated temperatures or after reviving cultures from cryptobiosis). We therefore examined several relevant genetic regulatory networks (GRNs) in the parthenogens to identify changes consequent on the evolution of polyploidy and loss of sex. We carried out a strict orthology analysis using OrthoMCL[20] with validation by Orthoinspector[21]. Using GRN defined by experimental intervention in *C. elegans* we identified homologous systems in the panagrolaims, and compared parthenogenetic and sexual *Panagrolaimus* species. Surprisingly, we found that many *C. elegans* genes acting in important developmental processes were absent not only in all *Panagrolaimus*, but also other Tylenchina species. For example, in the sex determination GRN, tylenchine nematodes lacked orthologues of the *C. elegans* master regulator *xol-1*, as well as *fem-1*, *fem-3*, *her-1*, and other genes. Similarly, in the endoderm- and mesoderm-forming GRN tylenchines lacked *med-1*, *med-2*, *end-1*, *end-3*, and *spn-4*, amongst others. A similar pattern was observed for genes acting in spermatogenesis, early axis definition, DNA repair, and oogenesis. This core divergence

from the systems as defined in *C. elegans* makes it difficult to identify changes associated with divergence in reproductive mode and ploidy.

While the parthenogenetic *Panagrolaimus* species had an excess of genes, many duplicated, in comparison to the amphimictic ones (**Fig. 1**, Supplement), the majority of genes were not present as duplicates. This is to be expected, as the commonest fate of duplicated loci is loss through mutagenic inactivation, and there will be selection against duplicates that interact deleteriously. To test this assumption, we identified clusters of orthologues that contained exactly one protein in *C. elegans*, both amphimictic, and all 3 parthenogenetic *Panagrolaimus* species (1:1:…:1 orthologues), as well as clusters where the parthenogenetic *P.* sp. PS1159 had more proteins than the amphimictic *P.* sp. ES5. Using the database of *C. elegans* RNAi phenotypes on wormbase.org we could retrieve phenotypes for 172 of the 703 single copy orthologues and 695 phenotypes for the 1,618 clusters with a higher number of proteins from the parthenogenetic species. A chi-square test confirmed (p-value = 0.0015) that the proportion of 91 lethal phenotypes (53%) in the 1:1 orthologues was significantly different from the proportion of 275 (40%) lethal phenotypes in the second group of clusters.

Horizontal gene transfer into panagrolaim genomes

We identified potential horizontal gene transfers (HGTs) using the AlienIndex[22] (AI) approach. We identified from 22 (*P. redivivus*) to 232 (*P.* sp. ES5) likely HGT candidates from non-metazoan donors into the genomes of panagrolaimids (Table 2). We used a series of additional screens to affirm HGT and rule out contamination, including a check for spliceosomal introns, codon usage, presence of evolutionarily conserved metazoan genes on the same contigs, and transcriptional support from RNA-Seq data. The inferred source taxa for the majority of high confidence HGT candidates were Bacteria followed by Fungi in all *Panagrolaimus* species except *P. davidi* (in which Fungi were the most frequent, followed by Bacteria).

OrthoMCL clusters composed of candidate HGT proteins that were either Panagrolaimidae-specific or shared with only one close outgroup. Twenty-six of these groups contained only one *Panagrolaimus* species and thus could represent either

species-specific HGT events, or more ancient HGT that were secondarily lost. The remaining 110 OrthoMCL groups contained at least two *Panagrolaimus* species. We reconstructed the relative timing of acquisition *via* HGT for the 110 orthogroups using parsimony (Methods, **Fig. 3**). Most (102; 93%) HGT events are predicted to have taken place in an ancestor of two or more *Panagrolaimus* species. The highest number of acquisitions (49 events) took place in the lineage leading to the last common ancestor of all the cryptobiotic *Panagrolaimus* species. One HGT was predicted to have been acquired twice independently, at two nodes, but this may be a case of frequent loss. Only seven HGT orthogroups were present in the last common ancestor of all eight panagrolaimid species analysed. Eight orthogroups appeared to derive from multiple independent acquisitions at leaves of the tree.

The 910 proteins in 56 orthogroups putatively acquired via HGT in the six *Panagrolaimus* species were annotated with 192 different Pfam domains (Table 2). Only three Pfam domains were identified as conserved in the candidate HGT sets of all six *Panagrolaimus* species and all corresponded to enzymatic functions (alcohol dehydrogenase GroES-like domain, pyridine nucleotide-disulphide oxidoreductase and zinc-binding dehydrogenase) (Supplement). HGTs conserved between *P. superbus*, *P.* sp. ES5, *P. davidi* and *P.* sp. PS1159 were annotated with thirty conserved Pfam domains that included GH43, GH28 and GH32 glycoside hydrolases, which have been identified as loci acquired by HGT in other nematodes[23].

Loci under positive selection in *Panagrolaimus*

How parthenogenetic species can adapt to environmental change in the absence of effective recombination remains unresolved. We identified genes under positive selection across the different reproductive modes.

We identified loci displaying dN/dS ratios consistent with positive selection (**Fig. 4**). InterProscan and GO annotation of these loci showed significant enrichment of functions such as carbohydrate transport and metabolism, cell wall/membrane/envelope biogenesis, and replication, recombination and repair (**Fig. 4**, Supplementary Excel Tables). Loci contributing to "replication, recombination and repair" annotations include several

connected to DNA integrity and processing of nucleic acids (**Fig. 4**). An orthologue of *C. elegans denn-4*, which is expressed in neurons and also involved in reproduction, was under positive selection in parthenogenetic compared to amphimictic *Panagrolaimus*. Homologues of the human DNA polymerase zeta (REV3L) were under positive selection. These might act in DNA repair, maintaining genome integrity in the nematodes. *Panagrolaimus* orthologues of *C. elegans fcd-2*, which acts on cross-linked DNA and DNA damage repair[24], also had a dN/dS >1 between *Panagrolaimus* and the outgroups, as did helicases similar to *C. elegans* WBGene00010061. In comparison to the parthenogenetic species *fcd-2* also has a signature of selection in the sexual species. It may therefore play a role in cryptobiosis, as the parthenogenetic species are less amenable to desiccation than the sexual congeners[25]. Similarly, the *Panagrolaimus* orthologue of the *C. elegans* fatty acid and retinol-binding protein[26] *far-6* appeared to be under positive selection. Fatty acids and retinol are required for the biosynthesis of lipids integral to the nematode cell membranes and cuticle. Thus, effective lipid binding and transport may have a significant role in the evolution of cryptobiosis.

Loci implicated in cryptobiosis in *Panagrolaimus*

We found that 47 Pfam domains functionally connected to cryptobiosis were associated to proteins putatively acuired via HGT. Some of these were acquired at ancestral nodes: 19 at the last common ancestor of *Panagrolaimus* and one predicted to have been present in the tylenchine ancestor (**Fig. 4**). Five were specific to all or several parthenogenetic species. Functions associated with these domains included peptidase, diapausin, killer toxins of the Kp4 family, LURP-one-related proteins, flavin-binding monooxygenase-like protein, caleosin, and photolyase (see Supplement for extended list). Proteins with these functions fulfil various roles in defence against pathogens and the removal of xenobiotics[27-30 31]. The photolyase (**Fig. 3**) was of particular interest, since photolyases are able to repair UV induced damage in DNA[32] and have been lost in several branches of Metazoa[33], and genome repair is important for desiccating organisms[34]. The caleosin found in *Panagrolaimus* is the first report to date of animal caleosins (Supplement), since these proteins have previously been described only in multicellular plants, green algae,

and fungi[35]. Plant caleosins expressed in non-seed tissues are responsive to a range of environmental stresses, particularly salinity and dehydration[36].

Adaptation to novel habitats can also involve expansion and diversification of gene families. We therefore analysed gene family expansions in the cryptobiotic *Panagrolaimus* species compared to related taxa. Fortythree Pfam domains were overrepresented a after Benjamini-Hochberg correction in *Panagrolaimus* compared to an outgroup containing *Prop.* sp. JU765 and *Panagrellus redivivus*, and 123 Pfam domains were overrepresented in Panagrolaimidae (*Panagrolaimus*, *Propanagrolaimus*, and *Panagrellus*) compared to the remaining taxa (Supplement).

We employed a support vector machine (SVM) classifier based on Pfam domain annotations to distinguish *Panagrolaimus* from *Propanagrolaimus* plus *Panagrellus*, and, independenly Panagrolaimidae from the outgroup. Classification performance was significantly better than control classification based on a permuted class-sample mapping in the training set (real *versus* permuted classifier performance in the *Panagrolaimus versus Propanagrolaimus*/*Panagrellus* SVM: *Panagrolaimus*: T = 7, *Propanagrolaimus*/*Panagrellus*: T= 7.4, real *versus* permuted classifier performance in the Panagrolaimidae *versus* outgroup SVM: Panagrolaimidae: T = 36.9, outgroup: T= 39, all $p < 0.001$, all df = 1999). In the *Panagrolaimus versus Panagrellus*/*Propanagrolaimus* SVM analyses, 64 domains were frequently associated with successful classification, of which 13 were common to the SVM and the two-sided Fisher's exact test approaches (**Fig. 5**). These 13 domain annotations included several that may play a role in cryptobiosis, including Hsp70 protein domains, serine protease inhibitors, ubiquitin family domains, lipase domains and enoyl-reductase domains. Annotations found to be overrepresented only in the two-sided Fisher's exact test analyses were proteolysis (subtilase family and matrixin domains) and maintenance of genome integrity (PIF1-like helicase, deoxyribonuclease II, linker histone H1 and H5 family).

In the SVM comparison of Panagrolaimidae *versus* outgroup species 80 domains were or were found to contribute to the differentiation between the taxa. Of these 24 were common with those identified in the two-sided Fisher's exact test analyses (**Fig. 5**),

including BTB/POZ domain, BTB and C-terminal Kelch, and ABC transporter. These three domains had previously been shown to be expanded in *P. redivivus* compared to *C. elegans*[14]. While C- type lectins were overrepresented in Panagrolaimidae as a whole, in the two-species comparison *C. elegans* had more C-type lectins than did *P. redivivus*[14]. C-type lectins are expressed during dehydration in the tylenchine nematode *Aphelenchus avenae*[6] and could thus also play a role in the dehydration process in *Panagrolaimus*. Other domains with potential roles in cryptobiosis were also identified as amplified in Panagrolaimidae, including heat shock proteins (HSPs), such as HSP70 and the collagen-specific serpin HSP47, the late embryogenesis abundant (LEA) domain (identified only in the NHST analysis), and antioxidant and reactive oxygen species (ROS) regulating domains such as cytochrome P450. The abundance of the DEAD/DEAH box domain, involved in RNA metabolism, may support roles in transcript and ribosome stability[37]. Small heat shock proteins (sHSPs) carry out chaperone functions in anhydriobiotic animals[38]. We identified a tylenchine-specific family of sHSPs that was distinct from the lineage specific expansion in *C. elegans*[39], but no subset specific to panagrolaims (Supplement).

Discussion

Understanding how parthenogenetic animals evolve and how (long) they persist through evolutionary time is of prime interest to the fundamental biological question of why sex is the ubiquitous mode of reproduction in animals. A similarly important question is how animal species adapt to unfavourable environmental conditions. It has been proposed that both traits are linked in geographical parthenogenesis.

Many parthenogenetic species have been shown to be of hybrid origin and polyploid[40,41], including tylenchine nematodes[41-43]. In animals it is assumed that polyploidy, especially when uneven, in general leads to non-viability, and that living hybrids may suffer a reduced fitness[44]. Altenatively, polyploidy might lead to a fitness advantage in buffering against Muller's ratchet[5]. Hybridisation leads to aberrant meiosis in animals, and a possible escape from this dilemma could be parthenogenesis[45]. In such systems, incorporation of a round of chromosomal endoreplication is a possible route to regain triploidy during oogenesis[46].

We have analysed the genomes and transcriptomes of parthenogenetic and amphimictic *Panagrolaimus* nematodes, which in contrast to the outgroup *Propanagrolaimus* and *Panagrellus* are both capable of surviving under prolonged anhydro- and cryobiosis. Our combined karyotypic and molecular data both indicate a triploid hybrid origin in parthenogenetic *Panagrolaimus*, with all analysed species showing n=12 (compared to diploid n=8 in amphimicts), an excess of duplicate genes that share a single phylogenetic origin, and variant frequency spectra compatible with triploidy. Triploidy could have arisen by autopolyploidy (duplication of one set of ancestral chromosomes, for example by fertilisation of a diploid oocyte by a haploid sperm from the same species) or allopolyploidy (where the additional haploid complement comes from a distinct lineage). These contrasting origins in turn predict contrasting patterns of divergence between the three copies of each locus in a triploid: the autopolyploid will have three closely related loci, while the allopolyploid will have a pair of closely related (allelic) copies and a more distantly related (homoeolog) copy. In parthenogenetic *Panagrolaimus* we find phylogenetic evidence for divergence of extra copies and probable acquisition from a donor species positioned between the *Propanagrolaimus* and *Panagrolaimus* species present in our analysis. These results support allotriploidisation.

Analysis of different taxa that can undergo anhydrobiosis has identified several different kinds of genes associated with desiccation survival. In plant embryos, LEA proteins, natively unstructured peptides that are believed to surround proteins like a "molecular shield" [47], are highly expressed. Natively unstructured proteins have been identified in anhydrobiotic responses in nematodes (*Aphelenchus avenae*[6]), arthropods (*Polypedilum vanderplanki*[48]) and tardigrades (*R. varieornatus*[44] and *H. dujardini*[49]). Other gene families associated with anhydrobiosis include protein kinases, proteasomal components, ubiqutin, protease inhibitors, proteases, DNA repair enzymes, enzymes with roles in oxidative stress protection, and heat shock proteins[48]. We identified overrepresentation of the above-mentioned gene families, as well as a number of others not previously identified, associated with cryptobiotic *Panagrolaimus* species. The set of *Panagrolaimus*-enhanced activities overlapped with the list of genes we identified as having been acquired through HGT. Many putative horizontal gene transfers were mapped to the root of the cryptobiotic *Panagrolaimus* species group. These candidate

HGT genes have acquired spliceosomal introns, have similar GC content and codon usage to resident *Panagrolaimus* genes, are assembled in the vicinity of *bona fide* nematode genes and are expressed as poly(A)+ RNA. They can be regarded as firmly domesticated and functionally integrated into the host genomes. Functional annotation of these putative HGT loci suggested that they contribute to anhydrobiotic physiology. Genes acquired through HGT were also implicated in cryptobiosis in the rotifer *A. vaga* and the tardigrade *R. varieornatus*, and a functional link between anhydrobiosis and HGT has been proposed[50].

The theory underpinning geographical parthenogenesis holds that parthenogenetic species have an advantage in adapting to new and/or extreme environments. Since we did not observe a signal for specific adaptations in the parthenogenetic species (including to the mode of reproduction itself) in comparison to sexual congeners it appears possible that a general-purpose genotype is indeed maintained in evolutionary young parthenogens, potentially by combining several differently-adapted genomes during hybridisation. Like most parthenogens, we estimate that the *Panagrolaimus* species are younger than the ancient evolutionary "scandals"[51]. Few parthenogenetic species are suitable for genomic and molecular research. *Panagrolaimus* nematodes are easily cultured in the laboratory and accessible to molecular genetic analysis[52], including gene knock-out techniques[6]. *Panagrolaimus* species may also be useful in understanding developmental diversity in Nematoda, diversification and speciation in parthenogens, the effects of triploidy on gene and genome evolution, and the biology of cryptobiosis. The existence of a well-studied model species (*C. elegans*) as a comparator should make the *Panagrolaimus* model even more attractive, as extensive functional annotation can be carried over. However, we found that many genes acting in *C. elegans* reproduction and development were neither present in the *Panagrolaimus* species, nor more broadly in Tylenchina. This may limit direct transfer of knowledge from the model, but *C. elegans* nevertheless constitutes a key reference for biological understanding.

Online Methods

Acknowledgements and Funding

The authors thank Isabel Goebel for generation time measurements, Janine Beyert and Andrea Kraemer-Eis for help in culturing nematodes. The authors are grateful to the GenoToul bioinformatics platform Toulouse Midi-Pyrenees for providing computing resources, as well as to Nordrhein-Westfalen for the use of the CHEOPS computing cluster at the University of Cologne.

PHS was funded by the VolkswagenStiftung in their initiative for evolutionary biology, the DFG in a grant to TW under SPP1819, and is supported by the European Research Council (ERC-2012-AdG 322790). AMB funded by Science Foundation Ireland (Project RFP/EOB2506). EAMS thanks Dr. Ryan Gregory (University of Guelph) for providing access to his laboratory to conduct genome size estimations. She also thanks CONACyT for her Postdoctoral Fellowship (2010), and UAZ-2016-37096. CJC was funded under the Science Foundation Ireland (SFI) Stokes lecturer scheme (07/SK/B1236A) and the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Grant, Rumen Systems Biology, (BB/E/W/10964A01).

Authors contributions

PHS: planned and conceived the study, managed the collaboration, analysed the data, supervised GS, wrote the paper

EGJD: analysed the HGT data, and planned and wrote the ms

AMB: analysed the anhydrobiosis and HGT data, conceived the study, supervised GO'M & BC, wrote the paper

A-MS: programmed the classifier, drew figures and contributed to writing

CJC: analysed the selection data, drew figures, and contributed to writing

SW: annotated the genomic data

ID: karyotyped species, drew figures, contributed to writing

GO'M: analysed anhydrobiosis data for *P. superbus*

BAC: conducted molecular experiments for transcriptomics

CR: analysed the HGT data

GS: conducted genome size measurements

EAMS: conducted genome size measurements, wrote and revised the manuscript

UT: analysed raw sequencing data

MK: conceived parts of the study, co-supervised PHS

MAST: conceived parts of the study, co-initiated the collaboration, analysed the anhydrobiosis data

ES: conducted microscopic analyses, supervised PHS, and revised the manuscript

TW: analysed generation time data, wrote and revised the ms

MB: planned and conceived the study, analysed the data, and wrote the ms

References

1. Lynch, M., Bürger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J Hered* **84,** 339–344 (1993).
2. Felsenstein, J. The Evolutionary Advantage of Recombination. *Genetics* **78,** 737–757 (1974).
3. Debortoli, N. *et al.* Genetic Exchange among Bdelloid Rotifers Is More Likely Due to Horizontal Gene Transfer Than to Meiotic Sex. *Curr Biol* **26,** 1–11 (2016).
4. Archetti, M. Recombination and loss of complementation: a more than two-fold cost for parthenogenesis. *J Evolution Biol* **17,** 1084–1097 (2004).
5. Lynch, M. Destabilizing Hybridisation, General-Purpose Genotypes and Geographic Parthenogenesis. *The Quarterly Review of Biology* 35 (1984).
6. Reardon, W. *et al.* Expression profiling and cross-species RNA interference (RNAi) of desiccation-induced transcripts in the anhydrobiotic nematode Aphelenchus avenae. *BMC Mol Biol* **11,** 6 (2010).
7. Gusev, O. *et al.* Comparative genome sequencing reveals genomic signature of extreme desiccation tolerance in the anhydrobiotic midge. *Nature Communications* **5,** 4784 (2014).
8. Hashimoto, T. *et al.* Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. *Nature Communications* **7,** 12808 (2016).
9. Tyson, T. *et al.* A molecular analysis of desiccation tolerance mechanisms in the anhydrobiotic nematode Panagrolaimus superbus using expressed sequenced tags. *BMC Research Notes* **5,** 68 (2012).
10. Pouchkina-Stantcheva, N. N. *et al.* Functional divergence of former alleles in an ancient asexual invertebrate. *Science* **318,** 268–271 (2007).
11. Shannon, A., Browne, J., Boyd, J., Fitzpatrick, D. & Burnell, A. The anhydrobiotic potential and molecular phylogenetics of species and strains of Panagrolaimus (Nematoda, Panagrolaimidae). *J Exp Biol* **208,** 2433–2445 (2005).
12. Lewis, S. *et al.* Molecular evolution in Panagrolaimus nematodes: origins of parthenogenesis, hermaphroditism and the Antarctic species P. davidi. *BMC Evol Biol* **9,** 15 (2009).
13. Thorne, M. A. S., Kagoshima, H., Clark, M. S., Marshall, C. J. & Wharton, D. A. Molecular analysis of the cold tolerant Antarctic nematode, Panagrolaimus davidi. *PLoS ONE* **9,** e104526 (2014).
14. Srinivasan, J. *et al.* The draft genome and transcriptome of Panagrellus redivivus are shaped by the harsh demands of a free-living lifestyle. *Genetics* **193,** 1279–1295 (2013).
15. Opperman, C. H. *et al.* Sequence and genetic map of Meloidogyne hapla: A compact nematode genome for plant parasitism. *Proceedings of the National Academy of Sciences* **105,** 14802–14807 (2008).
16. Kikuchi, T. *et al.* Genomic Insights into the Origin of Parasitism in the Emerging Plant Pathogen Bursaphelenchus xylophilus. *PLoS Pathog* **7,** (2011).
17. Thomas, G. W. C., Ather, S. H. & Hahn, M. W. *Gene-tree reconciliation with MUL-trees to resolve polyploidy events.* 1–21 (2016). doi:10.1101/058149
18. Lefort, V., Desper, R. & Gascuel, O. FastME 2.0: A Comprehensive, Accurate, and

Fast Distance-Based Phylogeny Inference Program. *Mol Biol Evol* **32,** 2798–2800 (2015).
19. Cutter, A. D. Divergence Times in Caenorhabditis and Drosophila Inferred from Direct Estimates of the Neutral Mutation Rate. *Mol Biol Evol* **25,** 778–786 (2008).
20. Li, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13,** 2178–2189 (2003).
21. Linard, B. *et al.* OrthoInspector 2.0: software and database updates. *Bioinformatics* **31,** btu642–448 (2014).
22. Rancurel, C., Da Rocha, M. & Danchin, E. G. J. Alienness: rapid detection of horizontal gene transfers in metazoan genomes. *European Conference on Computational Biology* 1–1 (2014). at
23. Haegeman, A., Jones, J. T. & Danchin, E. G. J. Horizontal gene transfer in nematodes: a catalyst for plant parasitism? *Mol. Plant Microbe Interact.* **24,** 879–887 (2011).
24. Vermeulen, W. Tissue specific response to DNA damage: C. elegans as role model. *DNA Repair* **32,** 141–148 (2015).
25. McGill, L. M. *et al.* Anhydrobiosis and Freezing-Tolerance: Adaptations That Facilitate the Establishment of Panagrolaimus Nematodes in Polar Habitats. *PLoS ONE* **10,** e0116084–24 (2015).
26. Garofalo, A. *et al.* The FAR protein family of the nematode Caenorhabditis elegans. *Journal of Biological Chemistry* **278,** 8065–8074 (2003).
27. Khalil, H. B. *et al.* Characterization of the caleosin gene family in the Triticeae. *BMC Genomics* **15,** 239 (2014).
28. Tanaka, H. *et al.* Insect diapause-specific peptide from the leaf beetle has consensus with a putative iridovirus peptide. *Peptides* **24,** 1327–1333 (2003).
29. Gu, F. *et al.* Structure and function of a virally encoded fungal toxin from Ustilago maydis: a fungal and mammalian Ca2+ channel inhibitor. *Structure* **3,** 805–814 (1995).
30. Kuwabara, C., Arakawa, K. & Yoshida, S. Abscisic acid-induced secretory proteins in suspension-cultured cells of winter wheat. *Plant Cell Physiol.* **40,** 184–191 (1999).
31. Eswaramoorthy, S., Bonanno, J. B., Burley, S. K. & Swaminathan, S. Mechanism of action of a flavin-containing monooxygenase. *P Natl Acad Sci Usa* **103,** 9832–9837 (2006).
32. Sancar, G. B. DNA photolyases: physical properties, action mechanism, and roles in dark repair. *Mutat Res* **236,** 147–160 (1990).
33. Lucas-Lledo, J. I. & Lynch, M. Evolution of Mutation Rates: Phylogenomic Analysis of the Photolyase/Cryptochrome Family. *Mol Biol Evol* **26,** 1143–1153 (2009).
34. Jönsson, K. I. & Schill, R. O. Induction of Hsp70 by desiccation, ionising radiation and heat-shock in the eutardigrade Richtersius coronifer. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* **146,** 456–460 (2007).
35. Partridge, M. & Murphy, D. J. Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in Arabidopsis. *Plant Physiol. Biochem.* **47,** 796–806 (2009).
36. Blée, E. *et al.* The reductase activity of the Arabidopsis caleosin RESPONSIVE TO DESSICATION20 mediates gibberellin-dependent flowering time, abscisic acid

sensitivity, and tolerance to oxidative stress. *Plant Physiol.* **166,** 109–124 (2014).
37. Guan, Q. *et al.* A DEAD box RNA helicase is critical for pre-mRNA splicing, cold-responsive gene regulation, and cold tolerance in Arabidopsis. *Plant Cell* **25,** 342–356 (2013).
38. Gusev, O., Cornette, R., Kikawada, T. & Okuda, T. Expression of heat shock protein-coding genes associated with anhydrobiosis in an African chironomid Polypedilum vanderplanki. *Cell Stress and Chaperones* **16,** 81–90 (2011).
39. Aevermann, B. D. & Waters, E. R. A comparative genomic analysis of the small heat shock proteins in Caenorhabditis elegans and briggsae. *Genetica* **133,** 307–319 (2008).
40. Castagnone-Sereno, P. & Danchin, E. G. J. Parasitic success without sex - the nematode experience. *J Evolution Biol* **27,** 1323–1333 (2014).
41. Lunt, D. H., Kumar, S., Koutsovoulos, G., Blaxter, M. L. & Blaxter, M. L. The complex hybrid origins of the root knot nematodes revealed through comparative genomics. *PeerJ* **2,** e356 (2014).
42. Lunt, D. H. Genetic tests of ancient asexuality in Root Knot Nematodes reveal recent hybrid origins. *BMC Evol Biol* **8,** 194–16 (2008).
43. Blanc-Mathieu, R. *et al.* Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLoS Genet* **13,** e1006777 (2017).
44. Burke, J. M. & Arnold, M. L. Genetics and the fitness of hybrids. *Annual Review of Genetics* **35,** 31–52 (2001).
45. Avise, J. *Clonality : The Genetics, Ecology, and Evolution of Sexual Abstinence in Vertebrate Animals*. (Oxford University Press, 2008).
46. Lutes, A. A., Neaves, W. B., Baumann, D. P., Wiegraebe, W. & Baumann, P. Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. *Nature* **464,** 283–286 (2010).
47. Chakrabortee, S. *et al.* Intrinsically disordered proteins as molecular shields. *Molecular BioSystems* **8,** 210–219 (2012).
48. Cornette, R. *et al.* Identification of anhydrobiosis-related genes from an expressed sequence tag database in the cryptobiotic midge Polypedilum vanderplanki (Diptera; Chironomidae). *J. Biol. Chem.* **285,** 35889–35899 (2010).
49. Yoshida, Y. *et al.* Comparative genomics of the tardigrades Hypsibius dujardini and Ramazzottius varieornatus. (2017). doi:10.1101/112664
50. Hespeels, B. *et al.* Gateway to genetic exchange? DNA double-strand breaks in the bdelloid rotifer Adineta vagasubmitted to desiccation. *J Evolution Biol* **27,** 1334–1345 (2014).
51. Smith, J. M. *The Evolution of Sex*. (CUP Archive, 1978).
52. Schiffer, P. H. *et al.* Developmental variations among Panagrolaimid nematodes indicate developmental system drift within a small taxonomic unit. *Development Genes and Evolution* **224,** 183–188 (2014).

Table 1: Pairwise divergence time estimates in *Panagrolaimus* spp. based on 8 days per generation (measured by us) and 50 days per generation as lower bound (see Supplement). Calibrated with divergence times calculated for *Caenorhabditis* spp. by Cutter (2008).

Species A	repro	Species B	repro	Myr (8 days/gen)	Myr (50 days/gen)
P. superbus	G	ES5	G	1	3.2
PS1159	P	ES5	G	1.33	8.25
PS1159	P	*P. superbus*	G	1.37	8.55
ES5	G	*P. redivivus*	G	1.35	8.35
PS1159	P	*P. redivivus*	G	1.45	9.3
P. superbus	G	*P. redivivus*	G	1.88	11.6
ES5	G	JU765	H	2.4	14.95
PS1159	P	JU765	H	2.4	15.1
P. superbus	G	JU765	H	2.5	15.8

Table 2: Alien indices and Pfam domains assigned to HGT candidates.

Species	*Total # proteins*	*possible HGT candidates (AI >0)*	*Candidates identified as contaminants**	*# high-confidence HGT candidates with AI >30*	*High confidence HGTs*					
					# on same contig with other gene	*# with neighbours hitting Metazoa*	*# with introns*	*# with expression >10/100/1000 TPM in RNASeq*	*# with Pfam domains* (Total number of Pfam domains)	*Distinct Pfam domains*
Panagrolaimus superbus	21,381	528	20	88	n/a	n/a	n/a	n/a	70 (91)	56
Panagrolaimus sp. ES5	24,756	975	141	232	215	73	132	185/155/97	183 (429)	136
Panagrolaimus sp. PS1159	27,350	899	19	182	159	98	133	174/164/112	141 (268)	93
Panagrolaimus davidi	31,646	1,177	18	201	194	84	136	178/153/84	131(172)	62
Propanagrolaimus sp. JU765	24,878	530	39	44	42	25	27	44/44/36	34 (100)	23
Panagrellus redivivus	26,372	516	12	22	21	21	16	n/a	14 (18)	12

*AI>0 and >70% id to non-metazoan protein

Figure 1. MUL-Tree cladogram, karyotypes, and genome assemblies

- **A.** A MUL-trees based gene-tree reconciliation based on 15,000 trees from Orthofinder found the most parsimonious phylogenetic position for proteins where the parthenogenetic species had multiple copies in a given orthologous group to be outside of a clade containing our gonochoristic and parthenogenetic *Panagrolaimus*. This topology supports the assumption of polyploid hybrids.
- **B.** All analysed parthenogenetic species had 12 chromosomes in comparison to 8 in the diploid state of the amphimictic species. The out group *Propanagrolaimus* has 2n=10.
- **C.** Assembly and gene prediction metrics. P. sp. PS1159 and *P. davidi* both have more predictions than the sexual species. The low coverage PS1579 genome appears to lack regions of the genome, and we thus included transcriptome based gene predictions(*) directly into downstream analysis.
- **D.** The presence of polar bodies in the parthenogenetic panagrolaims confirms the presence of functional meiosis.

Figure 2. Polyploidy and hybridization analysis, and species age

A. Mapping RNA-Seq reads to coding sequences of predicted in the respective genomes and counting the occurrence of variant frequencies, we found that in the parthenogenetic species many variant positions are supported by 1/3 of the mapped reads at the position. Variant frequencies in the homozygous data from the gonochoristic ES5 are not showing a distinct pattern, while the hermaphrodite JU765 shows the 50% peak expected under heterozygosity.

B. Using our own generation time measures, genome wide divergence calculated with Andi, and previous data for *Caenorhabditis* species as reference we calculated possible age of the origin of parthenogenesis in *Panagrolaimus*. 8 days per generation was measured in the laboratory and would indicate an approximate age of 1.3Myr for the PS1159/*P. superbus* split (see table 2).

Figure 3. Reconsruction of the timing of acquisition and loss of HGT candidates

- **A.** An alignment of the photolyase found in *Panagrolaimus*, and *A. suum* and *A. nanus* shows that distinct differences beteween the proteins. The *A. suum* and *A. nanus* domain annotations point towards *E. coli*, rendering it likely that the proteins are contaminations from the bacterial food source used commonly used in nematode cultures.
- **B.** A phylogenetic tree of the photolyase proteins in Panagrolaimus, A*. suum* and *A. nanus*, and bacterial photolyases identified with blast confirms that the *A. suum* and *A. nanus* proteins are *E. coli* contamination, while the *Panagrolaimus* proteins have evolved since their horizontal acquisition.
- **C.** Demonstrating the integration into the *Panagrolaimus* genomes the photolyase was found on one contig with eukaryotic (nematode) proteins in the *P. sp.* PS1159 assembly.
- **D.** Cladogram of the Panagrolaimidae species and two outgroup relatives. The distribution of HGT candidates across these species has been mapped to the phylogeny and ancestral states have been deduced with a parsimony approach. Internal nodes are numbered in black (2,3,4,5,6,7,8,9,14). The number of gene families acquired via HGT / lost at each node are

represented in green / red bubbles, respectively. The total number of gene families acquired via HGT present at each node is indicated within blue boxes.

E. Stacked bars of classes within the AI>30 HGT candidates for the different species. Candidates with introns and expression levels of TPM>100 can be regarded as most reliable.

Figure 4. Adaptive evolution analysis.

Average rates of evolution (dN/dS) are shown for comparisons between parthenogenic and gonochoristic *Panagrolaimus* (in blue), *Panagrolaimus* and the closest outgroups (*Propanagrolaimus* and *Panagrellus*) (in green), and between all species (in yellow) grouped by functional categories (see supplement). We find the while the majority of genes have a ratio of < 1 there are notable exceptions across several functional categories both between parthenogenetic and gonochoristic species and between panagrolaimids and their closest relatives.

Figure 5. SVM classification and overrepresentation analysis

Using a Fisher's exact test and a Support Vector Machine classification we compared Pfam annotations in the genus *Pangrolaimus* with thos of *Propanagrolaimus* and *Panagrellus* (combined), as well as these species taken together as Panagrolaimidae with outgroup species. This comparison revealed several protein families associated with cryptobyosis, e.g HSPs, helicases, and C-type lectins, to be important for *Panagrolaimus* biology.

Figure 1

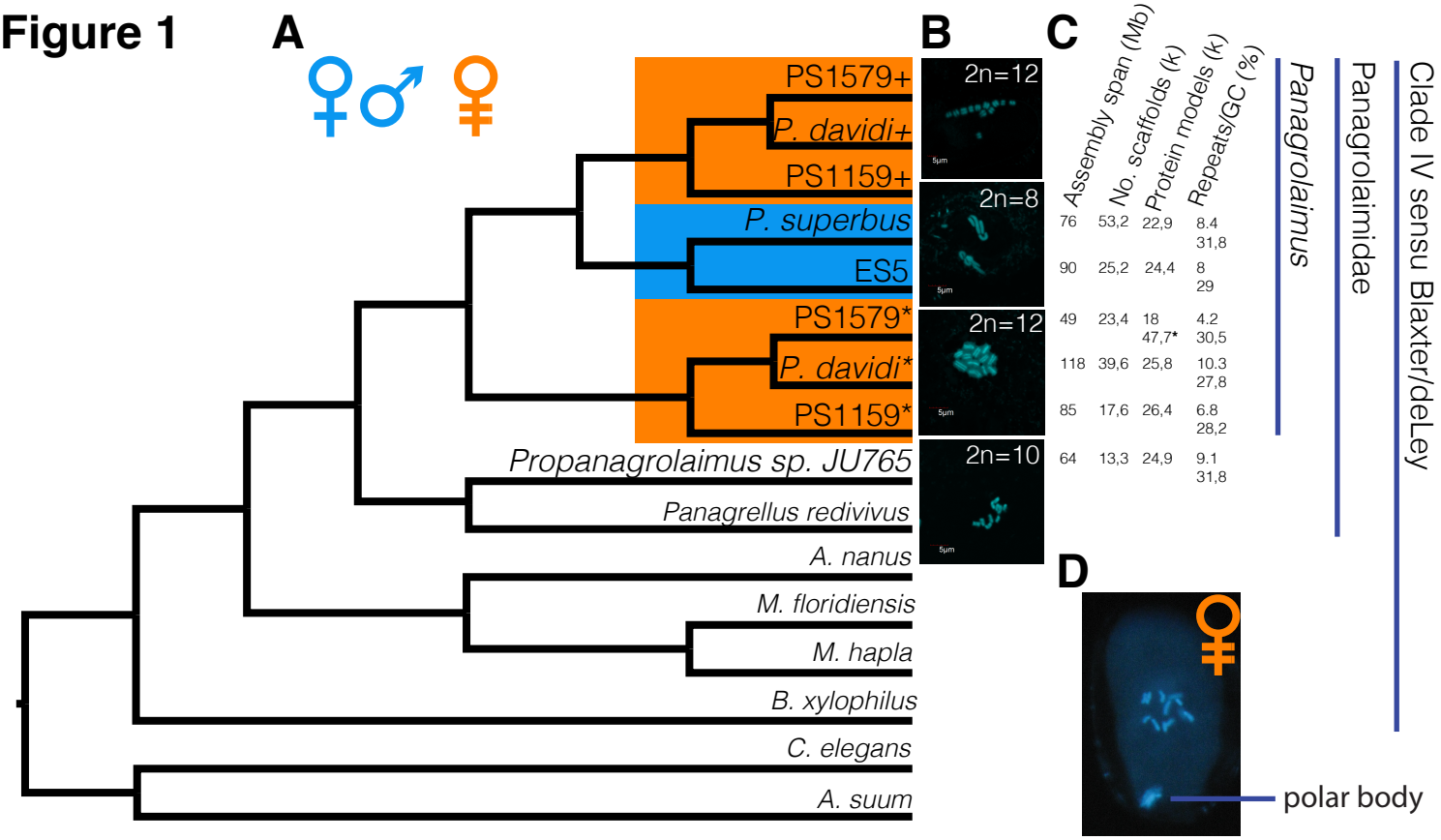
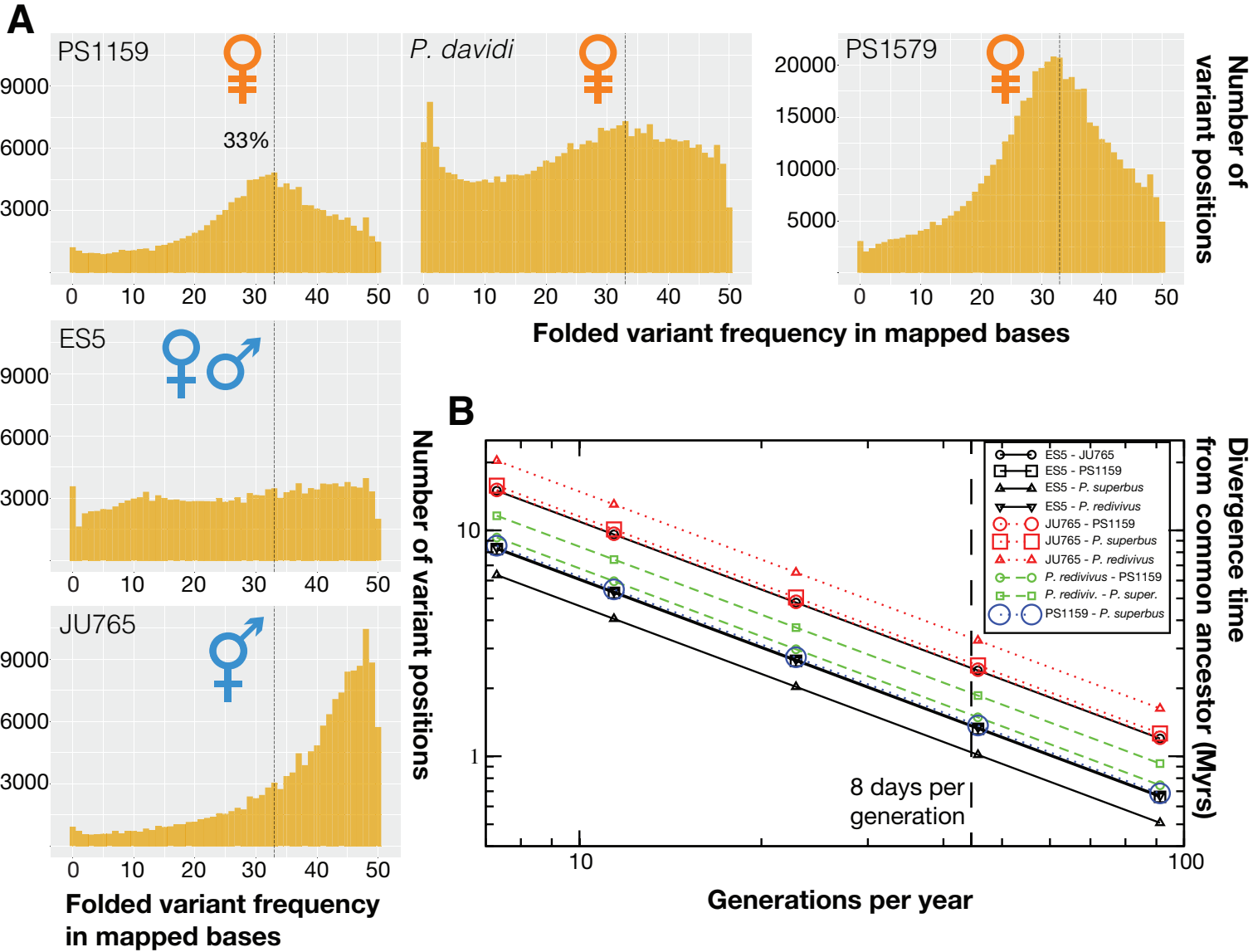

A
B
C
D
PS1579+
P. davidi+
PS1159+
P. superbus
ES5
PS1579*
P. davidi*
PS1159*
Propanagrolaimus sp. JU765
Panagrellus redivivus
A. nanus
M. floridiensis
M. hapla
B. xylophilus
C. elegans
A. suum
2n=12
2n=8
2n=12
2n=10
Assembly span (Mb)
No. scaffolds (k)
Protein models (k)
Repeats/GC (%)
76 53,2 22,9 8.4 31,8
90 25,2 24,4 8 29
49 23,4 18 47,7* 4.2 30,5
118 39,6 25,8 10.3 27,8
85 17,6 26,4 6.8 28,2
64 13,3 24,9 9.1 31,8
Panagrolaimus
Panagrolaimidae
Clade IV sensu Blaxter/deLey
polar body

Figure 2


A
PS1159
33%
P. davidi
PS1579
ES5
JU765
Number of variant positions
Folded variant frequency in mapped bases
B
ES5 - JU765
ES5 - PS1159
ES5 - P. superbus
ES5 - P. redivivus
JU765 - PS1159
JU765 - P. superbus
JU765 - P. redivivus
P. redivivus - PS1159
P. rediviv. - P. super.
PS1159 - P. superbus
8 days per generation
Generations per year
Divergence time from common ancestor (Myrs)

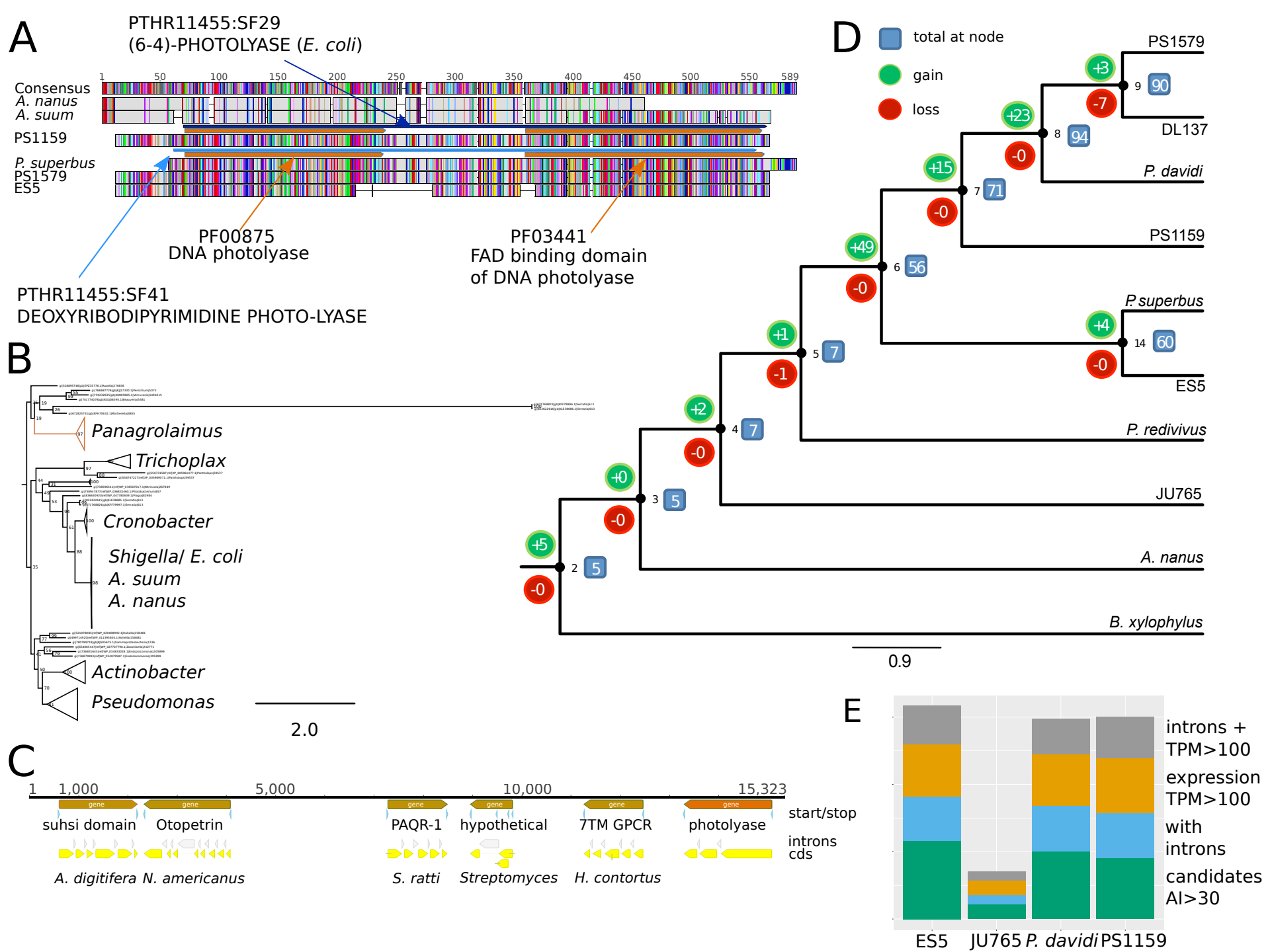

A
PTHR11455:SF29
(6-4)-PHOTOLYASE (E. coli)
Consensus
A. nanus
A. suum
PS1159
P. superbus
PS1579
ES5
PF00875
DNA photolyase
PF03441
FAD binding domain
of DNA photolyase
PTHR11455:SF41
DEOXYRIBODIPYRIMIDINE PHOTO-LYASE
B
Panagrolaimus
Trichoplax
Cronobacter
Shigella/ E. coli
A. suum
A. nanus
Actinobacter
Pseudomonas
2.0
C
1
1,000
5,000
10,000
15,323
gene
suhsi domain
Otopetrin
PAQR-1
hypothetical
7TM GPCR
photolyase
start/stop
introns
cds
A. digitifera
N. americanus
S. ratti
Streptomyces
H. contortus
D
total at node
gain
loss
PS1579
DL137
P. davidi
PS1159
P. superbus
ES5
P. redivivus
JU765
A. nanus
B. xylophylus
0.9
E
introns +
TPM>100
expression
TPM>100
with
introns
candidates
AI>30
ES5
JU765
P. davidi
PS1159

Figure 4

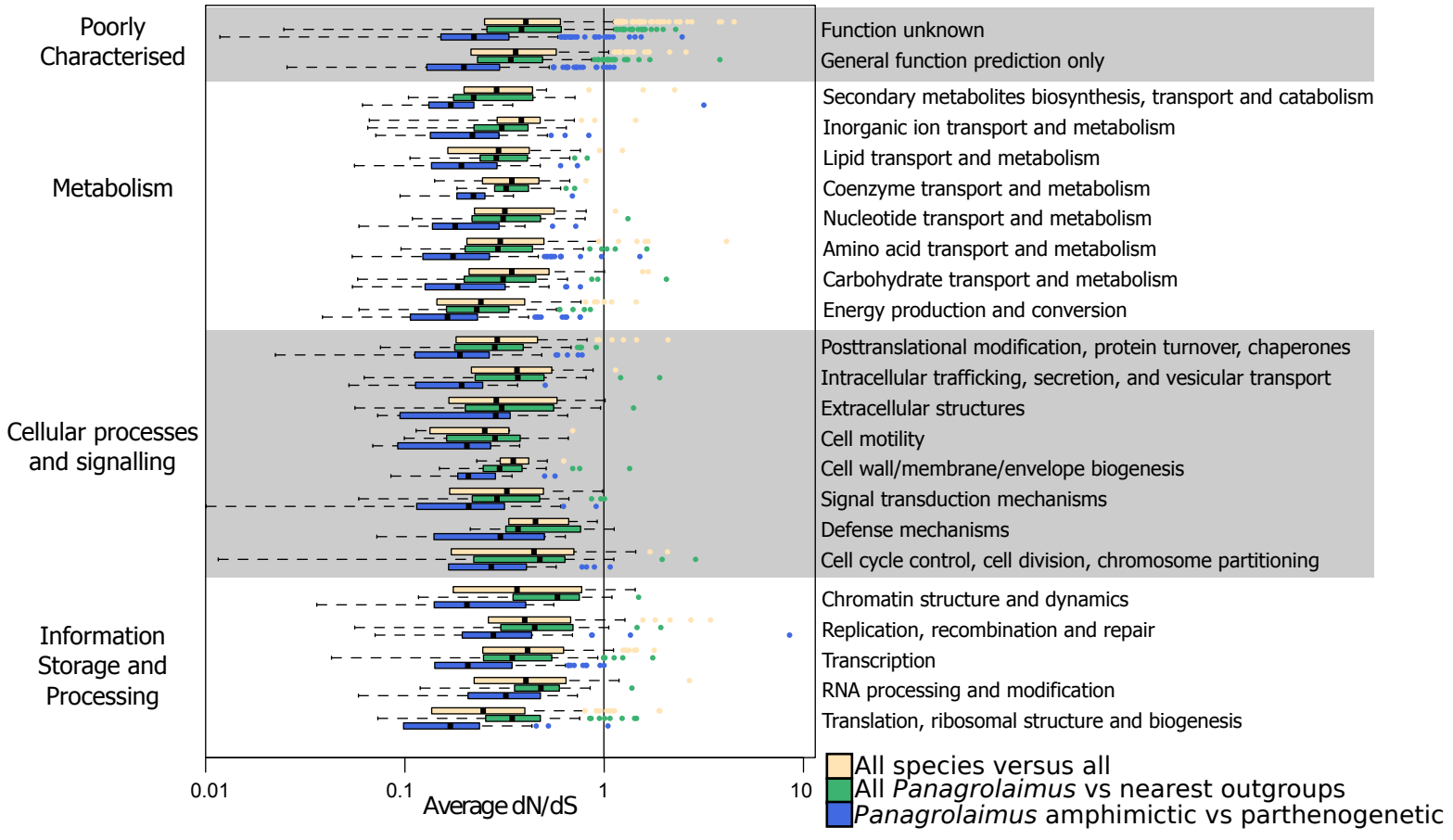
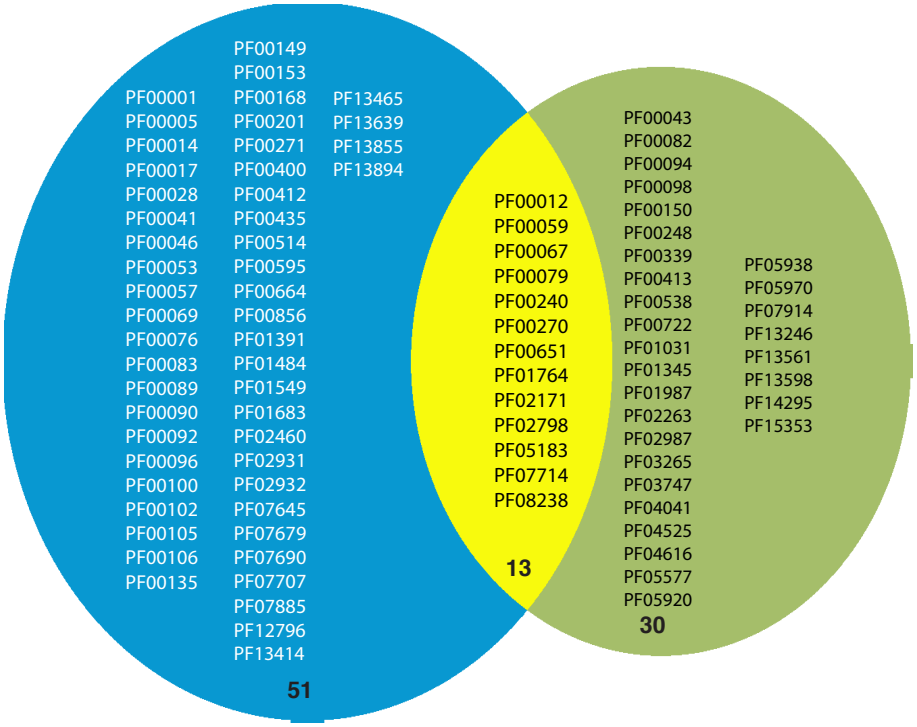
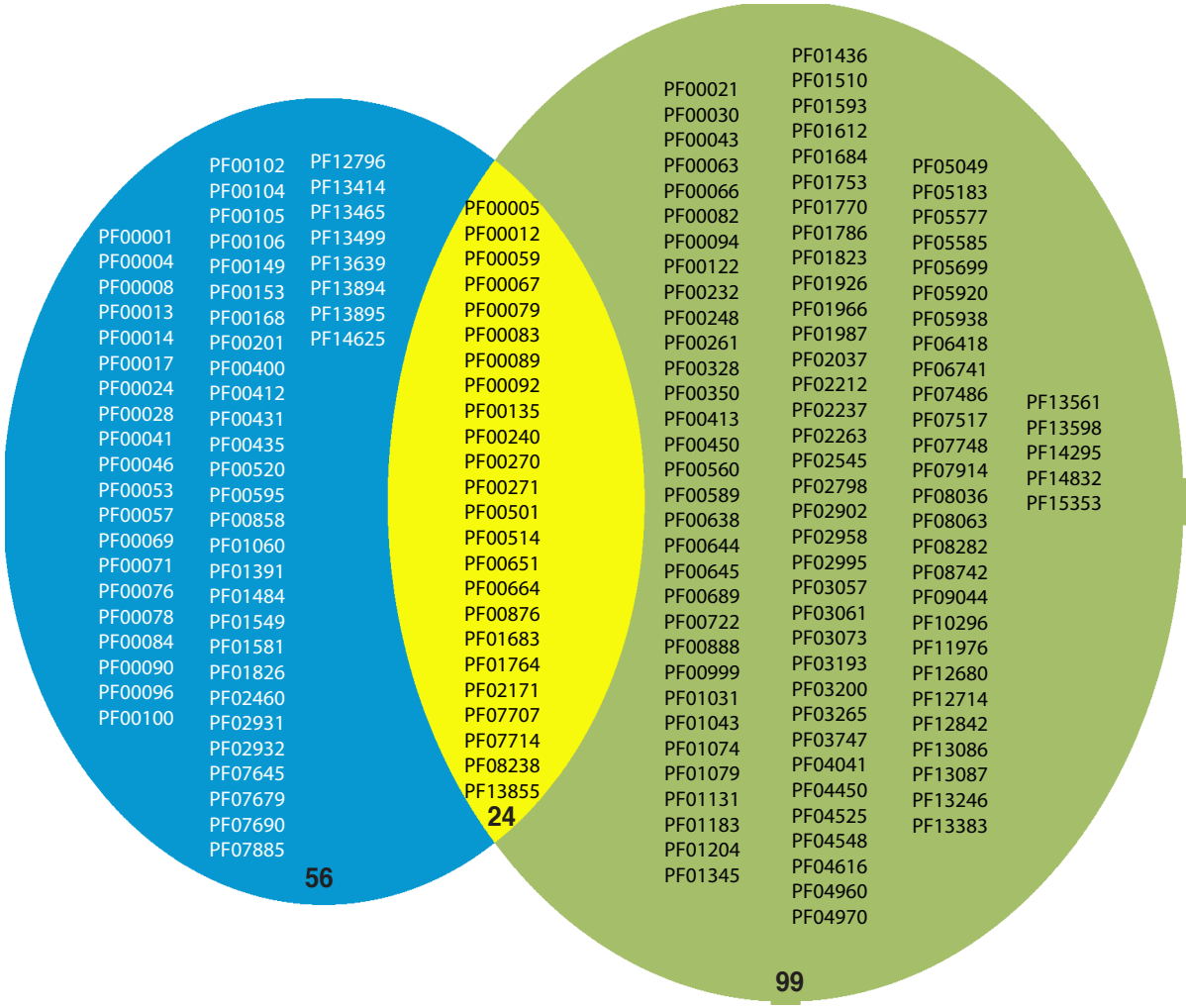

Poorly Characterised
Metabolism
Cellular processes and signalling
Information Storage and Processing
Function unknown
General function prediction only
Secondary metabolites biosynthesis, transport and catabolism
Inorganic ion transport and metabolism
Lipid transport and metabolism
Coenzyme transport and metabolism
Nucleotide transport and metabolism
Amino acid transport and metabolism
Carbohydrate transport and metabolism
Energy production and conversion
Posttranslational modification, protein turnover, chaperones
Intracellular trafficking, secretion, and vesicular transport
Extracellular structures
Cell motility
Cell wall/membrane/envelope biogenesis
Signal transduction mechanisms
Defense mechanisms
Cell cycle control, cell division, chromosome partitioning
Chromatin structure and dynamics
Replication, recombination and repair
Transcription
RNA processing and modification
Translation, ribosomal structure and biogenesis
0.01
0.1
1
10
Average dN/dS
All species versus all
All Panagrolaimus vs nearest outgroups
Panagrolaimus amphimictic vs parthenogenetic

Figure 5





Support Vector Machine (SVM)

Fisher's exact test (NHST)

Panagrolaimus vs. *Propanagrolaimus* & *Panagrellus*

Panagrolaimidae vs. Outgroup