Plant Signaling & Behavior 8:4, e23853; April 2013; © 2013 Landes Bioscience

Do TE activity and counteracting genome defenses, RNAi and methylation, shape the sex lives of smut fungi?

John D. Laurie,[1,†] Rob Linning,[1] Philip Wong[2,‡] and Guus Bakkeren[1,*]

[1]Agriculture &Agri-Food Canada; Pacific Agri-Food Research Centre; Summerland, BC Canada; [2]Helmholtz Zentrum München; German Research Center for Environmental Health; Institute of Bioinformatics and Systems Biology; Neuherberg, Germany

[†]Current Affiliation: University of Nebraska-Lincoln; Center for Plant Science Innovation; Lincoln, NE USA; [‡]University Health Network; Toronto, ON Canada

Keywords: gene silencing, genome research, plant pathogen, *Sporisorium*, *Ustilago*

Abbreviations: LTR, long terminal repeat; MSUD, meiotic silencing unpaired DNA; RdRp, RNA-dependent RNA polymerase; RIP, repeat-induced point; TE, transposable element

The availability of three genomes from smut fungi differing in mating, TE load, and genome defense mechanisms, allowed a comparative analyses and a discussion on evolutionary forces shaping them. A complex balance of selective forces seems at play. A bipolar mating system in *Ustilago hordei* promotes selfing, advantageous for successful niche occupation but favoring accumulation of repetitive DNA, including TEs. TE activity may have caused genome variations necessary for these obligate parasites under high host selection pressures. Higher TE activity is balanced by genome defenses through recombination, RNAi, methylation and RIP mutagenesis. In tetrapolar *U. maydis*, lacking silencing and possibly methylation mechanisms, reduced inbreeding potential favors removal of repetitive DNA, presumably by its highly-efficient recombination system.

Closely-Related Smut Species vary Widely in TE Content

We recently reported on the generation and analysis of the genome sequence of the barley-infecting basidiomycete smut fungus, *Ustilago hordei*.[1] In comparing the genome of this fungus to those of two closely-related fungi, the corn-infecting smuts *U. maydis*[2] and *Sporisorium reilianum*,[3] we revealed several striking differences. The genome of *U. hordei* contained 5- to 12-times more transposable elements (TEs) and repeats than those of *U. maydis* and *S. reilianum*, respectively. LTR (long-terminal repeat) retrotransposon and LINE (long-interspersed element) sequences were responsible for the bulk of the difference. We also found in the *U. hordei* genome clear examples of RIP (repeat-induced point)-like mutations, a genome defense mechanism originally described for the ascomycete fungus *Neurospora crassa*[4] and later in basidiomycetes, including *Microbotryum violaceum*[5,6] which introduces mutations in repeats (often TEs) to inactivate them. Coincidentally, mating behavior differs between these smuts. Earlier work on smut fungi helped define genes involved in mating and revealed two loci as crucial determinants: the homeodomain-containing, transcription factor-coding *b*-gene complex locus and the pheromone and pheromone receptor-coding *a*-gene complex locus (reviewed in 7). These gene complexes reside on separate chromosomes in *U. maydis*[2] and *S. reilianum*[8] and therefore genetically constitute what has been called a tetrapolar mating system: two genetic loci, one, the *a* locus, with two specificities (for *U. maydis*, and at least 3 for *S. reilianum*), and the *b* locus, having at least 30 specificities in nature (for *U. maydis*). *U. hordei*, in contrast, possesses a bipolar mating mechanism represented by one mating-type locus with only two specificities in nature (*MAT-1* and *MAT-2*). Years ago, we showed that this locus had similar mating complexes as found in *U. maydis* but that in *MAT-1*, these were physically linked on the largest chromosome, 527 kb apart. This large chromosomal region, suppressed in recombination with its estimated 430 kb *MAT-2* counterpart and having over 50% of its DNA sequence coding for TEs and repeats, resembled a nascent sex chromosome.[9-11] In the recent comparative genome paper, we additionally concluded that a single chromosome fusion event led to the bipolar condition in *U. hordei*.[1] There was another intriguing finding: comparison to the *U. hordei* genome showed that *U. maydis* lacked several genes

*Correspondence to: Guus Bakkeren; Email: guus.bakkeren@agr.gc.ca
Submitted: 12/17/12; Revised: 01/31/13; Accepted: 02/01/13
http://dx.doi.org/10.4161/psb.23853
Citation: Laurie JD, Linning R, Wong P, Bakkeren G. Do TE activity and counteracting genome defenses, RNAi and methylation, shape the sex lives of smut fungi? Plant Signal Behav 2013; 8: e23853

involved in genome defense such as components of the RNAi-silencing pathway (dicer, argonaut and three RdRps) and DNA methylation (chromodomain-containing proteins and a cytosine 5-specific methyltransferase). Even more astounding was the fact that these genes seemed to be cleanly deleted from the *U. maydis* genome, leaving the flanking genes in conserved synteny compared with the other two genomes. Remnants of TEs and small footprints reminiscent of TE activity were found immediately flanking these deletion sites. The emerging data tempted us to speculate further about possible correlations between TEs and repeats, silencing and genome defense, and their possible effect on mating type evolution.

TEs as Modifiers of Genomic Information

TEs are classified into two broad categories depending on their mode of replication. Type I TEs or retro-elements replicate via an RNA intermediate and require a reverse transcriptase to form a DNA replicate. Type II TEs, however, require no RNA intermediate and replicate directly from DNA.[12] Hypotheses have been put forward that TEs (retroelements and noncoding RNAs) are former viral agents derived from an ancient RNA world that have become neutralized (balanced after "endogenization") in a "persistent symbiotic lifestyle" in the genome and are "natural genome editors" and "agents of (beneficial) change" by playing important roles in the regulation of (all) cellular processes in the host organism (reviewed in 13-16). Most eukaryotes have genomes with a significant proportion of TEs that consequently have hampered assembly efforts following whole genome sequencing. TEs, their remnants and other repetitive DNA sequences are often found in a special form of chromatin called heterochromatin that is repressed in both transcription and recombination. Additionally, heterochromatin is replicated separately during the cell cycle from the more active euchromatin, where the majority of genes are housed. Having large amounts of repetitive DNA that is actively processed into discrete chromatin structures and replicates separately, likely comes at a cost for cells. Indeed, TEs are traditionally known for having negative impacts on the genomes of their hosts. These include in addition to complications in chromosome replication and energy costs, gene deletions or inactivation through TE insertions and ectopic recombination leading to gene loss or more severely to chromosome instability. On the other hand, and more in tune with modern views emanating from the genomics era, TEs are more-and-more seen as "genetic settlers" that bring about genetic and genomic innovations such as altered gene expression through their influence on neighboring genes, post-transcriptional RNA processing, translation control, epigenetic control, etc., that are beneficial to organisms by providing adaptations on which selective forces can act.[12,13,15-19] A non-exhaustive list of 30 documented examples by which a particular retroelement in humans can exert functional activity is given by von Sternberg and Shapiro.[14,20] Adaptations can be on a relatively short time scale such as in pathogens having to overcome host defense mechanisms as an immediate selection pressure (see below). It has been argued that under such specific conditions and for organisms with large genomes the cost of housing potential detrimental TEs is outweighed by their beneficial contributions.[19,21,22]

Control of TE Function and Activities

Unbridled proliferation of TEs in an organism though, is detrimental and is necessarily balanced by host immune mechanisms evolved to defend the genome and remove, inactivate or control competing TEs (and their modulating functions). Mating systems have been described affecting the spread and persistence of TE load in genomes.[23] One of the main factors believed to have favored sexual over asexual reproduction during evolution besides the generation of reshuffled alleles on which natural selection can act, is the removal of detrimental DNA, i.e., deleterious mutations including TEs, from a population (Fisher-Müller hypothesis).[24,25] At the same time, sex can accelerate dispersal of TEs in a population. Means to modulate activities (inactivate) TEs in organisms exist as part of their genome defense capabilities, include silencing mechanisms[26] (see below). An alternative perspective is that TEs evolve and can accumulate in genomes because of silencing mechanisms.[16,19] This perspective implies that the more variable and complex silencing strategies an organism has, the more likely it is to house numerous TEs. Consequently, TEs evolve together with genome silencing mechanisms, allowing genomes to expand. However, true to the modern view that repeat elements (including TEs) are integral functional components of the genome and not merely passive parasites and a dead weight to organisms, it is becoming clear that TEs can counteract host genome defense mechanisms, thus balancing the negative and positive consequences of their propagation. For example, in a fashion similar to viruses interfering with their silencing by their host, recent exciting work in rice demonstrates that miRNAs produced by TEs can interfere with the host epigenetic TE methylation/silencing mechanism.[27] It is likely the "balancing act," a deliberate (evolved) "sloppiness" of the self-nonself recognition mechanism (of introduced or transposed DNA/RNA molecules), that allows TEs to modify genomes and bring about (temporary) changes for selection forces to act on.

TEs in Fungi

Fungi represent good models for studying TEs because a large variability is found among sequenced genomes that either contain large amounts of TEs and repeats (e.g., 28,29), a moderate number,[1] or are nearly devoid of TEs (e.g., 2,3,30). With short generation times, both sexual and asexual representatives, and often easy handling as (molecular genetic) experimental systems, consequences of TE activity are therefore easier to follow at both the individual and the species level in an evolutionary context. Studying TEs in relation to evolutionary time is important since effects on genome diversity become more apparent.[12] From such a perspective, a role is seen for TEs in creating variability upon which selection can act. When species encounter changing or hostile environments, stress often leads to TE activation implying a special role for TEs under such conditions. For example, recent

genome sequences from plant-pathogenic oomycetes (fungus-like organisms) have revealed a role for TE activity in creating variability among effector proteins which are crucial during plant colonization.[22,31-33] Similarly, in plant pathogenic fungi, TE activity has been involved in overcoming evolutionary and environmental (host) pressures imposed by resistance genes.[34-37] It could be argued that evolutionary pressures are different in pathogen-host systems, thereby skewing existing theories and population-genetic models.

Silence and RIP'em

Fungi can possess both type I and type II TEs that can contribute differently to total genome size. Certain fungi have become model organisms for genetics and molecular biology and in particular, fungi have been useful in furthering our understanding of how TEs are controlled. Pioneering work on quelling in *N. crassa* helped support parallel work in RNA silencing in plants and animals. Key genes, including *dicer*, *argonaute* and *RdRp*, were identified showing involvement in a conserved eukaryotic defense mechanism acting to repress expression of duplicated sequences, including TEs (reviewed in 38). The RIP mutation mechanism has already been mentioned and an important aspect of RIP is that it works on duplicated sequences during the dikaryotic stage, which is after mating in *Neurospora*, but before nuclear fusion and meiosis, leading during sexual reproduction to the accumulation of more-and-more mutations in repetitive loci at each pass.[39] Also in *N. crassa*, components of the RNA silencing machinery were shown to act in silencing unpaired DNA during meiosis (MSUD).[40] Recent work in *Cryptococcus* showed that RNA silencing plays an important role in controlling TEs during sexual reproduction[41,42] but clearly also controls duplicated genes during vegetative growth.[43]

Experiments in fission yeast have provided valuable information on how TEs are packaged into heterochromatin and how this packaging is important for proper centromere functioning; centromeres often accumulate TEs (e.g., 44). Most notable is how heterochromatin is transcribed during the S-phase of the cell cycle and acted upon by the RNA silencing machinery to recruit repressive chromatin complexes including histone H3 lysine 9 methylation and HP1-like chromodomain proteins.[45,46] For many species, including many fungi, DNA methylation is also an important means to regulate gene expression including controlling TEs (reviewed in 47). Methylation of cytosine residues is often directed to repetitive DNA including TEs, especially in the CpG context, in a variety of organisms including in fungi.[48-52] Interestingly, in several fungi from diverse taxa, CpG dinucleotides are also targeted for mutation in repetitive DNA, such as the C-to-T transitions seen in RIP mutations. The consequence of C-to-T transitions is a general depletion of C/G nucleotides and this hallmark RIP mutation signature has now been seen in diverse taxa from both ascomycete and basidiomycete fungi, including in the *U. hordei* genome.[1,6,51] An essential component of the RIP mutation machinery is RID, a protein with homology to DNA methyltransferases. It has been suggested that the RIP mutation mechanism may have evolved from a preexisting DNA methylation pathway that recognizes CpG dinucleotides to actively mutate TEs upon their replication.[53] Collectively, the many studies on model fungi point to a need for mutational inactivation and control of TEs in genome defense during sexual reproduction.

Fungal Mating Systems, TEs and Recombination

Fungi are also excellent models for studying various modes of reproduction, representing both asexuals and sexuals. It should be noted that species typed as asexual may have a cryptic sexual cycle, not yet discovered. Indeed, in some species, recent sequenced genomes have revealed complete mating-related genes suggesting they are capable of sex (for example in the dandruff-causing fungus, *Malassezia globosa*[54]). The variation in sexual reproduction is astounding, encompassing homo- and heterothallic, selfing and outcrossing species with variable degrees of inbreeding (among both haploids or diploids) or outcrossing among them; the various modes are sometimes even found among closely-related species.[25,55-57] For sexual species, mating compatibility is governed by specific genetic loci and in its simplest form involved a transcription factor essential for regulating mating. Repressed recombination coinciding with a TE eventually led to sequence divergence and the formation of mating types; in the basidiomycete lineages, recruiting a separate self-nonself recognition system as a second genetic locus occurred (reviewed in 25,58,59). Interesting experimental data from the yeast, *Kluyveromyces lactis*, corroborates the notion that a TE could have evolved into or have become part of a functional control module to promote host sexual reproduction.[60] In a recent study, TE activity was implicated in the mating-type rearrangements that occurred during evolution in *Neurospora* species and which led in several independent occasions to the generation of homothallism ('same-clone mating system') from likely ancestral heterothallism (which prevents 'same-clone mating').[61] Similarly, in the basidiomycete lineage, rearrangements of mating-type loci to form large nascent sex chromosomes, which leads to increased selfing potential, have been described for multiple species and in all cases, repeats and (LTR) TEs were implicated.[1,9,10,62,63] TEs, therefore, are associated with steps in the evolution of sexual types.

...in *U. hordei*

Among the sequenced smut genomes, there was a difference in TE load and diversity and this seemed to be correlated with the bipolar and tetrapolar mating systems: both tetrapolar species *U. maydis* and *S. reilianum* were relatively devoid of TEs (2% and 0.8%, respectively), while the bipolar species *U. hordei* had significantly larger amounts of various types of both DNA and retroelements (up to 10%, with ~50% in the 527-kb *MAT-1* region suggested to be caused by recombination suppression). In addition, the TEs were mostly species-specific with a few very large 'families' of related elements indicating recent invasion and proliferation.[1] Of the retroelements, two groups of highly similar elements (Ty1/Copia-like and Ty3/Gypsy-like) were found scattered throughout the genome of only *U. hordei*. We cannot determine a

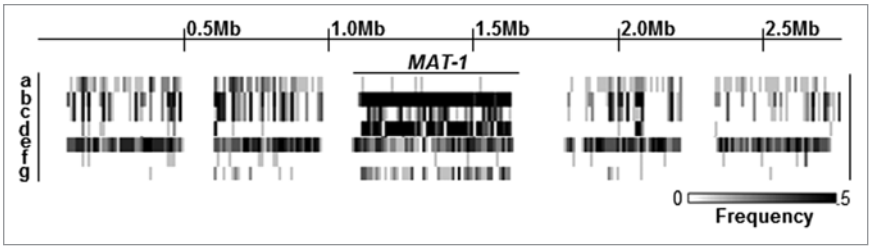

Figure 1. Heat map revealing distribution of various features on *U. hordei* chromosome I. Rows represent the frequency of the various elements as they occur in a 10 kb region of sequence, indicated by one block; the gray scale ranges from zero frequency (white) to 5 or more occurrences (black). The mating-type locus *MAT-1* is indicated. The row designations are: (a) Small sequencing gaps within supercontigs, (b) all repetitive and transposable element sequences combined, with the two largest groups: (c) a family of 2752 variants of LTR2, LTR5, LTR13, and composite TE, Tuh3 and Tuh5 elements that are related to the Ty1/copia-type elements previously identified in the *U. hordei MAT-1* mating-type region,[11] and (d) a family of 1377 variants of LTR1, LTR3, LTR6, LTR7, LTR8, LTR10, LTR12 and Tuh1-related sequences, which belong to a Ty3/Gypsy class. (e) All gene models predicted on the indicated contigs, (f) genes encoding predicted small candidate secreted effector proteins (CSEPs), (g) CG-depleted genes (CG LogRatio [observed/expected] < -1.0). See reference 1 for details.

likely origin of these elements. For example, a BLASTn search at the NCBI nr database of a representative Ty1/Copia-type element sequence did not return significant hits, although a BLASTx search matched gag-pol elements from Arabidopsis (e-150) to rice (e-131), but all having just only 30–35% amino acid identities. Obviously, gag-pol ORFs (proteins), originating from common ancestral viral elements, are widely maintained and relatively conserved in nature. However, a Ty3/Gypsy class representative sequence did not match any sequences, either by BLASTx or BLASTn. Since many of these *U. hordei* elements seem to be unique in the currently-available sequenced smut fungal genomes, many more sequenced genomes of related species need to become available (possibly identifying an ancestral genome) to solve phylogenies. These species-specific elements likely contain lineage-specific, self/nonself differentiation competencies allowing modulating effects only within the species.[16]

Perhaps the most interesting of the *U. hordei* retroelements was the Gypsy-like class that showed the highest density in the *MAT-1* locus (**Fig. 1**, row d). Gypsy-like elements are special in that they contain a chromodomain in the C-terminal region of the polyprotein that they encode.[64] It is thought that the chromodomain helps the Gypsy-like elements recognize heterochromatic DNA prior to insertion; the chromodomain interacts with H3K9 methylated histones, which are recruited to and are hallmarks of heterochromatin.[65] *U. hordei* Gypsy polyproteins contain a chromodomain at the C-terminal end (**Fig. 2**), but no such Gypsy-like elements or chromodomains were identified in the *S. reilianum* or *U. maydis* genomes (some weak similarity was found with the integrase domain in an *U. maydis* HobS element). The fact that Gypsy-like elements are enriched at the *U. hordei MAT-1* locus and that a low recombination frequency has been observed at this locus,[10] suggests that the locus is largely heterochromatic. A key question for the relationship of the Gypsy-like elements to the *U. hordei MAT-1* locus is whether the elements played a role in the original rearrangement(s)/translocation events that formed the nascent sex chromosome, or accumulated subsequently as recombination was suppressed, impeding purging them, or both. An interesting consequence of the formation of the *MAT* locus is that the resulting bipolar mating system yields progeny with a higher probability of diploid selfing than progeny produced through the tetrapolar system.[57] In bipolars, one out of two random progeny can mate with any other random progeny from a cross whereas one out of four can do this in tetrapolars; within the population at large, due to the bi-allelic pheromone-receptor locus, the success rate of compatibility should theoretically be 50% in both bipolar and tetrapolar species, thus not promoting outcrossing by one system over the other.[25] Of course, the extent of outcrossing in general depends on the availability of unrelated teliospores germinating in close proximity, which depends on disease severity. In any event, more frequent selfing in bipolars would lead to increased homozygosity and reduced ectopic recombination between non-syntenic repetitive elements. This would have been beneficial to *U. hordei* through an increase in genome stability and would have protected Gypsy-like elements from the purging process. Formation of the bipolar *MAT* locus therefore would have maintained more adaptive potential such as the Gypsy-like elements constituting a selective advantage over the tetrapolar mating system.

Selective Pressure from a Host Environment

Even though increased selfing might be advantageous in the short-term for populations occupying successfully a certain niche, it usually drives populations under (strong) selection pressure slowly to extinction likely because it prevents rapid adaptation.[66] This is especially true in (fast) co-evolving pathogen-host populations under higher selection pressures and experiments aimed at supporting the "Red Queen Hypothesis" (which posits that in an evolutionary system, continuing adaptation is needed in order for a species to maintain its relative fitness among the systems co-evolving) do not favor self-fertilizing mating systems.[67] For *U. hordei*, selective forces would have favored the presence of a compensating mechanism to create the diversity needed to adapt to the ever-changing, hostile host environment and this could be increased activity (and maintenance) of TEs. Indeed, in population genetic studies, when the effective population size decreases (e.g., due to mating restrictions or other bottleneck), TE insertions are found in higher frequencies.[12] In *Ustilago*, and other obligate parasites, the hostile host environment can be considered a bottleneck, leading to smaller effective population sizes.

RIP and Genome Defense in *U. hordei*

In the smuts, the observed differences in TE load between bipolar and tetrapolar species is likely resulting from a combination of mating system and genome defense strategies. In *U. hordei*,

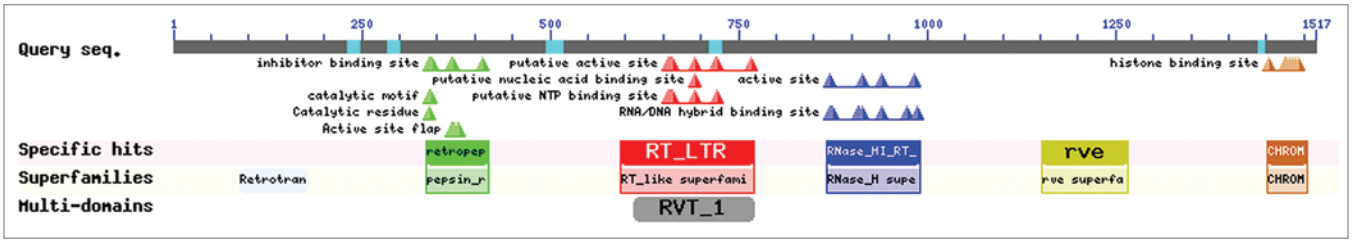

Figure 2. Example of a Gypsy-like TE element UHOR_14949 with specific protein domains. In the *U. hordei* genome, 44 homologs are identified using BLASTx with corresponding e-values varying from e-122 to e-10 (http://mips.helmholtz-muenchen.de/genre/proj/MUHDB). Eleven of these homologs are found in the 527 kb *MAT-1* region spanning the *a* and *b* mating-type loci. Note the histone binding chromo domain at the C-terminal end (annotated by the Conserved Domains Database at NCBI).

bipolarity favors selfing which would generally increase homozygosity. However, outcrossing is feasible and could introduce new elements, which could be purged through ectopic recombination in subsequent sexual cycles. The Gypsy-like elements may help enforce this bipolar mating system by helping the large *MAT* locus to remain heterochromatic and suppressed in recombination. TEs abundant within the *MAT* locus and elsewhere in the genome are then controlled by various means. A functional RNA silencing pathway likely plays an important role in controlling the abundant TEs and may work to enforce heterochromatic silencing during the S-phase of the mitotic cell cycle as in fission yeast or during sexual reproduction as in *C. neoformans*. As well, cytosine DNA methylation may also help control the TEs in *U. hordei*.

Our genome study identified a RIP-like mutation mechanism and a resultant observed depletion in C/G nucleotides, only in *U. hordei*. When all genes and TEs were analyzed, it was evident that the CG depletion was found primarily in TE sequences (see the Supplemental Figure 7 in ref. 1). Here we report on an expanded genome analysis. Calculations yielded a CG LogRatio observed/expected < -1.0 for 1,883 out of the 4,910 identified TEs, but a slight decrease in C/G nucleotides could also be observed in a subset of the *U. hordei* gene models. Further analysis of these revealed that 495 out of the predicted 7,043 *U. hordei* genes had a CG LogRatio observed/expected < -1.0 indicating a depletion whereas this was found for only 13 *S. reilianum* and none of the *U. maydis* genes (**Fig. 3**). When comparing these 495 *U. hordei* genes by BLASTn to each other, at least 96 genes appeared to match duplicated sequences over at least 50% of their length, mostly present in two copies in the genome but also some representing small gene families; these do not match TEs or repeats and code mostly for genes of unknown functions. When comparing these 495 genes to the collection of *U. hordei* TEs and repeat sequences, 378 genes (76%) matched TE and repeat sequences (at $e < 10^{-20}$) of which 253 (51%) represent the two largest TE families found in the genome: 183 are related to the Ty1/copia-type elements (**Fig. 1**, row c) and 70 to the Ty3/Gypsy class elements (**Fig. 1**, row d). Overall, approximately 200 genes match to TE sequences over at least 25% of their lengths, either at the 5'- or 3'-ends. These are likely chimeric genes and could represent TE insertions. When aligned, clear RIP mutation signatures are observed, including in the duplicated, non-TE genes (as in 1; data not shown).

We mapped these CG-depleted *U. hordei* genes on the genome and found clear groupings. An extreme example is found on chromosome I where a dense clustering can be seen at the *MAT-1* locus. Within the 527 kb *MAT-1* region delimited by the *a* and *b* mating-type loci and in which 142 gene models have been called,[11] 73 genes have a calculated CG LogRatio observed/expected < -1.0 (**Fig. 1**, row g). This suggests that gene duplication is common and that the RIP mutation machinery seems highly active at this locus. Although genes revealing a high degree of RIP mutations are common at the *MAT-1* locus, often in stretches of 3 to 5 genes (a stretch next to the *b* locus has 10), several genes showing no sign of RIP mutations, are frequently interspersed. Interestingly, these genes remained conserved among the smuts, even after the proposed chromosome fusion event in *U. hordei*, and this finding confirms that RIP-like mutations do not extend beyond duplicated sequences.

Can TE "Modifiers" Influence Mating Systems?

As mentioned, some of the observed duplications involve gene fragments fused to TE-related sequences. Recently, high expression of Gypsy-like elements was observed during ovule development in sexual but not apomictic (asexual) genotypes of plants.[68] More specifically, expression of entire chimeric Gypsy-like element-gene fusions occurred and it was noted that several of the gene fragments corresponded to genes known to play a role in apomixis. The authors suggested that an autocatalytic process could be at play in sexual genotypes where gene fragments fused to Gypsy-like elements help suppress apomictic development. Given that RNA silencing plays an important role in controlling TEs during sexual reproduction, one could speculate that TEs may coincidentally be active. Perhaps the Gypsy-like elements in *U. hordei*, upon which RNA silencing (RIP) could act, help enforce a particular sexual behavior by modulating expression of such TE-gene fusions that are involved in sexual reproduction (at the *MAT* locus). If the modulation is regulated through the TE part, this could potentially allow the Gypsy-like elements control over the nature of sex of its host. As hypothesized above, the bipolar mating system may constitute a selective advantage over the tetrapolar system and the changes that led to this system undoubtedly involved TEs. It has been shown that beneficial mutations can arise not only randomly but also through directed or adaptive mutagenesis, oftentimes due to stress and

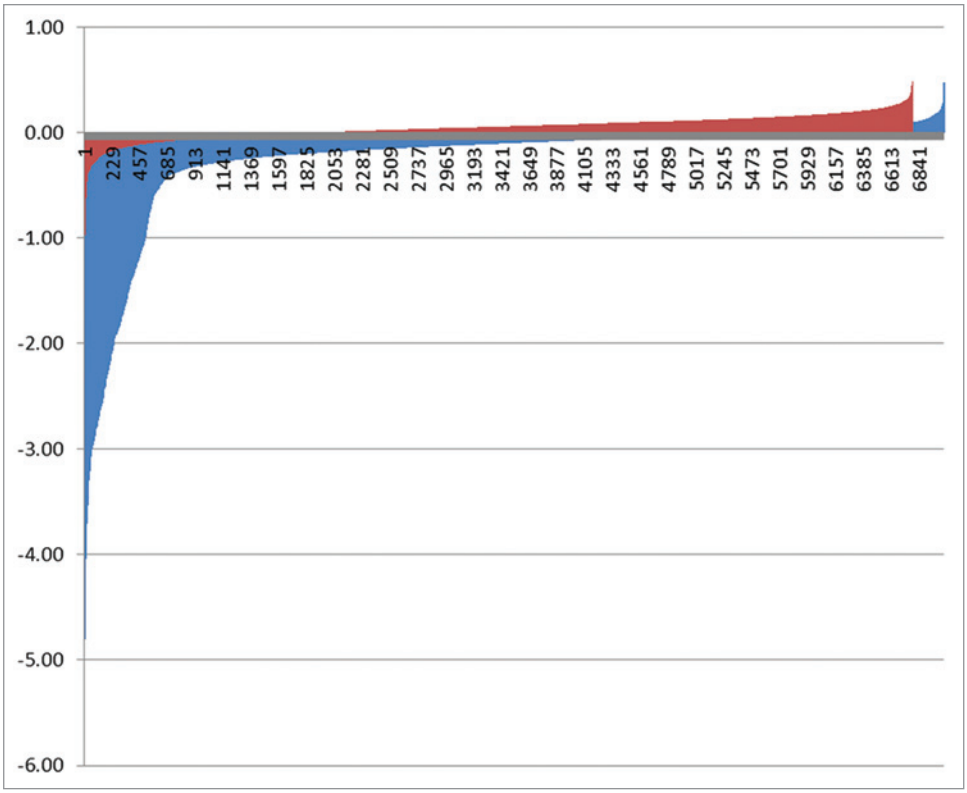

Figure 3. CG-depletion in all gene models predicted in the genomes of *U. hordei* and *U. maydis*. The CG LogRatio [observed/expected] is plotted on the Y-axis. In *U. hordei*, 495 out of the 7,043 genes (in blue) have a CG LogRatio [observed/expected] < -1.0 which we chose as cut-off value. *S. reilianum* has 13, but *U. maydis* (in red) has no such CG-depleted genes.

environmental pressures on the phenotype, e.g.[69,70] An in-depth analysis of the gene fragments fused to TEs is needed to confirm this possibility.

Conclusions

The lack of TEs and repeats in certain smuts is likely a consequence of an extremely efficient repetitive DNA purging mechanism. The near-absence of repetitive DNA in *U. maydis* might have resulted from the loss of RNA silencing and possibly the RIP mutation mechanism (if present in a progenitor of *U. maydis*). Indeed, disruption of the RNAi machinery in *S. pombe* negatively affects the repair of double-strand DNA breaks and increases homologous recombination.[71] Homologous recombination is highly efficient in *U. maydis* and is likely involved in genome maintenance. As well, *U. maydis* can survive extremely high doses of ionizing radiation that would be lethal to other eukaryotes. It therefore seems that *U. maydis* has an efficient mechanism to repair double-stranded DNA breaks. An interesting feature of the *U. maydis* genome is the high frequency of short, 10-bp repeats that populate intergenic sequences and were proposed to be footprints remaining from the purging of TEs.[1,3] Collectively, the short repeats, the ability to survive DNA damaging radiation and the efficient homologous recombination point to the existence of a very effective (alternative) genome defense and maintenance mechanism in light of the lack of certain epigenetic control mechanism(s). The mechanisms at play in *S. reilianum* remain to be determined; silencing pathway components were present in its genome but it is currently unknown whether these are functional.[3] This tetrapolar smut could very well represent an intermediate between *U. hordei* and *U. maydis* but is different in that multiple pheromone-receptor alleles evolved thereby increasing its outbreeding potential to more than 50%.[8] Active elimination of TEs may be achieved through increased outcrossing and recombination geared toward purging of repetitive DNA. The loss of RNA silencing genes in *U. maydis*, possibly through the action of TEs followed by recombination, may have created a selective advantage in this species. In this species, variability (among effectors), thought essential and caused by TEs in some other plant pathogens, may have been achieved through recombination or other means. Indeed, with the emerging view that the TE and repeat content, previously thought of as "selfish junk DNA" in organisms, actually represents integral functional components of the 'genome' (defined as the complex, informational 'organelle' which includes multilevel epigenetic control[14]), it is interesting to learn how organisms nearly devoid of them achieve and maintain such functional diversity. In conclusion, the smuts represent excellent models for studying the role of TEs, genome defenses and mating systems in creating adaptive potential. A tantalizing hypothesis that TEs have influenced some of these processes, including sexual behavior, needs testing. We are eagerly awaiting the sequencing of other smut genomes to help delineate the various proposed forces shaping them.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank Drs. T. Giraud and J.W. Kronstad for comments on the manuscript, and an anonymous reviewer for pointing out insightful literature.

References

1. Laurie JD, Ali S, Linning R, Mannhaupt G, Wong P, Güldener U, et al. Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. Plant Cell 2012; 24:1733-45; PMID:22623492; http://dx.doi.org/10.1105/tpc.112.097261.
2. Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis.* Nature 2006; 444:97-101; PMID:17080091; http://dx.doi.org/10.1038/nature05248.
3. Schirawski J, Mannhaupt G, Münch K, Brefort T, Schipper K, Doehlemann G, et al. Pathogenicity determinants in smut fungi revealed by genome comparison. Science 2010; 330:1546-8; PMID:21148393; http://dx.doi.org/10.1126/science.1195330.
4. Grayburn WS, Selker EU. A natural case of RIP: degeneration of the DNA sequence in an ancestral tandem duplication. Mol Cell Biol 1989; 9:4416-21; PMID:2531278.
5. Hood ME, Katawczik M, Giraud T. Repeat-induced point mutation and the population structure of transposable elements in *Microbotryum violaceum.* Genetics 2005; 170:1081-9; PMID:15911572; http://dx.doi.org/10.1534/genetics.105.042564.
6. Horns F, Petit E, Yockteng R, Hood ME. Patterns of repeat-induced point mutation in transposable elements of basidiomycete fungi. Genome Biol Evol 2012; 4:240-7; PMID:22250128; http://dx.doi.org/10.1093/gbe/evs005.
7. Brefort T, Doehlemann G, Mendoza-Mendoza A, Reissmann S, Djamei A, Kahmann R. *Ustilago maydis* as a Pathogen. Annu Rev Phytopathol 2009; 47:423-45; PMID:19400641; http://dx.doi.org/10.1146/annurev-phyto-080508-081923.
8. Schirawski J, Heinze B, Wagenknecht M, Kahmann R. Mating type loci of *Sporisorium reilianum*: novel pattern with three *a* and multiple *b* specificities. Eukaryot Cell 2005; 4:1317-27; PMID:16087737; http://dx.doi.org/10.1128/EC.4.8.1317-1327.2005.
9. Bakkeren G, Kronstad JW. Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. Proc Natl Acad Sci USA 1994; 91:7085-9; PMID:7913746; http://dx.doi.org/10.1073/pnas.91.15.7085.
10. Lee N, Bakkeren G, Wong K, Sherwood JE, Kronstad JW. The mating-type and pathogenicity locus of the fungus *Ustilago hordei* spans a 500-kb region. Proc Natl Acad Sci USA 1999; 96:15026-31; PMID:10611332; http://dx.doi.org/10.1073/pnas.96.26.15026.
11. Bakkeren G, Jiang G, Warren RL, Butterfield Y, Shin H, Chiu R, et al. Mating factor linkage and genome evolution in basidiomycetous pathogens of cereals. Fungal Genet Biol 2006; 43:655-66; PMID:16793293; http://dx.doi.org/10.1016/j.fgb.2006.04.002.
12. Tollis M, Boissinot S. The evolutionary dynamics of transposable elements in eukaryote genomes. Genome Dyn 2010; 7:68-91; PMID:22759814; http://dx.doi.org/10.1159/000337126.
13. von Sternberg R. On the roles of repetitive DNA elements in the context of a unified genomic-epigenetic system. Ann N Y Acad Sci 2002; 981:154-88; PMID:12547679; http://dx.doi.org/10.1111/j.1749-6632.2002.tb04917.x.
14. Shapiro JA, von Sternberg R. Why repetitive DNA is essential to genome function. Biol Rev Camb Philos Soc 2005; 80:227-50; PMID:15921050; http://dx.doi.org/10.1017/S1464793104006657.
15. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007; 8:272-85; PMID:17363976; http://dx.doi.org/10.1038/nrg2072.
16. Witzany G. Noncoding RNAs: persistent viral agents as modular tools for cellular needs. Ann N Y Acad Sci 2009; 1178:244-67; PMID:19845641; http://dx.doi.org/10.1111/j.1749-6632.2009.04989.x.
17. Britten RJ. Transposable element insertions have strongly affected human evolution. Proc Natl Acad Sci USA 2010; 107:19945-8; PMID:21041622; http://dx.doi.org/10.1073/pnas.1014330107.
18. Hua-Van A, Le Rouzic A, Boutin TS, Filée J, Capy P. The struggle for life of the genome's selfish architects. Biol Direct 2011; 6:19; PMID:21414203; http://dx.doi.org/10.1186/1745-6150-6-19.
19. Fedoroff NV. Presidential address. Transposable elements, epigenetics, and genome evolution. Science 2012; 338:758-67; PMID:23145453; http://dx.doi.org/10.1126/science.338.6108.758.
20. von Sternberg R, Shapiro JA. How repeated retroelements format genome function. Cytogenet Genome Res 2005; 110:108-16; PMID:16093662; http://dx.doi.org/10.1159/000084942.
21. Spanu PD. The genomics of obligate (and nonobligate) biotrophs. Annu Rev Phytopathol 2012; 50:91-109; PMID:22559067; http://dx.doi.org/10.1146/annurev-phyto-081211-173024.
22. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol 2012; 10:417-30; PMID:22565130.
23. Boutin TS, Le Rouzic A, Capy P. How does selfing affect the dynamics of selfish transposable elements? Mob DNA 2012; 3:5; PMID:22394388; http://dx.doi.org/10.1186/1759-8753-3-5.
24. Goddard MR, Godfray HC, Burt A. Sex increases the efficacy of natural selection in experimental yeast populations. Nature 2005; 434:636-40; PMID:15800622; http://dx.doi.org/10.1038/nature03405.
25. Lee SC, Ni M, Li W, Shertz C, Heitman J. The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev 2010; 74:298-340; PMID:20508251; http://dx.doi.org/10.1128/MMBR.00005-10.
26. Blumenstiel JP. Evolutionary dynamics of transposable elements in a small RNA world. Trends Genet 2011; 27:23-31; PMID:21074888; http://dx.doi.org/10.1016/j.tig.2010.10.003.
27. Nosaka M, Itoh JI, Nagato Y, Ono A, Ishiwata A, Sato Y. Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice. PLoS Genet 2012; 8:e1002953; PMID:23028360; http://dx.doi.org/10.1371/journal.pgen.1002953.
28. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 2010; 330:1543-6; PMID:21148392; http://dx.doi.org/10.1126/science.1194573.
29. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci USA 2011; 108:9166-71; PMID:21536894; http://dx.doi.org/10.1073/pnas.1019315108.
30. Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. Genome Res 1998; 8:464-78; PMID:9582191.
31. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans.* Nature 2009; 461:393-8; PMID:19741609; http://dx.doi.org/10.1038/nature08358.
32. Hogenhout SA, Van der Hoorn RA, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-associated organisms. Mol Plant Microbe Interact 2009; 22:115-22; PMID:19132864; http://dx.doi.org/10.1094/MPMI-22-2-0115.
33. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, et al. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science 2010; 330:1549-51; PMID:21148394; http://dx.doi.org/10.1126/science.1195203.
34. Kang S, Lebrun MH, Farrall L, Valent B. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Mol Plant Microbe Interact 2001; 14:671-4; PMID:11332731; http://dx.doi.org/10.1094/MPMI.2001.14.5.671.
35. Thon MR, Pan H, Diener S, Papalas J, Taro A, Mitchell TK, et al. The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Magnaporthe oryzae.* Genome Biol 2006; 7:R16; PMID:16507177; http://dx.doi.org/10.1186/gb-2006-7-2-r16.
36. Sacristán S, Vigouroux M, Pedersen C, Skamnioti P, Thordal-Christensen H, Micali C, et al. Coevolution between a family of parasite virulence effectors and a class of LINE-1 retrotransposons. PLoS ONE 2009; 4:e7463; PMID:19829700; http://dx.doi.org/10.1371/journal.pone.0007463.
37. Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, et al. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. Nat Commun 2011; 2:202; PMID:21326234; http://dx.doi.org/10.1038/ncomms1189.
38. Fulci V, Macino G. Quelling: post-transcriptional gene silencing guided by small RNAs in *Neurospora crassa.* Curr Opin Microbiol 2007; 10:199-203; PMID:17395524; http://dx.doi.org/10.1016/j.mib.2007.03.016.
39. Selker EU, Cambareri EB, Jensen BC, Haack KR. Rearrangement of duplicated DNA in specialized cells of *Neurospora.* Cell 1987; 51:741-52; PMID:2960455; http://dx.doi.org/10.1016/0092-8674(87)90097-3.
40. Shiu PK, Raju NB, Zickler D, Metzenberg RL. Meiotic silencing by unpaired DNA. Cell 2001; 107:905-16; PMID:11779466; http://dx.doi.org/10.1016/S0092-8674(01)00609-2.
41. Janbon G, Maeng S, Yang DH, Ko YJ, Jung KW, Moyrand F, et al. Characterizing the role of RNA silencing components in *Cryptococcus neoformans.* Fungal Genet Biol 2010; 47:1070-80; PMID:21067947; http://dx.doi.org/10.1016/j.fgb.2010.10.005.
42. Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J. Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. Genes Dev 2010; 24:2566-82; PMID:21078820; http://dx.doi.org/10.1101/gad.1970910.
43. Wang X, Wang P, Sun S, Darwiche S, Idnurm A, Heitman J. Transgene induced co-suppression during vegetative growth in *Cryptococcus neoformans.* PLoS Genet 2012; 8:e1002885; PMID:22916030; http://dx.doi.org/10.1371/journal.pgen.1002885.
44. Pidoux AL, Allshire RC. The role of heterochromatin in centromere function. Philos Trans R Soc Lond B Biol Sci 2005; 360:569-79; PMID:15905142; http://dx.doi.org/10.1098/rstb.2004.1611.
45. Moazed D, Bühler M, Buker SM, Colmenares SU, Gerace EL, Gerber SA, et al. Studies on the mechanism of RNAi-dependent heterochromatin assembly. Cold Spring Harb Symp Quant Biol 2006; 71:461-71; PMID:17381328; http://dx.doi.org/10.1101/sqb.2006.71.044.
46. Moazed D. Small RNAs in transcriptional gene silencing and genome defence. Nature 2009; 457:413-20; PMID:19158787; http://dx.doi.org/10.1038/nature07756.
47. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 2010; 11:204-20; PMID:20142834; http://dx.doi.org/10.1038/nrg2719.
48. Wälti MA, Villalba C, Buser RM, Grünler A, Aebi M, Künzler M. Targeted gene silencing in the model mushroom *Coprinopsis cinerea* (*Coprinus cinereus*) by expression of homologous hairpin RNAs. Eukaryot Cell 2006; 5:732-44; PMID:16607020; http://dx.doi.org/10.1128/EC.5.4.732-744.2006.

49. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus *Neurospora crassa.* Nature 2003; 422:859-68; PMID:12712197; http://dx.doi.org/10.1038/nature01554.
50. Kemppainen M, Duplessis S, Martin F, Pardo AG. RNA silencing in the model mycorrhizal fungus *Laccaria bicolor*: gene knock-down of nitrate reductase results in inhibition of symbiosis with Populus. Environ Microbiol 2009; 11:1878-96; PMID:19397683; http://dx.doi.org/10.1111/j.1462-2920.2009.01912.x.
51. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 2010; 328:916-9; PMID:20395474; http://dx.doi.org/10.1126/science.1186366.
52. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, et al. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science 2012; 337:1360-4; PMID:22984074; http://dx.doi.org/10.1126/science.1224839.
53. Freitag M, Williams RL, Kothe GO, Selker EU. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa.* Proc Natl Acad Sci USA 2002; 99:8802-7; PMID:12072568; http://dx.doi.org/10.1073/pnas.132212899.
54. Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, et al. Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci USA 2007; 104:18730-5; PMID:18000048; http://dx.doi.org/10.1073/pnas.0706756104.
55. Giraud T, Yockteng R, López-Villavicencio M, Refrégier G, Hood ME. Mating system of the anther smut fungus *Microbotryum violaceum:* selfing under heterothallism. Eukaryot Cell 2008; 7:765-75; PMID:18281603; http://dx.doi.org/10.1128/EC.00440-07.
56. Whittle CA, Nygren K, Johannesson H. Consequences of reproductive mode on genome evolution in fungi. Fungal Genet Biol 2011; 48:661-7; PMID:21362492; http://dx.doi.org/10.1016/j.fgb.2011.02.005.
57. Billiard S, López-Villavicencio M, Hood ME, Giraud T. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. J Evol Biol 2012; 25:1020-38; PMID:22515640; http://dx.doi.org/10.1111/j.1420-9101.2012.02495.x.
58. Butler G. The evolution of MAT: the ascomycetes. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA, eds. Sex in fungi. Washington, D.C.: ASM Press, 2007:3-18.
59. Fraser JA, Hsueh YP, Findley KM, Heitman J. Evolution of the mating-type locus: the basidiomycetes. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA, eds. Sex in fungi. Washington, D.C.: ASM Press, 2007:19-34.
60. Barsoum E, Martinez P, Aström SU. Alpha3, a transposable element that promotes host sexual reproduction. Genes Dev 2010; 24:33-44; PMID:20008928; http://dx.doi.org/10.1101/gad.557310.
61. Gioti A, Mushegian AA, Strandberg R, Stajich JE, Johannesson H. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. Mol Biol Evol 2012; 29:3215-26; PMID:22593224; http://dx.doi.org/10.1093/molbev/mss132.
62. Lengeler KB, Fox DS, Fraser JA, Allen A, Forrester K, Dietrich FS, et al. Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. Eukaryot Cell 2002; 1:704-18; PMID:12455690; http://dx.doi.org/10.1128/EC.1.5.704-718.2002.
63. Hood ME, Antonovics J, Koskella B. Shared forces of sex chromosome evolution in haploid-mating and diploid-mating organisms: *Microbotryum violaceum* and other model organisms. Genetics 2004; 168:141-6; PMID:15454533; http://dx.doi.org/10.1534/genetics.104.029900.
64. Novikova O. Chromodomains and LTR retrotransposons in plants. Commun Integr Biol 2009; 2:158-62; PMID:19513271.
65. Gao X, Hou Y, Ebina H, Levin HL, Voytas DF. Chromodomains direct integration of retrotransposons to heterochromatin. Genome Res 2008; 18:359-69; PMID:18256242; http://dx.doi.org/10.1101/gr.7146408.
66. Charlesworth D, Charlesworth B. Inbreeding depression and its evolutionary consequences. Annu Rev Ecol Syst 1987; 18:237-68; http://dx.doi.org/10.1146/annurev.es.18.110187.001321.
67. Morran LT, Schmidt OG, Gelarden IA, Parrish RC 2nd, Lively CM. Running with the Red Queen: host-parasite coevolution selects for biparental sex. Science 2011; 333:216-8; PMID:21737739; http://dx.doi.org/10.1126/science.1206360.
68. Ochogavía AC, Seijo JG, González AM, Podio M, Duarte Silveira E, Machado Lacerda AL, et al. Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual *Paspalum notatum* plants. Sex Plant Reprod 2011; 24:231-46; PMID:21394488; http://dx.doi.org/10.1007/s00497-011-0165-0.
69. Lenski RE, Mittler JE. The directed mutation controversy and neo-Darwinism. Science 1993; 259:188-94; PMID:7678468; http://dx.doi.org/10.1126/science.7678468.
70. Hall BG. Adaptive mutagenesis: a process that generates almost exclusively beneficial mutations. Genetica 1998; 102-103:109-25; PMID:9720275; http://dx.doi.org/10.1023/A:1017015815643.
71. Zaratiegui M, Castel SE, Irvine DV, Kloc A, Ren J, Li F, et al. RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. Nature 2011; 479:135-8; PMID:22002604; http://dx.doi.org/10.1038/nature10501.