Desiccation does not drastically increase the accessibility of exogenous DNA to nuclear genomes: Evidence from the frequency of endosymbiotic DNA transfer

Xi-Xi Li
Beijing Normal University

Cheng Fang
Beijing Normal University

Jun-Peng Zhao
Beijing Normal University

Xiao-Yu Zhou
Beijing Normal University

Zhihua Ni
Beijing Normal University

Deng-Ke Niu (✉ dkniu@bnu.edu.cn)
Beijing Normal University   https://orcid.org/0000-0003-4503-2658

Abstract

Background: Although horizontal gene transfer (HGT) is a widely accepted force in the evolution of prokaryotic genomes, its role in the evolution of eukaryotic genomes remains hotly debated. Some bdelloid rotifers that are resistant to extreme desiccation and radiation undergo a very high level of HGT, whereas in another desiccation-resistant invertebrate, the tardigrade, the pattern does not exist. Overall, the DNA double-strand breaks (DSBs) induced by prolonged desiccation have been postulated to open a gateway to the nuclear genome for exogenous DNA integration and thus to facilitate the HGT process, thereby enhancing the rate of endosymbiotic DNA transfer (EDT).

Results: We first surveyed the abundance of nuclear mitochondrial DNAs (NUMTs) and nuclear plastid DNAs (NUPTs) in five eukaryotes that are highly resistant to desiccation: the bdelloid rotifers *Adineta vaga* and *Adineta ricciae*, the tardigrade *Ramazzottius varieornatus*, and the resurrection plants *Dorcoceras hygrometricum* and *Selaginella tamariscina*. Excessive NUMTs or NUPTs were not detected. Furthermore, we compared 24 groups of desiccation-tolerant organisms with their relatively less desiccation-tolerant relatives but did not find a significant difference in NUMT/NUPT contents.

Conclusions: Desiccation may induce DSBs, but it is unlikely to dramatically increase the frequency of exogenous sequence integration in most eukaryotes. The capture of exogenous DNA sequences is possible only when DSBs are repaired through a subtype of non-homologous end joining, named alternative end joining (alt-EJ). Due to the deleterious effects of the resulting insertion mutations, alt-EJ is less frequently initiated than other mechanisms.

Background

Horizontal gene transfer (HGT, also termed lateral gene transfer) is the movement of genetic materials between organisms rather than from parent to offspring [1, 2]. HGT is well documented as the principal route of evolutionary innovation in bacteria and archaea, such as the acquisition of antibiotic resistance in pathogenic bacteria [3]. In contrast, the contribution of HGT to the evolution of eukaryotic genomes remains hotly debated [4-10]. An example of this controversy is the amount of human genes acquired via HGT [11]. In the first draft of the human genome, 223 proteins were found to be more similar to bacterial proteins than to any eukaryotic protein sequences available at that time, and these human proteins were suggested to have arisen from HGTs [12]. However, this result was quickly refuted by extensive analyses of the candidate genes [13, 14], and according to the opponents, gene loss in other eukaryotic lineages was the most likely explanation. Fourteen years later, Crisp et al [15] confirmed 17 previously reported foreign genes and found 128 additional foreign genes in the human genome in a large-scale comparative study; however, their conclusion was refuted by a case-by-case re-analysis of the 145 genes [16]. Almost at the same time, another group claimed that they had discovered 1467 HGT regions in the human genome, involving 642 known genes [17]. We do not expect that there will be an end to this debate in the near future.

Compared with HGT, gene transfer from the organellar genome to the nuclear genome, termed endosymbiotic gene transfer (EGT), has been widely observed and well accepted in eukaryotes [2, 18, 19]. Indeed, mitochondria and plastids are thought to be of endosymbiotic origin, and most of their genetic sequences have been transferred to the nuclear genome. Although the effective transfer of functional protein-coding genes has slowed or even stopped in some lineages, the process of DNA fragment transfer from organelle to nucleus still actively occurs in most eukaryotic lineages [19-22]. Mitochondrial DNA segments recently inserted into nuclear genomes are termed nuclear mitochondrial DNAs (NUMTs), whereas plastid DNA segments recently inserted into nuclear genomes are termed nuclear plastid DNAs (NUPTs). Furthermore, NUMTs and NUPTs are collectively termed nuclear organelle DNA (norgDNA) [23].

Both HGT and EGT are special cases of DNA transfer. In most cases, the transferred DNA is non-functional in the acceptor nucleus and is merely junk DNA. DNA transfer provides raw materials for gene transfer. Only when the transferred DNA contains entire genes and the transferred genes have been fixed in the evolution of the nuclear genome can we regard it as a case of HGT or EGT. If most of the transferred DNAs have been eliminated in natural selection, environmental factors that affect horizontal DNA transfer (HDT) and endosymbiotic DNA transfer (EDT) would have a very weak influence on the frequency of HGT and EGT. Only when the effects of the transferred DNAs are slightly deleterious or neutral to the host could the environmental factors that affect HDT and EDT have significant effects on the frequency of HGT and EGT.

In most multicellular animals and plants, the opportunity to directly pass genetic information to future generations is restricted to a small proportion of cells, i.e., germ line cells [24], which may serve as a strong barrier to HGT and HDT. However, in both animals and plants, a few weakly protected unicellular or early developmental stages exist in which there are some opportunities for environmental DNA to enter germ cells and be inherited [25]. Distinctive animals in the field of HGT are bdelloid rotifers. Although the frequency of HGT depends on the evaluation method used, the quality of the sequenced genome, and the specific species studied, a pattern of extensive HGT in bdelloid rotifers has consistently been reported [26-30]. Indeed, HGTs in bdelloid rotifers are so frequent that some researchers have suggested that this process can act as an alternate form of sex to facilitate the genetic exchange of rotifer DNA among these organisms [31]. Although evidence for genetic exchanges among bdelloid rotifers has been shown to be driven by cross-contamination in a recent paper [32], there is to our knowledge no controversy regarding the extensive pattern of HGT from non-metazoan species to bdelloid rotifers. Bdelloid rotifers are tiny invertebrates that live in ephemeral aquatic habitats and therefore frequently experience severe desiccation. When facing desiccation, bdelloid rotifers enter a form of dormancy called anhydrobiosis [33, 34], and the prolonged desiccation of the anhydrobiotic bdelloid rotifer *Adineta vaga* induces DNA double-strand breaks (DSBs) [35]. Such DSBs, which were vividly described as the "gateway to genetic exchange" by Hespeels et al. [35], are often believed to be the main cause of the high frequency of HGT in bdelloid rotifers [26-28, 30, 31, 35, 36]. The observed frequency of gene transfer in bdelloid rotifers has generally been attributed to the convenience of the transfer process rather than the fixation rate of the transferred genes. In this context, gene transfer has almost the same meaning as DNA transfer.

In addition to bdelloid rotifers, a number of other organisms also enter an anhydrobiotic state when their environment becomes dry. Another striking example is the phylum Tardigrada [34], in which prolonged desiccation has been shown to induce a small amount of DSBs [37]. Frequent HGT is naturally expected to occur in tardigrade species, by analogy with bdelloid rotifers, via a similar mechanism. Although no further support for the gateway hypothesis has been revealed by studies on tardigrades, the controversy of HGT in eukaryotes has been amplified. The first analysis of a tardigrade draft genome (*Hypsibius exemplaris*, originally misidentified with the congeneric *H. dujardini* [38]) suggested that HGT occurred more extensively in tardigrades than in bdelloid rotifers; 17.5% of its genes were identified as foreign and originating from bacteria, plants, fungi, and archaea [39]. However, subsequent independent sequencing of the same species and another tardigrade species, *Ramazzottius varieornatus*, consistently showed a much lower percentage of foreign genes in tardigrades than the prior study, with 2.3% reported as the upper bound [40-45]. As most previously identified foreign genes were found to be the result of DNA contamination from non-target organisms, tardigrades should not be regarded as a special group of animals with the high frequency of HGT [44]. Overall, the predominant objections to the claimed extensive HGT in tardigrades reject the gateway hypothesis and question whether prolonged desiccation truly facilitates HDT by inducing DSBs in the nuclear genome.

In this paper, we evaluated whether desiccation facilitates DNA transfer and provides increased raw materials for the fixation of gene transfer in evolution. Because donor sequences are readily available, EDT events are more conveniently identified than HDT events. From entering the nucleus to integrating into the nucleus genome, EDT and HDT processes share the same essential steps. For this reason, the environmental or physiological factors that affect the frequency of EDT likely have the same effect on HDT. Therefore, the effects of desiccation on EDT would provide direct evidence for the effects of desiccation on HDT, which, consequently, has implications for the effect of desiccation on the frequency of HGT when the frequency of HGT is assumed to be largely determined by HDT, as in bdelloid rotifers and tardigrades [26-28, 31, 35, 36, 39]. In this study, we sought to gain new insight into the debates on the effects of desiccation on HGT by analysing the abundance of NUMTs and NUPTs in eukaryotic organisms that frequently experience desiccation.

Results

A trade-off between genome quality and sample size

In the NCBI genome database [46], the assembly levels of nuclear genomes are divided into four categories: complete genome, chromosome, scaffold, and contig. For a trade-off between genome quality and sample size, we evaluated the effects of genome assembly level on the detected abundances of NUMTs/NUPTs using a phylogenetic comparative method, phylogenetic generalized least squares (PGLS). The contig level was designated 1, and the other three levels (scaffold, chromosome, and complete genome) were designated 2; the PGLS regression subsequently revealed that the assembly level was significantly correlated with the number and total length of NUPTs ($p < 0.05$ for both cases) and the number of NUMTs ($p = 0.037$) but not the total length of NUMTs (Table 1). These results indicate that the

abundances of NUMTs/NUPTs might be overestimated or underestimated when the analysed nuclear genomes are assembled at the contig level. Furthermore, we performed PGLS regression between NUMT/NUPT abundance and assembly level by assigning the scaffold level as 3 and the other two levels (chromosome and complete genome) as 4 (Table 1). The number of NUPTs was also significantly correlated with genome assembly level (p = 0.044). However, no significant correlations were detected between the assembly level and the number of NUMTs or the total length of NUMTs or NUPTs (Table 1). The scaffold level might also distort the estimation of the abundance of NUPTs, but the effects at the scaffold level are much weaker than those at the contig level. With a trade-off between the sample size and the accuracy of the estimation, we retained all of the nuclear genomes assembled at the scaffold, chromosome, and complete genome levels.

Table 1 Correlations between nuclear genome assembly level and the abundance of NUMTs/NUPTs.

	NUMT	NUPT
Number of species		
Contig (1)	16	8
Scaffold/chromosome/complete genome (2)	216	106
Number of NUMTs/NUPTs		
l	0.983	0.401
R^2	0.019	0.150
Slope	-0.713	1.393
p	0.037	2.1×10^{-5}
Total length of NUMTs/NUPTs		
l	0.942	0.718
R^2	0.008	0.046
Slope	-0.468	0.807
p	0.185	0.022
Number of species		
Scaffold (3)	151	56
Chromosome/complete genome (4)	65	50
Number of NUMTs/NUPTs		
l	0.985	0.507
R^2	4×10^{-4}	0.038
Slope	0.050	0.346
p	0.769	0.044
Total length of NUMTs/NUPTs		
l	0.935	0.745
R^2	0.004	0.015
Slope	-0.187	0.239
p	0.346	0.218

For one plant for which the data of both NUMTs and NUPTs were available, we regarded it as two samples, one NUMT sample and one NUPT sample, rather than combining them into one norgDNA sample. The correlations were analysed by phylogenetic generalized least squares (PGLS) regression. The numbers in the parentheses after the genome assembly levels are the numbers that were assigned to the assembly levels in the PGLS regression analysis. l is the phylogenetic signal. The approach of this value to one necessitates the use of phylogenetic comparative methods, such as PGLS.

The most desiccation-tolerant organisms do not have excessive NUMTs/NUPTs

Water deficiency is one of the most common abiotic stress factors for organisms living on land, and terrestrial organisms have evolved two solutions to cope with environmental drying [47]. The first is to conserve water and avoid severe body water deprivation, such as waxy coatings on plant shoots and the protective cocoon of the African lungfish, *Protopterus annectens* [48]. The second solution is to tolerate body water loss. The term anhydrobiosis is often used for the almost completely dehydrated but viable state of an organism experiencing extreme desiccation [49, 50]. Desiccation tolerance exists on a scale where the most desiccation-tolerant organisms are those that can enter anhydrobiosis and thus survive desiccation at any stage of their life cycle. The commonly studied organisms of this group are bdelloid rotifers, tardigrades, and resurrection plants [51-53].

DSBs induced by prolonged desiccation have been suggested to open the "gateway to genetic exchange" and account for the high frequency of HGT in bdelloid rotifers [26-28, 30, 31, 35]. However, prolonged desiccation in tardigrades does not result in an elevated level of HGT [44]. Using EDT as a proxy for HDT, we tested whether prolonged desiccation enhanced HDT in these desiccation-tolerant invertebrates. We detected 22 and 4 NUMTs, with total lengths of 32 kb and 14 kb, in the genomes of the desiccation-tolerant bdelloid rotifers *A. vaga* and *A. ricciae*, respectively. Similarly, in the genome of the desiccation-tolerant tardigrade *R. varieornatus*, we found 53 NUMTs, with a total length of 33 kb. If exceptionally high HGT levels in bdelloid rotifers are attributed to the anticipated high accessibility of their nuclear genomes to exogenous DNA, the NUMT content in bdelloid rotifers should also be exceptionally high compared with that in other invertebrates. In invertebrates, the NUMT density in honeybee *Apis mellifera* [54] is considered "exceptionally high". To reduce methodological artefacts, we re-surveyed the honeybee genome with our parameters and detected 1791 NUMTs, with a total length of 724 kb. We also surveyed the NUMT abundances of other invertebrates and confirmed that the NUMT contents of bdelloid rotifers and tardigrades are not exceptionally high (Table 2). The NUMT contents of the desiccation-tolerant bdelloid rotifers and tardigrades are similar to the median NUMT content of other invertebrates but are 13-185 times lower than the average NUMT number of other invertebrates and are 2-6 times lower than the average value of the total NUMT length in other invertebrates (Table 2). The high level of HGT in desiccation-tolerant bdelloid rotifers is not accompanied by a high frequency of EDT.

Table 2 The norgDNA contents of invertebrates and plants

	Type	Number of species	Number of norgDNAs	Total length of norgDNAs (kb)
Adineta vaga	NUMT		22	32
A. ricciae	NUMT		4	14
Ramazzottius varieornatus	NUMT		53	33
Other invertebrates	NUMT	193	742 ± 7810 (32)	97 ± 781 (15)
Dorcoceras hygrometricum	NUMT		3696	1232
	NUPT		1610	467
Selaginella tamariscina	NUPT		637	520
Other plants	NUMT	67	9309 ± 26177 (4182)	2735 ± 9044 (895)
	NUPT	118	3015 ± 3907 (2014)	940 ± 1280 (509)

The values of other invertebrates and other plants presented in this table are mean value ± S. D. (median value). Resurrection plants can survive extreme desiccation and maintain a quiescent state for months to years [55]. To examine the relationship between prolonged desiccation and the frequency of DNA transfer, we also surveyed the abundance of NUMTs/NUPTs in these plants. In the genome of the flowering plant *Dorcoceras hygrometricum*, we detected 3696 NUMTs with a total length of 1232 kb and 1610 NUPTs with a total length of 467 kb. In the genome of the spike moss *Selaginella tamariscina*, we detected 637 NUPTs with a total length of 520 kb. We did not obtain the mitochondrial genome sequences of *S. tamariscina* and consequently did not obtain any results regarding its NUMT content. The NUMT content of the resurrection plant *D. hygrometricum* was much higher than the NUMT contents of the desiccation-tolerant bdelloid rotifers and tardigrades (Table 2). Nonetheless, previous studies, together with our own survey, have shown that plant genomes generally have much higher NUMT contents than invertebrates [20, 21]. The norgDNA contents we detected in the resurrection plants were lower than the average and median values of the norgDNA contents in other plants except that the total length of NUPTs in *S. tamariscina* was a little longer than the median value of the lengths of NUPTs in other plants (520 *vs*. 509 kb, Table 2).

Pairwise comparison of organisms differing in desiccation tolerance

In the second grade of desiccation tolerance, anhydrobiosis is restricted to particular developmental stages, such as dormant eggs of the water flea *Daphnia*, cysts of primitive crustaceans such as the brine shrimp *Artemia salina*, and orthodox seeds of most angiosperms [56, 57]. In the last grade, all stages of the life cycle are desiccation sensitive, including the eggs of animals, recalcitrant seeds, and the embryos of viviparous plants. From this perspective, a viviparous plant or a plant with recalcitrant seeds is more sensitive to desiccation than a plant with orthodox seeds.

To reach a general conclusion regarding the relationship between desiccation and the frequency of EDT, we compared 24 pairs of lineages for which each pair of lineages differed in desiccation tolerance (Table 3). The closest relatives used as controls for desiccation-tolerant organisms were selected from widely used phylogenetic databases, including Timetree, NCBI taxonomy, and the Angiosperm Phylogeny

Website [58-60]. In cases where one phylogenetic branch contained two desiccation-tolerant species or two control species, we used the average value of the NUMT/NUPT contents of the two species. For example, the average value of desiccation-tolerant bdelloid rotifers *A. vaga* and *A. ricciae* was used to compare the average value of desiccation-sensitive bdelloid rotifers *R. magnacalcarata* and *R. macrura*. Nonparametric pairwise comparison did not reveal significant differences between desiccation-tolerant lineages and their controls in either the number of NUMTs/NUPTs (Wilcoxon signed-ranks test, $P = 0.269$, Fig. 1a) or the total length of NUMTs/NUPTs (Wilcoxon signed-ranks test, $P = 0.881$, Fig. 1b). Overall, a large nuclear genome is expected to have additional sites for NUMT/NUPT integration and is thus likely to contain additional NUMTs/NUPTs. To control for the influence of nuclear genome size, we compared the density of NUMTs/NUPTs in nuclear genomes and still found no significant differences between desiccation-tolerant lineages and their controls (Fig. 1c, d). These pairwise comparisons were also performed within invertebrates and within plants separately. Nevertheless, no significant differences were observed (Wilcoxon signed-ranks test, $P > 0.10$ for all the cases). Furthermore, we separated the 24 pairs into two groups according to the average values of the nuclear genome size of each pair. No significant differences were observed in either organisms with large nuclear genomes or organisms with small nuclear genomes (Wilcoxon signed-ranks test, $P > 0.10$ for all the cases).

Because insertions of norgDNAs into coding regions would disrupt gene function, a compact nuclear genome is a less accessible acceptor than a nuclear genome with a high percentage of noncoding sequences. Therefore, it was necessary to check whether the paired genomes had similar levels of compactness. From the NCBI genome database [46], the Ensembl Genomes database (Release 41) [61], and the Ensembl database (Release 94) [62], we got the annotation files of coding regions for six pairs of species. Pairwise comparison did not reveal a significant difference in genome compactness (Wilcoxon signed-ranks test, $P = 0.293$). As the NUMTs and NUPTs of *Sorghum bicolor* and *Zea mays* were regarded as independent samples, these six pairs of species have seven pairs of NUMT/NUPT data. The Wilcoxon signed-ranks test of these seven pairs did not reveal any significant differences in the abundances of NUMTs/NUPTs ($P > 0.10$).

Although most EDTs occur via direct transfer of organellar DNA to the nucleus, there is also evidence supporting EDT mediated by RNA molecules [63]. Therefore, we also compared the NUMTs/NUPTs contributed by the transcribed regions of organellar genomes, i.e., protein-coding genes, tRNA, and rRNA. Similar to the above results, no significant difference was detected between desiccation-tolerant species and their controls (Wilcoxon signed-ranks test, $P > 0.10$ for all cases, Additional file 1: Table S1).

Table 3 Desiccation-tolerant organisms and their controls used in this study

	Desiccation-tolerant species	Control
Invertebrates		
Bdelloidea	*Adineta ricciae* and *A. vaga* [29]*{Nowell, 2018 #30}*	*Rotaria macrura* and *R. magnacalcarata* [29]
Hypsibioidea	*Ramazzottius varieornatus* [44]	*Hypsibius dujardini* [44, 102]
Culicidae	*Aedes aegypti* [105, 106]*{Nowell, 2018 #30}*	*Culex quinquefasciatus* [107]
Anopheles	*Anopheles culicifacies* [108]	*Anopheles funestus* [109]
Dermatophagoides	*Dermatophagoides farina* [110]	*Dermatophagoides pteronyssinus* [110]
Meloidogyne	*Meloidogyne javanica* [111]	*Meloidogyne incognita* [111]
Polypedilum	*Polypedilum vanderplanki* [112]	*Polypedilum nubifer* [112]
Steinernema	*Steinernema carpocapsae* [113]	*Steinernema glaseri* [113]
Plants		
Pooideae	*Aegilops tauschii* [114]	*Lolium perenne* [115]
Fagales	*Betula nana* [101]	*Juglans regia* [116]
Malvaceae	*Corchorus capsularis* [101]	*Theobroma cacao* [117]
Pentapetalae	*Dorcoceras hygrometricum* [118]	*Fagopyrum esculentum* [119]
Malvoidae	*Hibiscus syriacus* [120]	*Gossypium hirsutum* [121]
Crotonoideae	*Manihot esculenta* [122]	*Hevea brasiliensis* [123]
Oryzeae	*Oryza sativa Japonica Group* [101]	*Zizania latifolia* [124]
Ericales	*Primula veris* [125]	*Actinidia chinensis* [126]
Selaginella	*Selaginella tamariscina* [127]	*Selaginella moellendorffii* [127]
Caryophyllales	*Silene latifolia* [128]	*Spinacia oleracea* [129]
Solanum	*Solanum commersonii* [130]	*Solanum tuberosum* [131]
Solanum	*Solanum pennellii* [132]	*Solanum melongena* [133]
Andropogonodae	*Sorghum bicolor* [134]	*Zea mays* [135]
Fungi		
Onygenales	*Coccidioides immitis* [136]	*Paracoccidioides brasiliensis* [137]
Hypocreales	*Fusarium graminearum* [138]	*Ophiocordyceps sinensis* [139]
Pseudogymnoascus	*Pseudogymnoascus pannorum* [140]	*Pseudogymnoascus destructans* [140]

The particular words or sentences supporting the desiccation tolerance/sensitivity of the above species were extracted from the references and are shown in Additional file 1: Table S3.

Discussion

Desiccation is a widely observed cause of DSB formation in bacteria, plants and animals, including bdelloid rotifers [33-35, 64-72]. Organisms that frequently encounter desiccation stress have evolved efficient mechanisms to repair DSBs [73]. Accordingly, the radioresistance of *Deinococcus radiodurans* and bdelloid rotifers has been attributed to their adaptation to desiccation stress [64, 74]. As the DSBs in nuclear genomes are opportunities for the integration of exogenous DNA, DSBs have been postulated to be a gateway for HGT [35]. If the DSBs of desiccation-tolerant organisms serve as an efficient gateway for exogenous DNA integration, the frequency of EDT should also be dramatically increased by desiccation. However, we did not observe excessive NUMTs/NUPTs in the genomes of the desiccation-tolerant organisms, the bdelloid rotifers *A. vaga* and *A. ricciae*, the tardigrade *R. varieornatus*, or the resurrection plants *D. hygrometricum* or *S. tamariscina*. Further pairwise comparison of 24 desiccation-tolerant lineages with their controls also did not reveal evidence for desiccation-elevated EDT frequency.

Notably, there are also a few exceptions to the general observations that desiccation could induce DSBs. In the anhydrobiotic tardigrade *Milnesium tardigradum*, the process of desiccation also induces DNA damage but at only minor levels [37]. In the tardigrade *R. varieornatus*, a protein that suppresses DNA breaks has been identified [45]. In some bacteria, *Nostoc commune*, *Chroococcidiopsis* sp. CCMEE 029 and possibly *Bacillus pumilus*, no evidence of extensive genome fragmentation was observed in cells following prolonged desiccation [69, 75, 76]. Some desiccation-tolerant eukaryotes might lack an elevated EDT frequency because they happen to be such exceptions in which desiccation does not induce DSBs. Furthermore, even in organisms in which desiccation does induce DSBs, frequent DSBs do not necessarily increase the frequency of exogenous DNA integration into nuclear genomes. Indeed, the integration of exogenous DNA, i.e., HDT and EDT, depends on the specific mechanisms of DSB repair, for which there are two main categories: homologous recombination (HR) and non-homologous end joining (NHEJ). The capture of exogenous DNA sequences is possible only when DSBs are repaired through a subtype of NHEJ, named alternative end joining (alt-EJ), which is initiated much less frequently than other mechanisms [77-79]. The choice between the HR and NHEJ repair pathways and among different subtypes of NHEJ depends on phylogenetic position, cell cycle phase and repetitive sequence content [80-82]. If an organism predominantly uses HR for DSB repair, desiccation-induced DSBs do not open a gateway for HDT and EDT. In general, organisms with compact genomes, such as yeasts, preferentially use HR; in contrast, NHEJ is the predominant DSB repair pathway in mammals and plants with large and complex genomes. We did not observe significant differences between desiccation-tolerant plants and their controls. In addition, significant differences were not detected in organisms with relatively larger nuclear genomes. Exceptions to the general rule between genome size and HR/NHEJ usage, bdelloids have more copies of gene involved in NHEJ than other animals and thus preferentially use NHEJ in their DSB repairs [83]. However, we did not observed a higher NUMT content in bdelloids than other invertebrates.

If desiccation-induced DSBs open the gateway for exogenous DNA but the availability of organellar DNA fragments is under strong constraints, the frequency of NUMTs/NUPTs would not be elevated by desiccation. Organisms such as *Ostreococcus tauri* have a single plastid or mitochondrion in each cell. The lysis of these organelles would almost certainly result in cell death. These organisms, with a limited availability of donor organellar DNA, have been shown to have much lower amounts of NUMTs or NUPTs than organisms with multiple organelles per cell [84]. However, all of the organisms we studied have multiple organelles per cell, and the lysis of organelles and release of organellar DNA fragments are not under strong constraints during desiccation.

As NUMTs/NUPTs are insertion mutations, their abundance and the frequency of EGT/HGT depend not only on the insertion rate but also on the fixation possibility during evolution [85]. For a compact nuclear genome with a limited amount of junk DNA sequences, most insertions would disrupt the nuclear genetic information and thus be selected against. Similar to previous reports [20, 84, 86], we observed a significant positive correlation between the abundance of NUMTs/NUPTs and nuclear genome size in our dataset (Spearman's rho = 0.737 and 0.803 for NUMT/NUPT number and total length, respectively, and *P* < 0.01 for both cases).

Our study was based on an assumption that the frequency of EDT could be used as a proxy for the frequency of HDT, which in turn is probably correlated with the frequency of HGT. This assumption would be invalidated if organellar DNAs could not exit organellar membranes, while environmental DNAs could cross the plasma membrane during desiccation. From the properties of protectants (intrinsically disordered proteins, trehalose, etc.) that can prevent the loss of membrane integrity [87-90], the preferential protection of organellar membranes were not expected. Moore et al. [91] suggested that intrinsically disordered proteins could offset the damage during the desiccation of liposomes that mimic the lipid compositions of the mitochondrial membrane and the plasma membrane. More importantly, they found that liposome stabilization by intrinsically disordered proteins was not dependent on the lipid composition. Thus, there is no evidence for the preferential protection of the mitochondrial membrane over the plasma membrane or the plasma membrane over the mitochondrial membrane.

It should be noted that the diversity of environmental DNA sequences is theoretically unlimited, significantly larger than that of the organellar DNA sequences. Meanwhile, for most eukaryotes, the copy number of the organellar DNA sequences in a cell is larger than most environmental DNA sequences. We are not interested in comparing the absolute amount of NUMT/NUPT with the nuclear DNA sequences obtained recently from environments. Based on the assumption is that environmental DNA and organellar DNA share the same mechanisms from entering the nucleus to the integration into the nuclear DNA, we proposed that a eukaryote with higher NUMT/NUPT content should acquire more environmental DNA than other eukaryotes.

In this study, we examined whether desiccation facilitates the integration of exogenous DNA into nuclear genomes. In principle, desiccation could also induce DSBs in mitochondrial and plastid genomes. It will be interesting to determine whether desiccation can facilitate the transfer of exogenous DNA segments (including both nuclear DNA and DNA from the environment) into organellar genomes. Mitochondrial and plastid DNA of nuclear origin are termed MINC and PLNC, respectively [92]. With references to the organellar/nuclear genes or nuclear retrotransposons, some MINCs and PLNCs could be differentiated from NUMTs and NUPTs. Compared with plastid genomes and animal mitochondrial genomes, plant mitochondrial genomes are unique in their excessive size relative to their coding capacity [93]. The significant amount of noncoding DNA in plant mitochondrial genomes makes it difficult to differentiate NUMTs and MINCs in most shared nuclear−mitochondrial segments [92]. As far as we know, the shared nuclear−mitochondrial segments are simply regarded as NUMTs in all previous genome-wide studies of NUMTs [20, 21, 54, 84, 94-96]. In the present study, we adopted the same approximation. We suggest that MINCs are significantly less frequent than NUMTs; therefore, such approximations are unlikely to affect the conclusions qualitatively. This conjecture was inspired by the paucity of NUMTs/NUPTs in organisms that have only one mitochondrion or one plastid per cell [84, 94]. The lysis of a single mitochondrion or plastid would almost certainly result in cell death; therefore, mono-mitochondria and monoplastids restrict the production of norgDNAs. With the same logic, a single nucleus also restricts the formation of MINCs and PLNCs because the lysis of the nucleus and release of nuclear DNA segments into the cytoplasm would result in cell death. Nevertheless, we hope to distinguish NUMTs and MINCs accurately in the future once additional unique plant mitochondrial sequences are identified [97].

Conclusions

In this study, we did not find evidence that desiccation significantly increased the norgDNA content. The results are inconsistent with the hypothesis that prolonged desiccation could open a gateway to the nuclear genome for exogenous DNA integration. Therefore, the prolonged desiccation experienced by bdelloid rotifers, tardigrades, and resurrection plants likely do not provide additional raw materials (i.e., HDTs) for HGT. However, notably, the fate of HDTs might be entirely different from that of EDTs. After a long evolutionary history, most organellar genes have been transferred to nuclear genomes, and only tens to a few hundreds of protein-coding genes are retained in organellar genomes. There is an upper limit of further successful EGT. In contrast, there is no such an upper limit for HGT. Any DNA sequences acquired from other species or even other phylogenetic domains that could help the host survive under severe stresses, such as desiccation, will be favoured in natural selection and fixed in the host nuclear genome. Therefore, our results should not be regarded as evidence against frequent HGT in bdelloid rotifers and tardigrades. Readers interested in this subject should refer to recent studies on HGT, such as [44] and [29].

Methods

From the widely used genome databases [46, 61, 98], we first retrieved a list of eukaryotes whose nuclear genomes and at least one organellar genome (mitochondrion or plastid) had been sequenced. To avoid interference from the effects of a single plastid or single mitochondrion per cell on the abundance of NUMTs or NUPTs [84], protists were not included in our study. As vertebrates generally use the strategy of avoiding severe body water deprivation to survive in water deficient environments rather than mechanisms to adapt cellular dehydration [48], they are not suitable materials to examine the effects of desiccation on organelle DNA accumulation in nuclear genomes.

Selection of desiccation-tolerant species and their controls

Then, we searched the desiccation-tolerant species and their controls using the Latin names of the above species together with desiccation tolerant, drought tolerant, desiccation sensitive, drought sensitive, desiccation resistant, drought resistant, orthodox seed, recalcitrant seed, viviparous plant, resurrection plant, xerophyte, xerotolerance, xerophilic, and anhydrobiosis as keywords in Google [99] and Google Scholar [100]. In addition, we retrieved a species list of the recalcitrant and orthodox seeds from the Seed Information Database [101]. Plants with recalcitrant seeds and viviparous plants do not have a desiccation-tolerant seed stage in their life cycle. These plants can be regarded as more sensitive to desiccation than other plants with dry seeds, i.e., orthodox seeds. For each desiccation-tolerant lineage, we selected the most closely related and relatively less tolerant lineage to pair with it. The phylogenetic relatedness was determined using three phylogenetic databases: Timetree, NCBI taxonomy, and the Angiosperm Phylogeny Website [58-60]. Each lineage contained one or two species from a genus.

Although tardigrades are often regarded as models of desiccation-tolerant organisms, clear differences in desiccation tolerance have been demonstrated among them. The species *H. dujardini* is a poor

anhydrobiote compared with other tardigrades [44, 102]. Therefore, *H. dujardini* was used as the control of the desiccation-tolerant tardigrade *R. varieornatus*.

During the above literature searching, we also found some desiccation-tolerant species (and/or their controls) whose mitochondrial genome sequences were not available in public databases. For these species, we acquired four mitochondrial genome sequences (*A. vaga*, *A. ricciae*, *R. magnacalcarata*, and *R. macrura*) from the authors of the reference [29]. Then, we assembled the mitochondrial genome sequences for one invertebrate, six fungi and eight plants using NOVOPlasty [103] and for another invertebrate (*Polypedilum nubifer*) using Geneious Prime (2019.2.1). The IDs of the whole-genome sequencing datasets, the mitochondrial genome sequence seeds, the plastid genome sequences, and the mitochondrial genomes used as references in assembling the mitochondrial genomes are shown in Additional file 1: Table S2. The mitochondrial genome sequences that we assembled are shown in Additional file 2: Data S1.

Finally, we obtained 24 pairs of desiccation-tolerant *vs.* control lineages (Table 3 and Additional file 1: Table S3).

Detection and filtering of NUMTs and NUPTs

In addition to the mitochondrial genome sequences mentioned above, we downloaded other organellar genome sequences from the NCBI nucleotide database [98]. The nuclear genome sequences were downloaded from the NCBI genome database [46], the Ensembl Genomes database (Release 41) [61], and the Ensembl database (Release 94) [62]. The assembly accession number and exact source of each genome are shown in Additional file 1: Table S4.

The NUMTs and NUPTs of some model species have been well explored [20, 21, 84, 95, 96]. Because the different parameters used in previous studies and different levels of genome assembly would likely lead to discrepancies in NUPT/NUMT abundance [21, 84, 104], we performed our own detection using the same set of parameters and with the expectation that over- or underestimation affects all species consistently. Organellar genome sequences were used as the queries to search against nuclear genomes using blastn (version 2.4.0+) with a threshold E value of 0.0001, a word size of 11, a match score of 2, a mismatch score of −3, and gap cost values of 5 (existence) and 2 (extension), as described by Smith et al. [84]. To reduce false-positive hits resulting from the contamination of organellar sequences in nuclear genome sequences, short nuclear contigs that completely matched to organellar genome sequences were discarded. Redundancy in the blastn results was eliminated as formulated below. M_1 and M_2 were assumed to be two sequences residing in different positions of the same mitochondrial genome, and four nuclear sequences (N_1, N_2, N_3, and N_4) were assumed to independently reside in four different chromosomes. In the blastn results, M_1 was matched to N_1 and N_2, whereas M_2 was matched to N_2, N_3, and N_4. Both M_1 and M_2 might be the parent sequence of N_2. However, N_2 originated from either M_1 or M_2. Therefore, N_2 was found twice but should be counted only once in the calculation of the NUMT content. In total, the two mitochondrial sequences produced four NUMTs. The distribution of the identities

between the NUMTs/NUPTs and their parent organellar sequences for all species presented in Table 3 are shown in Additional file 3: Fig. S1.

We also separately detected the NUMTs and NUPTs contributed by organellar sequences with known functions, including protein-coding genes, rRNA genes, and tRNA genes, using the above method and parameters.

Abbreviations

HGT: horizontal gene transfer; HDT: horizontal DNA transfer; EGT: endosymbiotic gene transfer; EDT: endosymbiotic DNA transfer; NUMT: nuclear mitochondrial DNA; NUPT: nuclear plastid DNA; MINC: mitochondrial DNA of nuclear origin; PLNC: plastid DNA of nuclear origin; norgDNA: nuclear organelle DNA; DSB: double-strand break; NHEJ: non-homologous end joining; alt-EJ: alternative end joining; HR: homologous recombination; PGLS: phylogenetic generalized least squares.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Authors' contributions

D.K.N. conceived the study and wrote the manuscript. X.X.L. retrieved the data from online databases and performed the analyses. C.F., J.P.Z., X.Y.Z., and Z. N. each verified some of the results. All authors read and approved the final manuscript.

References

1. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet. 2008;9:605-18.

2. Soucy SM, Huang JL, Gogarten JP. Horizontal gene transfer: building the web of life. Nat Rev Genet. 2015;16:472-82.

3. Koonin E. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. F1000Res. 2016;5:1805.

4. Moreira D, López-García P. Protist evolution: Stealing genes to gut it out. Curr Biol. 2017;27:R223-R5.

5. Danchin EGJ. Lateral gene transfer in eukaryotes: tip of the iceberg or of the ice cube? BMC Biol. 2016;14:101.

6. McInerney JO. Horizontal gene transfer is less frequent in eukaryotes than prokaryotes but can be important (retrospective on DOI 10.1002/bies.201300095). BioEssays. 2017;39:e201700002.

7. Ku C, Martin WF. A natural barrier to lateral gene transfer from prokaryotes to eukaryotes revealed from genomes: the 70 % rule. BMC Biol. 2016;14:89.

8. Leger MM, Eme L, Stairs CW, Roger AJ. Demystifying eukaryote lateral gene transfer. BioEssays. 2018;40:5.

9. Martin WF. Too much eukaryote LGT. Bioessays. 2017;39:1700115.

10. Husnik F, McCutcheon JP. Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol. 2017;16:67.

11. Sieber KB, Bromley RE, Dunning Hotopp JC. Lateral gene transfer between prokaryotes and eukaryotes. Exp Cell Res. 2017;358:421-6.

12. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860-921.

13. Stanhope MJ, Lupas A, Italia MJ, Koretke KK, Volker C, Brown JR. Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature. 2001;411:940-4.

14. Salzberg SL, White O, Peterson J, Eisen JA. Microbial genes in the human genome: Lateral transfer or gene loss? Science. 2001;292:1903-6.

15. Crisp A, Boschetti C, Perry M, Tunnacliffe A, Micklem G. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. Genome Biol. 2015;16:50.

16. Salzberg SL. Horizontal gene transfer is not a hallmark of the human genome. Genome Biol. 2017;18.

17. Huang WZ, Tsai LL, Li YL, Hua N, Sun C, Wei CC. Widespread of horizontal gene transfer in the human genome. BMC Genomics. 2017;18:274.

18. Ku C, Nelson-Sathi S, Roettger M, Garg S, Hazkani-Covo E, Martin WF. Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. Proc Natl Acad Sci USA. 2015;112:10139-46.

19. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004;5:123-35.

20. Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: Mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genet. 2010;6:e1000834.

21. Hazkani-Covo E, Martin WF. Quantifying the number of independent organelle DNA insertions in genome evolution and human health. Genome Biol Evol. 2017;9:1190-203.

22. Kleine T, Maier UG, Leister D. DNA transfer from organelles to the nucleus: The idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol. 2009;60:115-38.

23. Leister D. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends Genet. 2005;21:655-63.

24. Jensen L, Grant JR, Laughinghouse HD, Katz LA. Assessing the effects of a sequestered germline on interdomain lateral gene transfer in Metazoa. Evolution. 2016;70:1322-33.

25. Huang JL. Horizontal gene transfer in eukaryotes: The weak-link model. Bioessays. 2013;35:868-75.

26. Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EGJ et al. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. Nature. 2013;500:453-7.

27. Eyres I, Boschetti C, Crisp A, Smith TP, Fontaneto D, Tunnacliffe A et al. Horizontal gene transfer in bdelloid rotifers is ancient, ongoing and more frequent in species from desiccating habitats. BMC Biol. 2015;13:90.

28. Gladyshev EA, Meselson M, Arkhipova IR. Massive horizontal gene transfer in bdelloid rotifers. Science. 2008;320:1210-3.

29. Nowell RW, Almeida P, Wilson CG, Smith TP, Fontaneto D, Crisp A et al. Comparative genomics of bdelloid rotifers: Insights from desiccating and nondesiccating species. PLoS Biol. 2018;16:e2004830.

30. Boschetti C, Carr A, Crisp A, Eyres I, Wang-Koh Y, Lubzens E et al. Biochemical diversification through foreign gene expression in bdelloid rotifers. PLoS Genet. 2012;8:e1003035.

31. Debortoli N, Li X, Eyres I, Fontaneto D, Hespeels B, Tang CQ et al. Genetic exchange among bdelloid rotifers is more likely due to horizontal gene transfer than to meiotic sex. Curr Biol. 2016;26:723-32.

32. Wilson CG, Nowell RW, Barraclough TG. Cross-contamination explains "Inter and intraspecific horizontal genetic transfers" between asexual bdelloid rotifers. Curr Biol. 2018;28:2436-44.e14.

33. Ricci C. Bdelloid rotifers: 'sleeping beauties' and 'evolutionary scandals', but not only. Hydrobiologia. 2017;796:277-85.

34. Gusev O, Okuda T. "Life without water": the sleeping chironomid and other anhydrobiotic invertebrates and their utilization in astrobiology. In: Hanslmeier A, Kempe S, Seckbach J, editors. Life on Earth and other Planetary Bodies. Dordrecht: Springer Netherlands; 2012. p. 121-38.

35. Hespeels B, Knapen M, Hanot-Mambres D, Heuskin AC, Pineux F, Lucas S et al. Gateway to genetic exchange? DNA double-strand breaks in the bdelloid rotifer *Adineta vaga* submitted to desiccation. J Evol Biol. 2014;27:1334-45.

36. Gladyshev EA, Arkhipova IR. Genome structure of bdelloid rotifers: Shaped by asexuality or desiccation? J Hered. 2010;101:S85-S93.

37. Neumann S, Reuner A, Brümmer F, Schill RO. DNA damage in storage cells of anhydrobiotic tardigrades. Comp Biochem Physiol A Mol Integr Physiol. 2009;153:425-9.

38. Gasiorek P, Stec D, Morek W, Michalczyk L. An integrative redescription of *Hypsibius dujardini* (Doyere, 1840), the nominal taxon for Hypsibioidea (Tardigrada: Eutardigrada). Zootaxa. 2018;4415:45-75.

39. Boothby TC, Tenlen JR, Smith FW, Wang JR, Patanella KA, Osborne Nishimura E et al. Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. Proc Natl Acad Sci USA. 2015;112:15976-81.

40. Arakawa K. No evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. Proc Natl Acad Sci USA. 2016;113:E3057.

41. Bemm F, Weiß CL, Schultz J, Förster F. Genome of a tardigrade: Horizontal gene transfer or bacterial contamination? Proc Natl Acad Sci USA. 2016;113:E3054-E6.

42. Delmont TO, Eren AM. Identifying contamination with advanced visualization and analysis practices: metagenomic approaches for eukaryotic genome assemblies. PeerJ. 2016;4:e1839.

43. Koutsovoulos G, Kumar S, Laetsch DR, Stevens L, Daub J, Conlon C et al. No evidence for extensive horizontal gene transfer in the genome of the tardigrade *Hypsibius dujardini*. Proc Natl Acad Sci USA. 2016;113:5053-8.

44. Yoshida Y, Koutsovoulos G, Laetsch DR, Stevens L, Kumar S, Horikawa DD et al. Comparative genomics of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus*. PLoS Biol. 2017;15:e2002266.

45. Hashimoto T, Horikawa DD, Saito Y, Kuwahara H, Kozuka-Hata H, Shin-I T et al. Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. Nat Commun. 2016;7:12808.

46. NCBI genome database. ftp://ftp.ncbi.nih.gov/genomes/.

47. Alpert P, Oliver MJ. Drying without dying. In: Black M, Pritchard HW, editors. Desiccation and Survival in Plants: Drying without Dying. Wallingford: CABI Publishing; 2002.

48. Procko C, Shaham S. Stress. Curr Biol. 2011;21:R908-R10.

49. Crowe JH, Hoekstra FA, Crowe LM. Anhydrobiosis. Annu Rev Physiol. 1992;54:579-99.

50. Wharton DA. Anhydrobiosis. Curr Biol. 2015;25:R1114-R6.

51. Giarola V, Hou QC, Bartels D. Angiosperm plant desiccation tolerance: hints from transcriptomics and genome sequencing. Trends Plant Sci. 2017;22:705-17.

52. Costa MCD, Cooper K, Hilhorst HWM, Farranta JM. Orthodox seeds and resurrection plants: Two of a kind? Plant Physiol. 2017;175:589-99.

53. Farrant JM, Moore JP. Programming desiccation-tolerance: from plants to seeds to resurrection plants. Curr Opin Plant Biol. 2011;14:340-5.

54. Pamilo P, Viljakainen L, Vihavainen A. Exceptionally high density of NUMTs in the honeybee genome. Mol Biol Evol. 2007;24:1340-6.

55. VanBuren R. Desiccation tolerance: Seedy origins of resurrection. Nature Plants. 2017;3:17046.

56. Farnsworth E. The ecology and physiology of viviparous and recalcitrant seeds. Annu Rev Ecol Syst. 2000;31:107-38.

57. Watanabe M. Anhydrobiosis in invertebrates. Appl Entomol Zoolog. 2006;41:15-31.

58. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol. 2017;34:1812-9.

59. Federhen S. The NCBI Taxonomy database. Nucleic Acids Res. 2012;40:D136-D43.

60. Angiosperm Phylogeny Website http://www.mobot.org/MOBOT/research/APweb/welcome.html.

61. Ensembl Genomes. ftp://ftp.ensemblgenomes.org/pub/release-41.

62. Ensembl. ftp://ftp.ensembl.org/.

63. Timmis JN. Endosymbiotic evolution: RNA intermediates in endosymbiotic gene transfer. Curr Biol. 2012;22:R296-R8.

64. Mattimore V, Battista JR. Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. J Bacteriol. 1996;178:633-7.

65. Pitcher RS, Green AJ, Brzostek A, Korycka-Machala M, Dziadek J, Doherty AJ. NHEJ protects mycobacteria in stationary phase against the harmful effects of desiccation. DNA Repair. 2007;6:1271-6.

66. Waterworth WM, Drury GE, Bray CM, West CE. Repairing breaks in the plant genome: the importance of keeping it together. New Phytol. 2011;192:805-22.

67. Faria JMR, Buitink J, van Lammeren AAM, Hilhorst HWM. Changes in DNA and microtubules during loss and re-establishment of desiccation tolerance in germinating *Medicago truncatula* seeds. J Exp Bot. 2005;56:2119-30.

68. Gusev O, Nakahara Y, Vanyagina V, Malutina L, Cornette R, Sakashita T et al. Anhydrobiosis-associated nuclear DNA damage and repair in the sleeping Chironomid: Linkage with radioresistance. PLoS ONE. 2010;5:e14008.

69. Yang Y, Yokobori S, Yamagishi A. Bacterial survival in response to desiccation and high humidity at above zero and subzero temperatures. Adv Space Res. 2009;43:1285-90.

70. Waterworth WM, Footitt S, Bray CM, Finch-Savage WE, West CE. DNA damage checkpoint kinase ATM regulates germination and maintains genome stability in seeds. Proc Natl Acad Sci USA. 2016;113:9647-52.

71. Parreira JR, Balestrazzi A, Fevereiro P, Araújo SDS. Maintaining genome integrity during seed development in *Phaseolus vulgaris* L.: Evidence from a transcriptomic profiling study. Genes. 2018;9:463.

72. Dupuy P, Gourion B, Sauviac L, Bruand C. DNA double-strand break repair is involved in desiccation resistance of *Sinorhizobium meliloti*, but is not essential for its symbiotic interaction with *Medicago truncatula*. Microbiology. 2017;163:333-42.

73. de Groot A, Dulermo R, Ortet P, Blanchard L, Guérin P, Fernandez B et al. Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium *Deinococcus deserti*. PLoS Genet. 2009;5:e1000434.

74. Gladyshev E, Meselson M. Extreme resistance of bdelloid rotifers to ionizing radiation. Proc Natl Acad Sci USA. 2008;105:5139-44.

75. Billi D. Subcellular integrities in *Chroococcidiopsis sp*. CCMEE 029 survivors after prolonged desiccation revealed by molecular probes and genome stability assays. Extremophiles. 2008;13:49.

76. Shirkey B, McMaster NJ, Smith SC, Wright DJ, Rodriguez H, Jaruga P et al. Genomic DNA of *Nostoc commune* (Cyanobacteria) becomes covalently modified during long-term (decades) desiccation but is protected from oxidative damage and degradation. Nucleic Acids Res. 2003;31:2995-3005.

77. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair pathway choices and consequences at the double-strand break. Trends Cell Biol. 2016;26:52-64.

78. Ono R, Ishii M, Fujihara Y, Kitazawa M, Usami T, Kaneko-Ishino T et al. Double strand break repair by capture of retrotransposon sequences and reverse-transcribed spliced mRNA sequences in mouse zygotes. Sci Rep. 2015;5:12281.

79. Dupuy P, Sauviac L, Bruand C. Stress-inducible NHEJ in bacteria: function in DNA repair and acquisition of heterologous DNA. Nucleic Acids Res. 2019;47:1335-49.

80. Her J, Bunting SF. How cells ensure correct repair of DNA double-strand breaks. J Biol Chem. 2018;293:10502-11.

81. Lieber MR, Karanjawala ZE. Ageing, repetitive genomes and DNA damage. Nat Rev Mol Cell Biol. 2004;5:69-75.

82. Mao Z, Bozzella M, Seluanov A, Gorbunova V. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair. 2008;7:1765-71.

83. Hecox-Lea BJ, Mark Welch DB. Evolutionary diversity and novelty of DNA repair genes in asexual Bdelloid rotifers. BMC Evol Biol. 2018;18:177.

84. Smith DR, Crosby K, Lee RW. Correlation between nuclear plastid DNA abundance and plastid number supports the limited transfer window hypothesis. Genome Biol Evol. 2011;3:365-71.

85. Schonknecht G, Weber APM, Lercher MJ. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. Bioessays. 2014;36:9-20.

86. Yoshida T, Furihata HY, Kawabe A. Patterns of genomic integration of nuclear chloroplast DNA fragments in plant species. DNA Res. 2014;21:127-40.

87. Koshland D, Tapia H. Desiccation tolerance: an unusual window into stress biology. Mol Biol Cell. 2019;30:737-41.

88. Shivaraj YN, Barbara P, Gugi B, Vicre-Gibo M, Driouich A, Govind SR et al. Perspectives on structural, physiological, cellular, and molecular responses to desiccation in resurrection plants. Scientifica. 2018;2018:9464592.

89. Janis B, Belott C, Menze MA. Role of intrinsic disorder in animal desiccation tolerance. Proteomics. 2018;18:e1800067.

90. Kalemba EM, Bagniewska-Zadworna A, Ratajczak E. Multiple subcellular localizations of dehydrin-like proteins in the embryonic axes of common beech (*Fagus sylvatica* L.) seeds during maturation and dry storage. J Plant Growth Regul. 2015;34:137-49.

91. Moore DS, Hansen R, Hand SC. Liposomes with diverse compositions are protected during desiccation by LEA proteins from *Artemia franciscana* and trehalose. BBA-Biomembranes. 2016;1858:104-15.

92. Mower JP, Jain K, Hepburn NJ. The role of horizontal transfer in shaping the plant mitochondrial genome. In: Marechal Drouard L, editor. Mitochondrial Genome Evolution. Advances in Botanical Research. London: Academic Press Ltd-Elsevier Science Ltd; 2012. p. 41-69.

93. Morley SA, Nielsen BL. Plant mitochondrial DNA. Front Biosci. 2017;22:1023-32.

94. Smith DR. Extending the limited transfer window hypothesis to inter-organelle DNA migration. Genome Biol Evol. 2011;3:743-8.

95. Richly E, Leister D. NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. Mol Biol Evol. 2004;21:1972-80.

96. Richly E, Leister D. NUMTs in sequenced eukaryotic genomes. Mol Biol Evol. 2004;21:1081-4.

97. Wynn EL, Christensen AC. Repeats of unusual size in plant mitochondrial genomes: identification, incidence and evolution. G3 Genes Genom Genet. 2019;9:549-59.

98. NCBI nucleotide database. https://www.ncbi.nlm.nih.gov/nucleotide/.

99. Google. http://www.google.cn.

100. Google Scholar. http://scholar.google.com.cn.

101. Seed Information Database (SID). Version 7.1 [database on the Internet]. 2019. Available from: http://data.kew.org/sid/. Accessed: May 2019

102. Wright JC. Desiccation tolerance and water-retentive mechanisms in tardigrades. J Exp Biol. 1989;142:267-92.

103. Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 2017;45:e18-e.

104. Hazkani-Covo E, Martin WF. Quantifying the number of independent organelle DNA insertions in genome evolution and human health. Genome Biol Evol. 2017.

105. Rezende GL, Martins AJ, Gentile C, Farnesi LC, Pelajo-Machado M, Peixoto AA et al. Embryonic desiccation resistance in *Aedes aegypti*: presumptive role of the chitinized serosal cuticle. BMC Dev Biol. 2008;8:82.

106. Faull KJ, Williams CR. Intraspecific variation in desiccation survival time of *Aedes aegypti* (L.) mosquito eggs of Australian origin. J Vector Ecol. 2015;40:292-300.

107. Bhattacharya S, Basu P, Sajal Bhattacharya C. The southern house mosquito, *Culex quinquefasciatus*: profile of a smart vector. J Entomol Zool Studies. 2016;4:73-81.

108. Sharma S, Prasad R. Bionomics of *Anopheles culicifacies* giles in riverine tract rural areas of district Shahjahanpur, Uttar Pradesh. Indian J Malariol. 1991;28:19-28.

109. Ogola EO, Odero JO, Mwangangi JM, Masiga DK, Tchouassi DP. Population genetics of *Anopheles funestus*, the African malaria vector, Kenya. Parasites Vectors. 2019;12:15.

110. Zock J, Heinrich J, Jarvis D, Verlato G, Norback D, Plana E et al. Distribution and determinants of house dust mite allergens in Europe: the European Community Respiratory Health Survey II. J Allergy Clin Immunol. 2006;118:682-90.

111. Ephytia. http://ephytia.inra.fr/en/C/10822/Tobacco-Root-knot-nematodes-Meloidogyne-spp.

112. Gusev O, Suetsugu Y, Cornette R, Kawashima T, Logacheva MD, Kondrashov AS et al. Comparative genome sequencing reveals genomic signature of extreme desiccation tolerance in the anhydrobiotic midge. Nat Commun. 2014;5:4784.

113. Shapiro-Ilan DI, Brown I, Lewis EE. Freezing and desiccation tolerance in entomopathogenic nematodes: diversity and correlation of traits. J Nematol. 2014;46:27.

114. Sohail Q, Inoue T, Tanaka H, Eltayeb AE, Matsuoka Y, Tsujimoto H. Applicability of *Aegilops tauschii* drought tolerance traits to breeding of hexaploid wheat. Breed Sci. 2011;61:347-57.

115. Climate technology center & network. https://www.ctc-n.org/products/drought-tolerant-ryegrass.

116. Leslie C, McGranahan G. Micropropagation of Persian walnut (*Juglans regia* L.). High-Tech and Micropropagation II. Springer; 1992. p. 136-50.

117. Lahay RR, Misrun S, Sipayung R, editors. The storage capacity of cocoa seeds (*Theobroma cacao* L.) through giving Polyethylene Glycol (PEG) in the various of storage container. Iop Conference Series: Earth & Environmental Science; 2018.

118. Xiao LH, Yang G, Zhang LC, Yang XH, Zhao S, Ji ZZ et al. The resurrection genome of *Boea hygrometrica*: A blueprint for survival of dehydration. Proc Natl Acad Sci USA. 2015;112:5833-7.

119. Pirzadah TB, Malik B, Tahir I, Rehman Ru. Vivipary in *Fagopyrum esculentum*. Folia Biol Geol. 2016;57:41-3.

120. Missouri Botanical Garden Staff. Missouri Botanical Garden Plant Finder. 2017. http://www.missouribotanicalgarden.org/plantfinder/plantfindersearch.aspx.

121. Iqbal M, Khan MA, Naeem M, Aziz U, Afzal J, Latif M. Inducing drought tolerance in upland cotton (*Gossypium hirsutum* L.), accomplishments and future prospects. World Appl Sci J. 2013;21:1062-9.

122. Bakker J. Seeds, ecology, biogeography and evolution of dormancy, and germination. C.C. Baskin & J.M. Baskin. Plant Ecol. 2001;152:204-5.

123. Charloq, Yazid A, Gustiansya PP. Viability of shelled rubber (*Hevea brasiliensis* Muell. Arg.) seed treated with PEG 6000 at different drying times in storage period. J Agron. 2017;16:83-6.

124. Berjak P, Pammenter N. From *Avicennia* to *Zizania*: seed recalcitrance in perspective. Ann Bot. 2007;101:213-28.

125. Brys R, Jacquemyn H. Biological flora of the British Isles: *Primula veris* L. J Ecol. 2009;97:581-600.

126. Zhang Y-b, Tang W, Wang L-h, Hu Y-w, Lui X-w, Lui Y-s. Kiwifruit (*Actinidia chinensis*) R1R2R3-MYB transcription factor *AcMYB3R* enhances drought and salinity tolerance in *Arabidopsis thaliana*. J Integr Agr. 2019;18:417-27.

127. Xu Z, Xin T, Bartels D, Li Y, Gu W, Yao H et al. Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. Mol Plant. 2018;11:983-94.

128. White campion. https://www.gardenorganic.org.uk/weeds/white-campion.

129. Georgieva K, Maslenkova L. Drought induced changes in PS II activity in leaves from desiccation-sensitive plant *Spinacia oleracea* and desiccation-tolerant 'resurrection' fern *Polypodium polypodioides*. CR Acad Bulgare Sci. 2001;54:2:67.

130. Aversano R, Contaldi F, Ercolano MR, Grosso V, Iorizzo M, Tatino F et al. The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell. 2015;27:954-68.

131. Pino MT, Ávila A, Molina A, Jeknic Z, Chen TH. Enhanced in vitro drought tolerance of *Solanum tuberosum* and *Solanum commersonii* plants overexpressing the *ScCBF1* gene. Cien Inv Agr. 2013;40:171-84.

132. Egea I, Albaladejo I, Meco V, Morales B, Sevilla A, Bolarin MC et al. The drought-tolerant *Solanum pennellii* regulates leaf water loss and induces genes involved in amino acid and ethylene/jasmonate metabolism under dehydration. Sci Rep. 2018;8:2791.

133. Tani E, Kizis D, Markellou E, Papadakis I, Tsamadia D, Leventis G et al. Cultivar-dependent responses of eggplant (*Solanum melongena* L.) to simultaneous *Verticillium dahliae* infection and drought. Front Plant Sci. 2018;9:1181.

134. Craufurd PQ, Peacock JM. Effect of heat and drought stress on sorghum (*Sorghum Bicolor*). II. grain yield. Exp Agric. 1993;29:77.

135. Aslam M, Maqbool MA, Ceng R. Drought stress in maize (*Zea mays* L.). Springerbriefs in Agriculture. 2015.

136. Hoog SD, Zalar P, Van Den Ende BG, Gunde-Cimerman N. Relation of halotolerance to human-pathogenicity in the fungal tree of life: an overview of ecology and evolution under stress. In: Gunde-

Cimerman N, Oren A, Plemenitaš A, editors. Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya. Cellular Origin, Life in Extreme Habitats and Astrobiology. Dordrecht: Springer; 2005.

137. Bagagli E, Theodoro RC, Bosco SMG, McEwen JG. *Paracoccidioides brasiliensis*: phylogenetic and ecological aspects. Mycopathologia. 2008;165:197-207.

138. Son H, Lee J, Lee Y-W. Mannitol induces the conversion of conidia to chlamydospore-like structures that confer enhanced tolerance to heat, drought, and UV in *Gibberella zeae*. Microbiol Res. 2012;167:608-15.

139. Negi CS, Joshi P, Bohra S. Rapid vulnerability assessment of yartsa gunbu (*Ophiocordyceps sinensis* [Berk.] G.H. Sung et al) in Pithoragarh District, Uttarakhand State, India. Mt Res Dev. 2015;35:382-91.

140. Chaturvedi V, Defiglio H, Chaturvedi S. Phenotype profiling of white-nose syndrome pathogen *Pseudogymnoascus destructans* and closely-related *Pseudogymnoascus pannorum* reveals metabolic differences underlying fungal lifestyles. F1000Res. 2018;7:665.

Supplementary Files Legend

Additional file 1: Table S1. The NUMT/NUPT contents (contributed by organellar transcribed sequences) of desiccation-tolerant species and their controls. **Table S2.** The information used in assembling mitochondrial genomes. **Table S3.** The information used in assigning desiccation-tolerant species and their controls. **Table S4.** Genome information of the eukaryotes used in this study. **Table S5.** The NUMT/NUPT contents (contributed by organellar genomes) of desiccation-tolerant species and their controls. (XLSX 63 kb)

Additional file 2: Data S1. The mitochondrial genome sequences assembled in this study. (FASTA 1477 kb)

Additional file 3: Fig. S1. The distribution of the identities between NUMTs/NUPTs and their parental organellar sequences. (PDF 5930 kb)
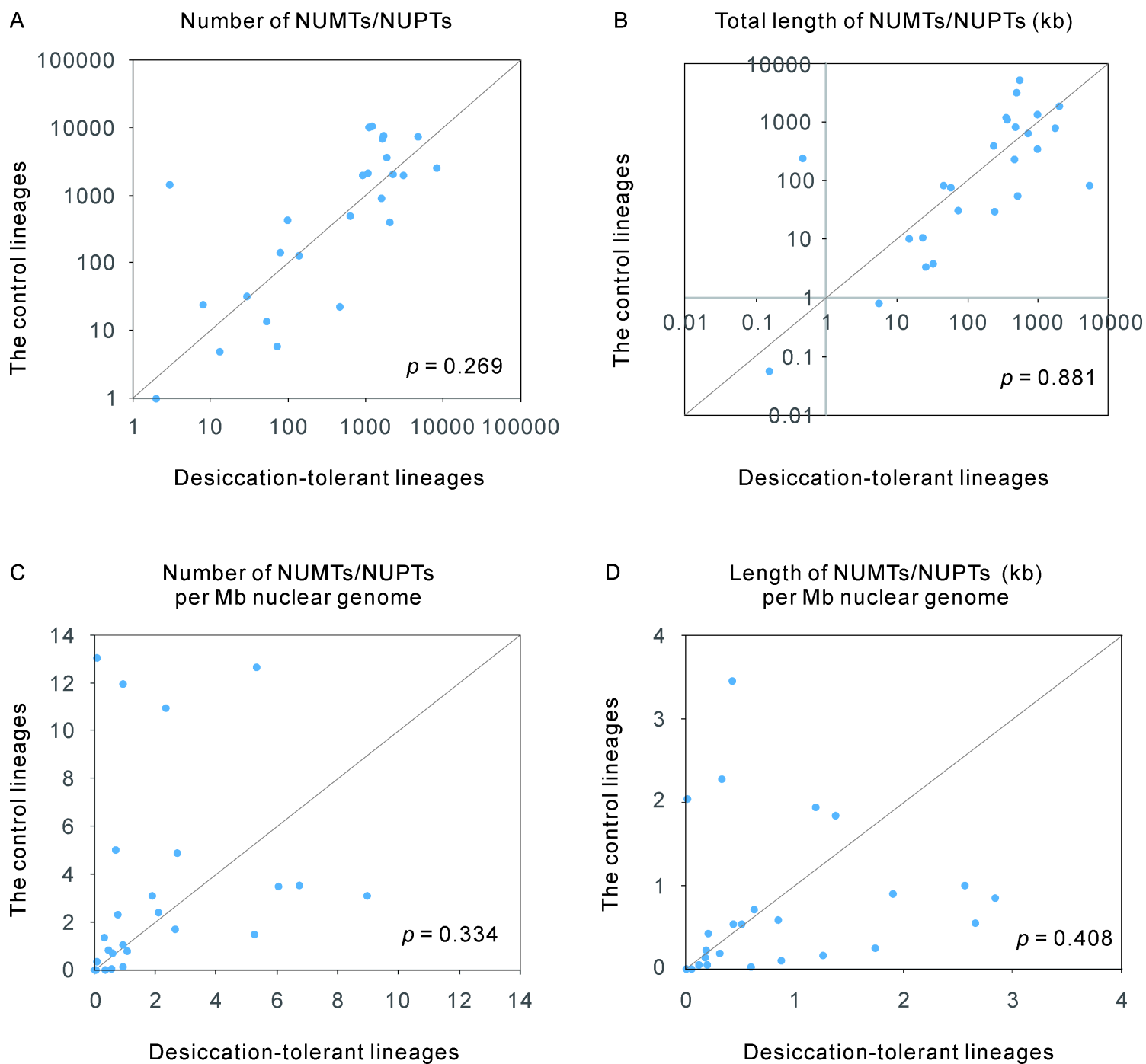
Figures

Figure 1

Comparison of the NUMT/NUPT contents between desiccation-tolerant organisms and their relatively less desiccation-tolerant relatives. The diagonal line represents cases in which desiccation-tolerant organisms and their paired relatives have the same NUMT/NUPT contents. Points above the line represent cases in which desiccation-tolerant organisms have lower NUMT/NUPT contents than their paired relatives, while points below the line indicate the reverse. All the p values were calculated using two-tailed Wilcoxon signed-rank tests. The exact values of the NUMT/NUPT contents are shown in Additional file 1: Table S5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- Additional2DataS1.fasta
- Additionalfile3.pdf