doi:10.1111/evo.13424

Cis-regulator runaway and divergence in asexuals

Frédéric Fyon[1] and Thomas Lenormand[1,2]

[1] *CEFE, CNRS, Univ Montpellier, Univ Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France*

[2] *E-mail: thomas.lenormand@cefe.cnrs.fr*

Received October 11, 2017
Accepted December 9, 2017

With the advent of new sequencing technologies, the evolution of gene expression is becoming a subject of intensive genomic research, with sparking debates upon the role played by these kinds of changes in adaptive evolution and speciation. In this article, we model expression evolution in species differing by their reproductive systems. We consider different rates of sexual versus asexual reproduction and the different type of parthenogenesis (apomixis and the various modes of automixis). We show that competition for expression leads to two selective processes on *cis*-regulatory regions that act independently to organism-level adaptation. Coevolution within regulatory networks allows these processes to occur without strongly modifying expression levels. First, *cis*-regulatory regions such as enhancers evolve in a runaway fashion because they automatically become associated to chromosomes purged from deleterious mutations ("Enhancer Runaway process"). Second, in clonal or nearly clonal species, homologous *cis*-regulatory regions tend to diverge, which leads to haploidization of expression, when they are sufficiently isolated from one another ("Enhancer Divergence process"). We show how these two processes cooccur and vary depending on the level of outcrossing, gene conversion, mitotic recombination, or recombination in automictic species. This study offers thus a baseline to understand patterns of expression evolution across the diversity of eukaryotic species.

KEY WORDS: **Allele-specific expression, asexuality, degeneration, enhancers, gene expression, gene regulation evolution, parthenogenesis, self-fertilization, silencing.**

The regulation of gene expression controls many aspects of phenotypes. Regulatory elements and their interactions with signaling cascades determine where and when proteins are expressed. Furthermore, they not only regulate the quantity in which these proteins are produced but also influence their final folding. Evolution of the regulome undoubtedly contributes to adaptation (Carroll 2005) as many studies have observed changes in regulatory elements or networks that correlate with adaptive novelties in bacteria (Cooper et al. 2003), yeasts (Ferea et al. 1999), insects (Raymond et al. 1998), fish (Shapiro et al. 2004), and birds (Abzhanov et al. 2004) to cite a few. Stabilizing selection on levels of expression have also been widely documented (Ludwig et al. 2000; Denver et al. 2005; Gilad et al. 2006; Whitehead and Crawford 2006; Fay and Wittkopp 2008).

Gene expression regulators belong to two categories: *cis*-acting DNA sequences (enhancers, repressors, insulators, 5'UTRs, and core promoters) and *trans*-acting RNA or protein intermediates (transcription factors and cofactors). *Cis*-acting sequences regulate the gene copy located next to them (i.e., on the same chromosome) rather than the gene copy on the homologous chromosome. Transcription factors regulate in *trans*, that is genes copies on both homologous chromosomes (for a review of gene expression regulatory architecture, see Wray et al. 2003).

The central role of *cis*-regulatory changes in adaptation has been strongly advocated (Wittkopp et al. 2004; Carroll 2005; Wray 2007). In this view, *cis*-regulatory sequences contribute disproportionately to adaptation because they are thought to have (1) a codominant effect on transcript abundance, which make them more visible to selection than coding sequence mutations (Wray 2007); (2) a greater chance to be beneficial compared to coding sequences since they are organized in distinct modules, each regulating a different aspect of expression profiles, limiting mutation pleiotropy (Carroll 2008; Wittkopp and Kalay 2012). This view is controversial and may be premature given the lack of strong theoretical and empirical validation (Hoekstra and Coyne 2007; Lynch and Wagner 2008). Alternatively, many regulatory

changes have been argued to evolve (quasi) neutrally (Oleksiak et al. 2002; Khaitovich et al. 2004; Yanai et al. 2004). This thought stems from the fact that some regulatory mutations may have no effect on expression profiles and that different networks can result in similar expression profiles, leading to similar phenotypes (Weirauch and Hughes 2010). Most regulation indeed relies on the proper matching between several regulatory components, whose precise identity or quantity is somehow arbitrary as long as they match (as in signal/receiver, or key/lock situations), allowing for considerable "evolutionary freedom" (Lenormand et al. 2009). Hence, the fitness landscape associated with regulatory networks is likely to present fitness ridges rather than a single well-defined peak as seen in stabilizing selection scenarios. Different networks may lead to the same optimal phenotype. Evolutionary divergence along those ridges can occur by the spread of slightly deleterious mutations and the occurrence of compensatory mutations restoring optimal expression profiles of targeted genes (Tautz 2000; Kuo et al. 2010; Weirauch and Hughes 2010; Coolon et al. 2014). The structural features of regulatory networks (complexity, connectivity, redundancy, . . .) may also be largely shaped by neutral processes (Force et al. 2005; Lynch 2007).

The debates over the role of regulatory changes in evolution have strongly focused on their role in organismal adaptation. Yet, self-promoting regulatory sequences may evolve in a runaway process even without optimizing individual level traits, which considerably complicate the evaluation of their usually envisioned role in adaptation. This process generates selection pressures that are not considered in neutral models. This was described using models explicitly considering mutations in regulatory (*cis*- or *trans*-) or genic regions in diploids (Fyon et al. 2015). We often use the short-hand term "enhancer" below (enhancers are typical examples of *cis*-regulators), but what we say applies to all kinds of *cis*-regulators. To summarize Fyon et al. (2015), this runaway occurs when enhancers compete for expression with different "strength." Differences in strength of the *cis*-regulatory region can arise because the "fit" of binding sequences toward transcription factor(s) differ or because the number of binding sequence (to transcription factor) differ, or because chromatin state differ etc. Stronger enhancers (i.e. those that activate more transcription, for a given cellular environment, including all transcription factors) tend to be preferentially associated with good genetic background. Indeed, stronger enhancers express a larger share of proteins than their weaker homologs. As a result, their associated gene copy (in *cis*) tends to contribute more to final phenotype, is more exposed to selection, and thus better purged from deleterious mutations. This association between stronger enhancers and good genes allows for the stronger enhancers to invade. Overall, the enhancers' strength escalates, leading to an endless "Enhancer Runaway" process.

This process should occur very generally for most genes across the genomes of most eukaryotes exposed to selection during their diploid phase. It is expected to happen, however, only for enhancers close enough to their target genes (so that the positive genetic association between the stronger enhancer and the purged gene is not broken down too often by recombination). It is important to note that Enhancer Runaway does not necessarily lead to increased expression levels, which would be deleterious. Due to the occurrence of multiple regulatory networks ensuring the same regulation, coevolution between regulators could ensure optimal expression levels despite ever-stronger enhancers (Fyon et al. 2015). For instance, the spread of stronger enhancers could be compensated by the concomitant spread of "weaker" associated transcription factors. Because of Enhancer Runaway, evolution through those multiple networks is expected to not be neutral, but rather biased for networks with strong enhancers close to the genes. Hence, the evolution of regulatory regions may not be simply driven by optimization of gene expression levels or quasi-neutral dynamics, but also by the endless spread of self-promoting *cis*-regulators. Enhancer Runaway may also have contributed to shape many features of expression control and regulatory networks (see discussion in Fyon et al. 2015). For example, some of the complexity of regulatory architecture might stem from the accumulation of *cis*-regulatory elements driven by this process.

Enhancer Runaway depends on allele-specific expression and should therefore be stronger when double heterozygotes are more frequent in the population, at the regulatory and at the regulated loci. However, recombination is less efficient when there are fewer double heterozygotes, and low recombination between these loci favors Enhancer Runaway. In addition, enhancer runaway may occur differently in asexuals where recombination is reduced and stronger enhancers may not be able to escape from their lineage of origin. We know already that moderate rates of self-fertilization slows Enhancer Runaway (Fyon et al. 2015). To obtain a more comprehensive view of the dependence of Enhancer Runaway process on the reproductive mode, we investigate in this article how this process varies in a wide variety of such systems, considering in particular the major forms of automixis and apomixis found in parthenogenetic species (see Asher 1970; Suomalainen et al. 1987; Schön et al. 2009; Nougué et al. 2015 for technical details about these reproductive systems). This may serve as a basis for the comparative analysis of the evolution of regulatory regions across sister species differing by their mode of reproduction and shed light on the evolution of regulatory regions in asexuals.

Methods

The model builds on the individual-based stochastic model used in Fyon et al. (2015). An individual genome is represented by two loci: one gene and its associated *cis*-regulatory region (e.g., an enhancer locus). *Cis*-regulatory sequences include many elements like enhancers, core promoters, or 5'UTRs. Results derived from

the model are valid for any type of *cis*-regulatory sequence, provided that it can mutate and alter expression levels. The process we study here highly depends on linkage disequilibrium between the gene and the *cis*-regulator, and therefore on the rate at which they recombine. We often use the term enhancers simply because they represent typical *cis*-regulatory sequences that can be found at different recombination distances from their target gene (Wray 2007). Individuals go through a simple life cycle: diploid selection, meiosis with recombination, mutations, and syngamy.

SELECTION

In this model, fitness depends on the presence of deleterious alleles on the gene and on their relative expression levels. However, we assume that total (absolute) expression levels do not influence fitness. This is similar to assuming that the gene is embedded into a negative feedback loop, which ensures constant optimal absolute expression levels. This assumption allows us to focus on the effects of expression asymmetry between homologs without the confounding effects of selection on absolute expression levels. Therefore, studying relative expression only requires focusing on *cis*-regulators. Studying the interplay between relative (between alleles) and absolute expression levels is beyond the scope of this article, as it would require considering more loci (other *cis*-regulators, trans-acting regulators, loop regulators) as several regulators are necessarily involved in total expression levels.

The fitness of an individual i is calculated as:

$$W_i = w_{i,1} + h_i \left(w_{i,2} - w_{i,1}\right), \quad (1)$$

where $w_{i,1}$ is the fitness of the fittest gene allele of individual i, $w_{i,2}$ the fitness of the other gene allele of individual i, and h_i the dominance coefficient of the least fit allele in individual i. Note that when $h_i = 0.5$, $W_i = (w_{i,1} + w_{i,2})/2$. Without loss in generality, we suppose that dominance coefficients depend on the relative expression level of the least fit allele, which depends on the relative strengths of homologous enhancers. We express dominance coefficients using the following relationship, which ensures that dominance of an allele increases with its relative expression and that dominance coefficient is always equal to the parameter h for equally expressed alleles:

$$h_i = \left(\frac{e_{i,1}}{e_{i,1} + e_{i,2}}\right)^{-\frac{Log(h)}{Log(2)}}, \quad (2)$$

where $e_{i,1}$ and $e_{i,2}$ are the strengths of enhancers associated with gene alleles of fitness $w_{i,1}$ and $w_{i,2}$, respectively, and h being the dominance coefficient in individuals that are homozygotes at the enhancer locus (and where both gene alleles are thus equally expressed). In the following, we set $h = 0.25$, which corresponds to the empirical consensus for the dominance of mildly deleterious mutations (Manna et al. 2012). With such partial recessivity, dominance of an allele increases more than linearly with its proportion of expression.

To model selection and reproduction, we sampled individuals in the population with replacement, accepting them with a probability equal to their fitness until we obtained two parents. One (possibly recombined) chromosome (one gamete) in each parent is then sampled to form a new offspring, and this procedure is repeated until we obtain N_{pop} offspring. In each parent, a recombination event between the gene and the enhancer can occur with a probability equal to the recombination rate per individual per generation R_{EA}. A recombination event can also take place between the gene-enhancer pair and the centromeres at a rate R_c per meiosis.

MUTATIONS

Mutations may occur in gametes. The gene locus undergoes recurrent, recessive deleterious mutations at a rate u_A. This is implemented by drawing the total number of mutations occurring in the population with a Poisson distribution of a mean equal to $2N_{pop}u_A$. The fitness effect of mutations is drawn from a negative exponential distribution of mean s. The enhancer locus also undergoes mutations at a rate u_E. Similarly, the number of enhancer mutations is drawn from a Poisson distribution with a mean equal to $2N_{pop}u_E$. As it is our focus, we assume that these mutations only affect expression levels of alleles of the gene. In particular, we do not consider that they have other effects (e.g., on expression timing or localization). Such mutations are expected to occur since *cis*-regulatory changes are thought to exhibit low pleiotropy (Carroll 2005; Wray 2007). Also, different regulatory networks can lead to similar phenotypes (Weirauch and Hughes 2010), indicating that pleiotropic effects of mutations can be easily compensated. Mutations change additively the $\log_{10}$ of enhancer strength, which ensures that the mutational effect size remains constant irrespectively of arbitrarily chosen absolute values of enhancer strength (see Fyon et al. 2015 for details). Mutation effects are drawn from a Normal distribution of mean zero and standard deviation σ_E. These assumptions of enhancer mutations allow us to observe a linear increase of enhancer strength under the Enhancer Runaway process (Fyon et al. 2015). Other mutational regimes can be considered, and are expected to lead to quantitative but not qualitative differences. Enhancer Runaway only needs mutations altering relative expression levels of homologous gene copies to occur.

REPRODUCTIVE MODES

In Fyon et al. (2015), we showed that self-fertilization reduced the rate of the ER process. Here, we provide a much more comprehensive study of the impact of the reproductive modes on Enhancer Runaway and we consider the case of complete absence of recombination, which leads to qualitatively very different

Table 1. Description of the different breading systems that allow a single individual to reproduce without any mate.

N°	Reproductive mode	# meiosis	Details	% of heterozygosity retention
1	Apomixis	0	Mitosis	100%
2	Premeiotic doubling	1	Duplication of chromosomes before meiosis, identical chromosomes resulting from duplication pair during meiosis I	100%
3	Postmeiotic doubling	1	Endomitosis of the meiotic product	0%
4	Central fusion	1	Central fusion in ordered tetrads or suppression of meiosis I	From 100% at centromere to 66% at large genetic distances (in Morgan) from centromere
5	Terminal fusion	1	Terminal fusion in ordered tetrads or suppression of meiosis II	From 0% at centromere to 66% at large genetic distances (in Morgan) from centromere
6	Random fusion	1	Random fusion in tetrads = mixed fusion with terminal and central fusion in proportion 2/3 and 1/3, respectively	Intermediate between central and terminal fusion
7	Self-fertilization	2	Syngamy of a male and a female gamete from two independent meioses within the same individual	50%

Apomixis (1) does not involve meoisis. Automixis (2–6) involves the production of an offspring by fusion of the two products from a single meiosis (unlike self-fertilization where offspring are produced by fusion of two products from two independent meioses, in the male and the female gametes, respectively). Central and terminal fusions are usually distinguished. Central (resp. terminal) fusion corresponds to the fusion of meiotic products derived from the first (resp. second) meiotic division. Central fusion retains heterozygosity at centromere positions while terminal fusion leads to the loss of heterozygosity at centromere positions. In both cases heterozygosity is reduced by one third at positions far away from the centromeres. Thus, automixis through central fusion combined with very low recombination rates leaves a genetic signature very similar to that of apomixis (with maintenance of high level of heterozygosity). In contrast, central, terminal, and mixed fusions combined with very high recombination rates leaves a genetic signature very similar to self-fertilization (with nearly complete loss of heterozygosity). This table is adapted from (Nougué et al. 2015).

outcomes. Different modifications of the simple life cycle described above were considered to implement different reproductive systems. These reproductive systems are presented in Table 1 and illustrated on Figure 1. They include self-fertilization with various forms of automixis and apomixis. These reproductive modes correspond to the major cases found in Eukaryotes, extensively studied for the last 50 years: see (Asher 1970; Mittwoch 1978; Nougué et al. 2015) for a classification and (Bell 1982; Suomalainen et al. 1987; Schön et al. 2009) for extensive reviews and taxonomic surveys.

With self-fertilization, the two chromosomes transmitted to the offspring are chosen from two independent meiosis events from a single parent individual. With automixis, the two chromosomes forming a new offspring are sampled without replacement among the four meiotic products of a single meiosis. The precise sampling scheme depends however on the mode of automixis. With Central Fusion, the two chromosomes sampled come from meiotic products that are split at meiosis I. With Terminal Fusion, they come from meiotic products that are split at meiosis II. In Random Fusion, they come at random from the four meiotic products. In premeiotic doubling, the DNA is replicated twice (instead of once in normal meiosis) before meiosis and one of the diploid meiotic products becomes the zygote. In postmeiotic doubling, the genome is doubled in the meiotic products of a normal meiosis, and becomes the diploid zygote. Finally, in apomictic reproduction, no meiosis occurs, and individuals are formed from a mitotic division in the parent. These modified life cycles occur with a given probability to each individual at every generation to model-mixed mating systems.

In many simulations, individuals have mixed reproduction. In each generation, a proportion α of individuals reproduce through sex (outcrossing), while others reproduce through another mode: self-fertilization and the different known forms of parthenogenesis (see Table 1). These different modes correspond to a different loss

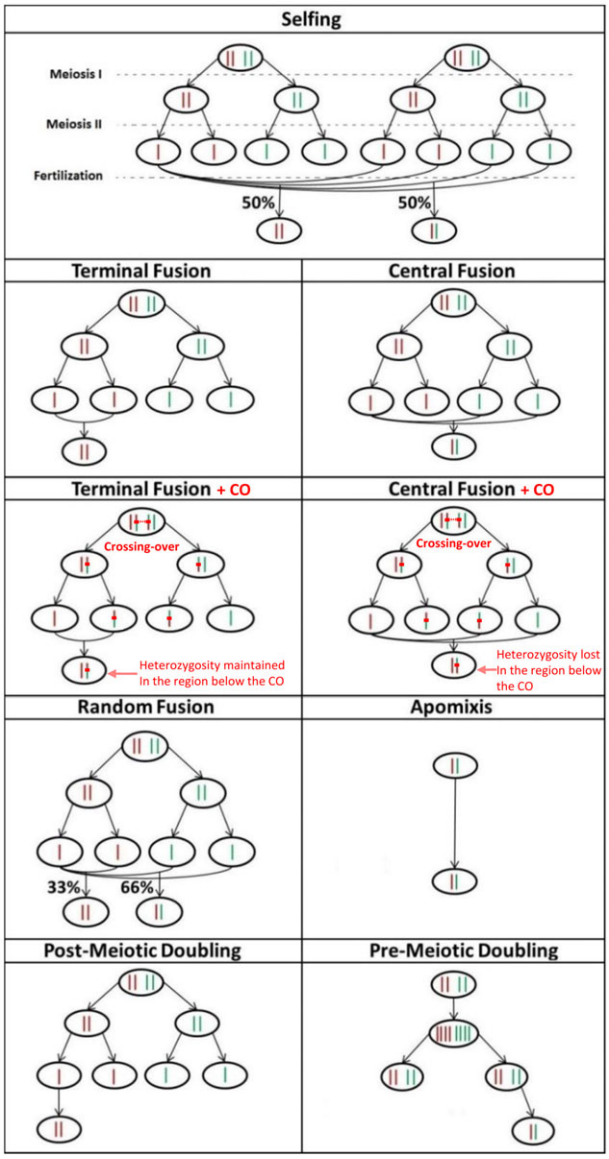

Figure 1. Schematic illustration of the various reproductive systems implemented in our models. In selfing, gametes from two independent meiosis fuse to form a new individual. In automictic reproduction, gametes from the same meiosis fuse to form a new individual. Those gametes may come from the same Meiosis II division (Central Fusion) or from different Meiosis II division (Terminal Fusion). They may also randomly come from the same or different Meiosis II divisions (Random Fusion). In automictic Premeiotic doubling, DNA content is doubled before Meiosis I, such that products of meiosis are directly diploid individuals. In automictic Postmeiotic doubling, DNA content is doubled in gametes to give birth to a diploid individual. Finally, apomixis is a mode of reproduction where new individuals are produced mitotically without any meiosis division. The consequence of recombination are illutrated for terminal and central fusion. With terminal/central fusion, recombination can lead to the maintenance/loss of heterozygosity in the portion of the chromosome distal from the crossing-over position (illustrated in red). On the figure, the position of the centromere is at assumed to be at the top of chromosomes.

of heterozygosity k per generation. At equilibrium, the inbreeding coefficient F for a neutral allele in such systems is

$$F = \frac{k\,\alpha}{1-(1-k)\,\alpha} \tag{3}$$

In strictly asexual species, other important mechanisms such as mitotic recombination or gene conversion, even if they occur at a low rate, may influence the long-term evolution of the different loci. To address this, we included gene conversion on the regulatory region, on the gene, or on both. Apart from the fact that conversion on genes versus regulatory regions may have varying dynamical effects in our model, there is strong empirical evidence indicating that rates of recombination/conversion may markedly differ in these regions (Lenormand et al. 2016). For simplicity, we assume that when conversion occurs at a locus, the individual becomes homozygous for one of its (randomly chosen) alleles. In reality, gene conversion events do not lead to a full loss of heterozygosity (across the whole gene or regulatory region) as conversion tract length are typically shorter than genes or regulatory regions. This has to be carefully taken into account when reading the figures with quantitative estimates in mind. Finally, meiotic recombination was modeled simply by allowing recombination events between chromatids. Note that mitotic recombination is equivalent to meiotic recombination in central fusion automicts, although typical rates of mitotic recombination are much lower than meiotic ones. In both cases, recombination with the centromere leads to loss-of-heterozygosity in the portion of the chromosome distal to the position of the crossing-over. For simplicity, we grouped the two similar situations under the acronym RWC (Recombination with the Centromere).

SIMULATIONS AND PARAMETER VALUES

Our objective is to study how cis-regulatory sequences evolve with various reproductive modes and with very low levels of recombination. To study this variation, we considered a typical situation where the enhancer runaway is well characterized and relatively fast (Fyon et al. 2015): (1) high rates of mutations ($u_A = u_E = 10^{-4}$) that allow for significant polymorphism; (2) high intensity of purifying selection ($s = 0.1$) on the gene and low recombination rates ($R_{EA} = 10^{-6}$ unless under apomixis where $R_{EA} = 0$), to allow for substantial indirect selection on *cis*-regulatory sequences; (3) relatively large population size, to reduce stochastic effects of genetic drift ($N_{pop} = 10\,000$). We considered the different modes of reproduction, as described above, and for each, we varied the rate of nonrandom mating. For the special case of automixis through central fusion, we first assumed no RWC ($R_c = 0$; i.e., no recombination between the enhancer-gene pair and the centromere). We then relaxed this assumption and introduced RWC at various rates. We also considered other small departures from strict cloning by introducing gene conversion or

mitotic recombination. Simulations were repeated 200 times, for 10^5–10^6 generations depending on cases.

At the start of simulations, all individuals are homozygotes for both loci. For 1000 generations, mutations on the gene locus occur. After these burn-in generations, the gene locus is close to selection-mutation equilibrium and mutations on the enhancer locus occur. In simulations, we measured the frequency of heterozygotes in the population and the rates of enhancer strength escalation. The latter is computed as the slope of mean enhancer log-strength through time (which increases linearly, as in Fyon et al. 2015). When investigating the question of enhancer and gene divergence in cases of full apomixis, full central fusion, and full premeiotic doubling, we follow the mean enhancer log-strength of the stronger enhancer-bearing chromosome (log *E*+), mean enhancer log-strength of the weaker enhancer-bearing chromosome (log *E*-), and mean fitness effect of gene alleles on corresponding chromosomes (*s*+ and *s*–). To provide a measurement of allele-specific expression, we calculate and display on Figure 3B the log-ratio of homologous enhancers' strengths log(E+/E–). By definition, this log-ratio is always positive.

Results

EFFECT OF HETEROZYGOSITY ON ENHANCER RUNAWAY

We first checked the effects of the different reproductive systems on heterozygosity levels. Results are shown on Fig. 2A, plotted against the rate of asexual reproduction per generation (a rate of 0.2 apomixis, for example, indicates that, each generation, 20% of individuals reproduce without meiosis, that is by mitosis, while 80% of them outcross). We first ignore the recombination between the gene-enhancer pairs and the centromeres, which has no influence except for terminal and central fusion automixis. For the latter, we start considering that this rate is zero, but this assumption will be relaxed later. First, we see, as expected, that apomixis (or mitosis), premeiotic doubling and central fusion automixis do not alter heterozygosity compared to fully outcrossing populations (except at low levels of outcrossing, when chromosomes diverge, see below). On the contrary, all other reproductive systems tend to decrease heterozygosity. As expected, heterozygosity is lowest when reproductive systems entirely eliminate heterozygosity with postmeiotic doubling or terminal fusion automixis with no recombination with centromere. Equilibrium heterozygosity is then gradually higher with reproductive systems reducing offspring heterozygosity by one half (self-fertilization) or one third (random fusion, which also corresponds to terminal and central fusion occurring with probability 1/3 and 2/3, respectively).

We plot on Figure 2B enhancer strength escalation against the rate of asexual reproduction. Apomixis, premeiotic doubling and central fusion give similar results as they exhibit constant

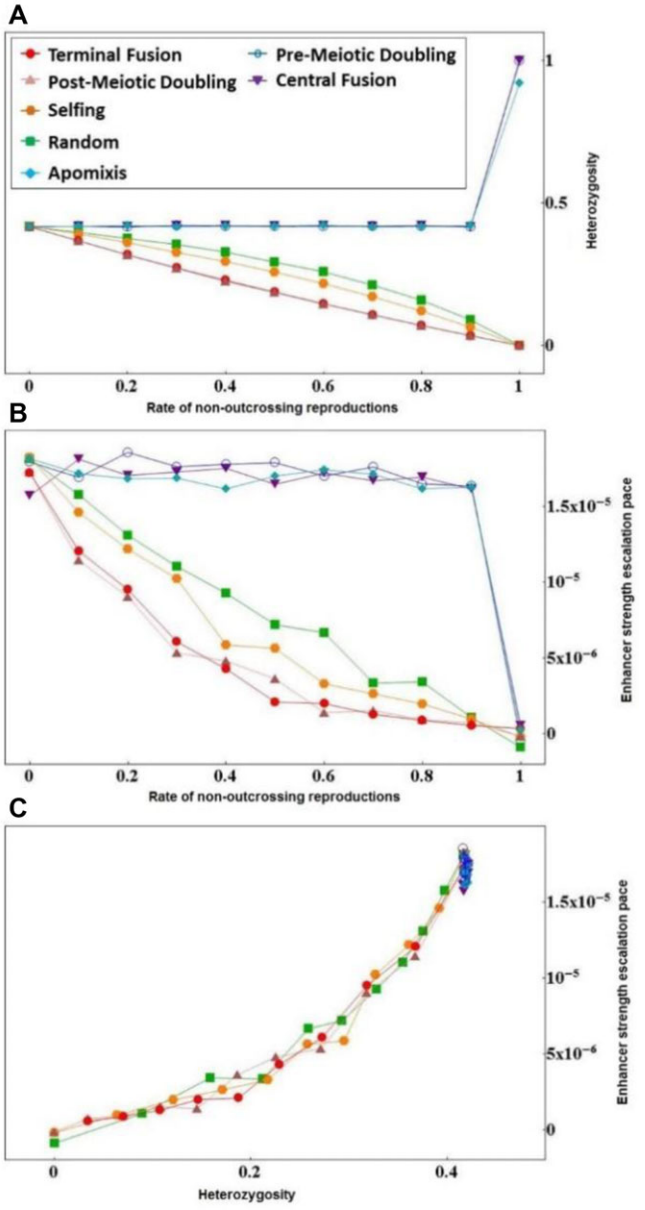

Figure 2. Heterozygosity and enhancer strength escalation rate for various reproductive systems. Simulated populations reproduce partly through outcrossing and partly through other reproductive systems. (A) Mean heterozygosity in function of the rate of nonoutcrossing. (B) Enhancer strength escalation rate as a function of the rate of nonoutcrossing. (C) Enhancer strength escalation rate against mean heterozygosity. Enhancer strength escalation rate is computed as the slope of the linear regression fitting the mean increase of enhancer log-strength through time. Individuals that do not reproduce through outcrossing reproduce through terminal fusion (red circles), postmeiotic doubling (pink upright triangles), selfing (orange hexagons), random fusion (green squares), apomixis (light blue diamonds), Premeiotic doubling (blue rings), or central fusion (purple downright triangles). Enhancer strength escalation rate increases quadratically with mean heterozygosity, except under full clonality (Premeiotic doubling or apomixis or central fusion with no recombination).

escalation rates that correspond to the rate under random mating. The only difference with random mating occurs at very low level of outcrossing. In that case, escalation rates drop precipitously to zero when the outcrossing rate tends to zero. With other mating systems, the rate of escalation gradually reduces with reduced levels of outcrossing and tends to zero when outcrossing rate tends to zero. For a given level of outcrossing, escalation rates are high for random fusion, intermediate for self-fertilization, and low for postmeiotic doubling/central fusion. Overall, the level of heterozygosity maintained in the population by the different mating systems fully predicts escalation rates (except for cases of full apomixis/premeiotic doubling/central fusion). Figure 2C illustrates this relationship by plotting the escalation rates against heterozygosity in the populations (full apomixis/premeiotic doubling/central fusion points are not shown, as they exhibit a specific behavior, which is examined below). This scaling clearly shows that there is no other effect of mating systems once heterozygosity level has been taken into account. Unless outcrossing rate tends to zero, this analysis shows that reproductive systems impact Enhancer Runaway only by changing heterozygosity levels in the population. Lower heterozygosity levels reduce the rate of enhancer strength escalation.

A NEW OUTCOME: "ENHANCER DIVERGENCE" PROCESS

In cases of full apomixis, full premeiotic doubling, and full central fusion (with no RWC), Enhancer Runaway does not occur. Indeed, we observe maximal levels of heterozygosity, and yet enhancer strength does not escalate. To understand what happens in these cases, we tracked mean enhancer strength on each homologous chromosome separately in every individual. To do so, we defined a stronger enhancer-bearing chromosome as E+ chromosome and a weaker enhancer-bearing chromosome as E– chromosome in every individual, and calculated the population mean enhancer strength for each of these chromosomes. Results for apomixis are shown on Figure 3, comparing a case of full apomixis to a case of partial apomixis. Figure 3A shows that these two cases lead to very different results. In partial apomixis, Enhancer Runaway occurs as usual: all chromosomes increase in strength. In full apomixis, however, there is a divergence between chromosomes in every clonal lineage as one chromosome accumulates enhancer strength-increasing mutations while the other accumulates enhancer strength-decreasing mutations. Consequences on the gene locus can be readily seen in Figure 3B. We computed *s*+ and *s*–, the averages across individuals of the fitness effects of alleles on the gene for E+ and E– chromosomes, respectively. In partial apomixis, both *s*+ and *s*– reached an equilibrium. The two values differed as stronger enhancers tend to be purged and are associated with less deleterious alleles. However, this difference does not increase through time. In contrast, with full apomixis, *s*+ and *s*– diverged through time. On an E+ chromosome, the gene locus is well purged from deleterious alleles, as it progressively becomes the only expressed chromosome (*s*+ decreases). On E– chromosome, deleterious mutations accumulate as they are hidden due to low expression (*s*– increases). Cases of full premeiotic doubling and full central fusion give the same results. Overall, in cases of full apomixis, full premeiotic doubling and full central fusion with no recombination with the centromeres (clonal lineages), selection on enhancers results in a divergence of enhancer and gene copies. We refer to this process as the "Enhancer Divergence" process. This process results in a progressive haploidization of expression as well as the simultaneous degeneration of the unexpressed gene copy.

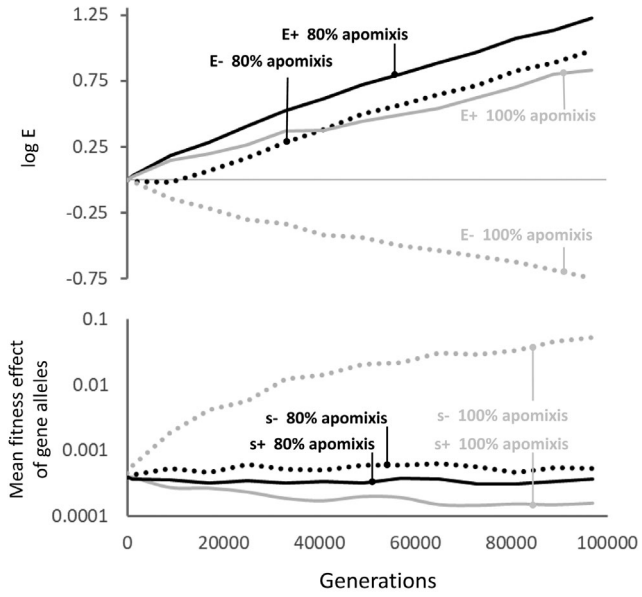

Figure 3. **Time dynamics of mean *cis*-regulator strength (top panel) and mean fitness effects of alleles on the gene (lower panel). Results are shown for chromosomes with stronger *cis*-regulator (E+ and s+; plain lines) and for chromosomes with weaker *cis*-regulator (E- and s-; dotted lines). In black, the reproductive system is 80% apomixis/20% outcrossing. In gray, the reproductive system is 100% apomixis (i.e., equivalent to mitosis). With partial apomixis, *cis*-regulator strength escalates on both chromosomes. The mean fitness effect of alleles of the gene also reaches an equilibrium, with more deleterious alleles on E$^-$ chromosome (weaker *cis*-regulators tend to get preferentially associated with more deleterious alleles). With 100% apomixis, chromosomes diverge. E+ chromosomes accumulate stronger *cis*-regulator whereas E– chromosomes accumulate weaker *cis*-regulator. E+ chromosomes also become purged from deleterious alleles (s+ decreases), as they become more expressed than E– chromosome. Conversely, E– chromosomes accumulate deleterious mutations (s– increases) and become silenced.**

Why do chromosomes diverge in clonal lineages? First, it can occur because there is no genetic shuffling. Without

recombination, gene conversion or gamete shuffling (through outcrossing or inbreeding), a single mutant enhancer cannot become homozygote in any individual (like for any beneficial mutation in diploid asexuals, see Kirkpatrick and Jenkins 1989; Mandegar and Otto 2007). Enhancer mutations, as well as gene mutations, are restricted to the chromosome on which they reside. In our infinite-allele model, clonal lineages all end up being heterozygous due to enhancer and gene mutations. Hence, selection on the enhancer locus tends to favor individuals expressing less the deleterious allele that happens to occur on one of the copies of the gene. In other words, genetic association between stronger enhancers and more viable gene alleles is selectively favored, as in Enhancer Runaway, but the result is different. Rather than invading the whole population, stronger enhancers concentrate on the chromosome bearing the most favorable gene allele in every individual. The process is self-reinforcing: as one gene becomes underexpressed and thus partially hidden from selection, it becomes even more likely to accumulate new deleterious mutations. In turn, as one gene accumulates deleterious mutations, enhancer strength-decreasing mutations become more favorable. This process repeats itself until haploid expression is reached at this particular gene. Note that the process occurs independently for each gene, so that each homolog may maintain expression for different genes. Thus, this process is unlikely to lead to the entire degeneration of one chromosome (like with, e.g., Y chromosomes) as each of the two homolog chromosomes is expected to preserve a different set of genes.

To fully understand the differences between Enhancer Runaway and Divergence, we compared enhancer strength escalation between both processes on a longer timescale than in Figure 3.

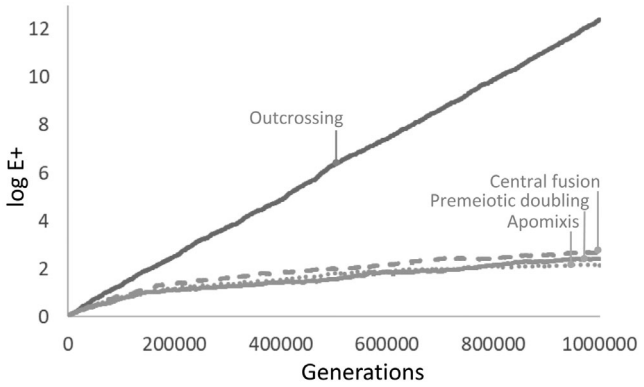

Figure 4. Comparison of *cis*-regulator strength escalation with different reproductive systems. *x*-axis: number of generations. *y*-axis: mean log strength of the stronger enhancer bearing chromosome (log E+). With outcrossing (dark gray), mean strength of all *cis*-regulator increase linearly with time. In fully parthenogenetic lineages (light gray), it increases less than linearly. Note that in the case illustrated with central fusion (dashed line), there is no recombination with the centromere ($R_c = 0$).

Figure 4 illustrates this comparison. Though Enhancer Runaway ends up in a linear increase of enhancer strength, it is not the case of mean enhancer strength of stronger enhancer bearing chromosomes in Enhancer Divergence process. In cases of clonal lineages, we see that mean enhancer strength of E+ chromosomes increases logarithmically. After a sufficient amount of time, the marginal increase in enhancer strength becomes negligible. This is because the Enhancer Divergence is driven by the progressive hiding of deleterious alleles on the chromosomes having the weaker enhancer. In our model, extinction of one gene is asymptotic and the process stops when extinction is nearly complete (i.e., when haploidization of expression is nearly achieved).

ENHANCER DIVERGENCE IN NONSTRICT ASEXUALS

Figure 2 shows that the transition between Enhancer Runaway and Divergence occurs at low rates of outcrossing. To better understand the transition between these two regimes, we studied the effect of various departures from strict cloning (outcrossing, RWC, gene conversion). Figure 5 shows the escalation rate, enhancer strength divergence ratio, $\mathrm{Log}(E^+/E^-)$, and heterozygosity of populations after 250 000 generations.

There is a transition between Enhancer Runaway and Divergence at outcrossing rates between 10^{-5} and 10^{-3}. High rate of outcrossing results in a residual level of divergence. This level only reflects the fact that the $\mathrm{Log}(E^+/E^-)$ measure is positive as long as there is heterozygosity in the population. As outcrossing rates become small, escalation rates decrease (Enhancer Runaway leads to slower escalation at lower levels of heterozygosity, see Fig. 2), while heterozygosity and enhancer divergence increase (Enhancer Divergence becomes stronger as homologous chromosomes become more and more isolated). The transition shows that Enhancer Divergence only occurs at very low rates of outcrossing. In the range 10^{-5}–10^{-3}, both Enhancer Runaway and Divergence occur: there is some divergence of homologous sequences, while enhancer strength globally increases.

With RWC, we see that Enhancer Runaway only occurs within a small intermediate range of recombination rates (around 10^{-4} on Fig. 5). At lower recombination rates, homologous chromosomes are too isolated for Enhancer Runaway to occur, but homologs diverge. At higher recombination rates, recombination with the centromeres causes a strong reduction in heterozygosity, which prevents both Enhancer Runaway and Divergence to occur. Figure 6 illustrates the time-dynamics of these different regimes. In particular, we can see runaway occurring at intermediate RWC (orange curves). The time dynamics also shows that detectable divergence takes place relatively quickly over few thousand generations ($\mathrm{Log}(E^+/E^-) = 0.1/0.2$ corresponds already to 25%/60% difference in expression levels).

With or without gene conversion, Enhancer Runaway does not occur at all in fully clonal lineages. This is expected as gene

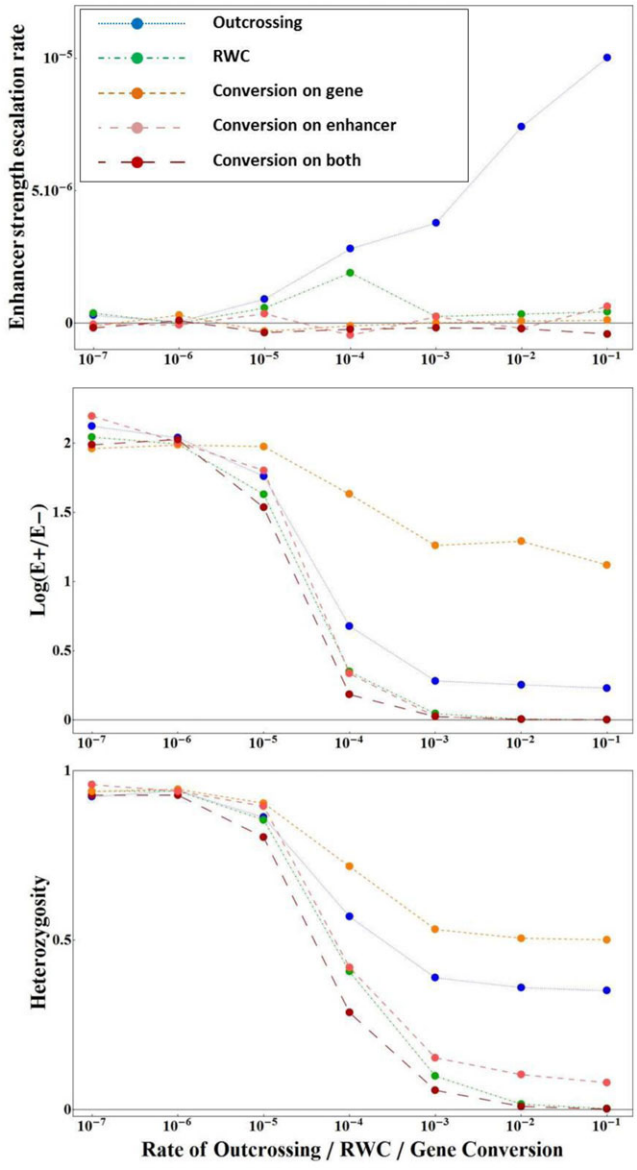

Figure 5. *Cis*-regulator runaway and divergence in nonstrict asexuals. *y*-axis: (A) *Cis*-regulator strength escalation rate; (B) Homologous *cis*-regulator divergence, measured as Log(E+/E-); (C). Average heterozygosity of both *cis*-regulators and gene loci. The curves represent different departures from strict clonality. Blue circles and dotted line: populations reproduce partly through outcrossing (outcrossing rates on *x*-axis). Green circles with dotted and dashed line: Central fusion automixis with Recombination with Centromere (RWC) (RWC rate on *x*-axis). Orange circles with small dashed line: apomixis, with gene conversion occurring on the gene (conversion rate on *x*-axis). Pink circle with medium dashed line: apomixis with gene conversion occur on the *cis*-regulator (idem). Red circles with large dashed line: apomixis with gene conversion occur on both the gene et *cis*-regulator (idem). See Methods for the calculation of *cis*-regulator strength escalation rate, *cis*-regulator divergence, and heterozygosity. *Cis*-regulator divergence and heterozygosity results correspond to divergence and heterozygosity after 250 000 generations.

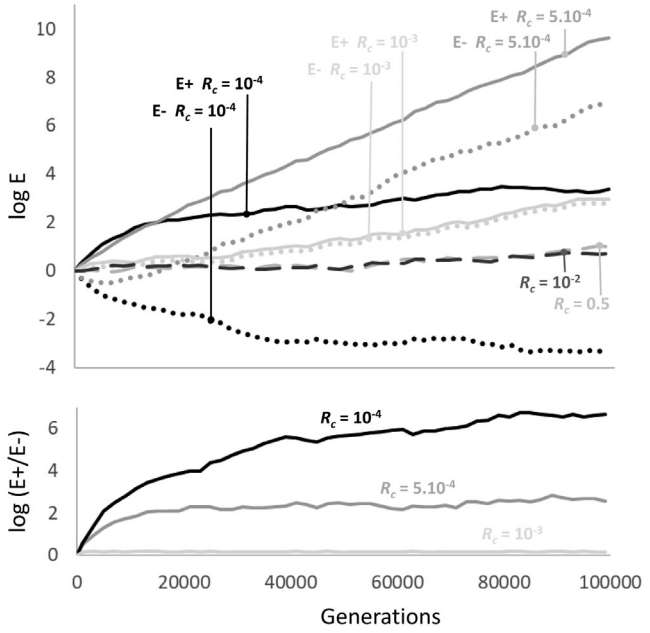

Figure 6. Time dynamics of *cis*-regulator runaway and divergence in asexuals with RWC. These results can correspond either to 100% central fusion automicts with rate of recombination R_c between enhancer-gene loci and the centromere or to 100% clonal (mitotic) reproduction with mitotic recombination occurring at the rate R_c. The color and symbol code for the R_c values are indicated directly on the graph. Top panel: Mean *cis*-regulator log strength of stronger (E+, plain curves) or weaker (E-, dotted curves) *cis*-regulator bearing chromosomes. Results strongly depend on R_c values, with three regimes: divergence without escalation ($R_c = 10^{-4}$), small divergence and escalation ($R_c = 5.10^{-4}$), no divergence and no escalation (R_c above 0.01). Note that because there is no significant divergence between E+ and E- chromosomes for R_c values above 0.01, a single dashed curves are presented for $R_c = 0.01$ and $R_c = 0.5$. Lower panel: Mean enhancer strength divergence between homologous chromosomes. Divergence is given as the $\log_{10}$ of the ratio of stronger (E+) over weaker (E–) mean *cis*-regulator strength. No curve is indicated for R_c values 0.01 and 0.5 as there is no significant divergence in these cases. For instance, with a value of 0.1/0.2 expression already differs by 25%/60% between the two alleles, which would be in principle detectable with transcriptomic data. Divergence between homologues decreases as R_c increases. When divergence occurs, it is relatively fast and occurs in a few thousand generations.

conversion reduces heterozygosity and genetic associations and both effects play against the ER process. Gene conversion in a clonal lineage can largely minimize the Enhancer Divergence as well (Fig. 5). This is also expected as divergence/heterozygosity is directly reduced by conversion. Conversion has a much stronger effect when it occurs on the regulatory region than on the gene. Indeed, conversion on the regulatory region directly reduces divergence, while conversion on the gene only removes

the cause of the divergence (i.e., heterozygosity for deleterious mutations).

Discussion

ENHANCER RUNAWAY PROCESS IN PARTIAL ASEXUALS

In sexual diploid species, *cis*-regulators compete for expression, which tends to selectively favor stronger *cis*-regulatory alleles. Here, stronger means that in a heterozygote, the gene associated to the stronger *cis*-regulator will contribute a larger share of expressed protein than its homolog. The cause of this selection-pressure is that genes associated with a stronger *cis*-regulator are more exposed to selection and therefore better purged from deleterious mutations. We termed this process "Enhancer Runaway," as it leads to a runaway-like increase of *cis*-regulator strength (Fyon et al. 2015). It is a very general process that should occur on every gene in every diploid species. It works, however, only for *cis*-regulatory sequences located at small recombination distance to the regulated gene. By construction, this process depends on the occurrence of heterozygotes on both the regulated and regulatory sequences. It also depends on recombination, whose efficacy relies on the occurrence of double heterozygotes. Here, we studied the evolution of *cis*-regulators in species with a large diversity of reproductive modes. The distribution of genetic variation and recombination levels is indeed strongly influenced by these different modes. An important result is that the impact of the various reproductive modes on Enhancer Runaway is only through their effect on heterozygosity levels. Lower heterozygosity leads to a slower Enhancer Runaway. With lower heterozygosity, there is less room for homologous regulatory sequences to compete for expression. Figure 2C further shows that the rate of Enhancer Runaway increases more than linearly with heterozygosity levels, as expected from the fact that both the *cis*-regulator and gene need to be heterozygous for the process to operate.

ENHANCER DIVERGENCE PROCESS IN CLONAL LINEAGES

In fully clonal lineages, mutations are bound to the genetic background in which they first occur. They cannot spread to the whole population as they stay confronted and only compete with the same other homologous allele within a clonal lineage. If a newly arising, stronger, *cis*-regulatory mutation occurs in a clone homozygous at the gene, it will be neutral. If it becomes homozygous in that lineage, it will also be neutral in the population. Thus, Enhancer Runaway cannot occur. However, while it is still heterozygous within a lineage, it can be associated in *cis* or *trans* to a deleterious mutation. If it is associated in *cis*, it will lead to the overexpression of the deleterious mutation and will therefore be selected against. If it is associated in *trans*, it will, on the contrary, lead to the favorable underexpression of the deleterious mutation. Hence, a stronger *cis*-regulator can spread within a clonal lineage, but, it stays heterozygous and associates to the gene copy less loaded in deleterious mutations. Eventually, a second stronger *cis*-regulatory mutation will occur, and for the same reason, will either reinforce the previous mutation or be lost, and so on. In the meantime, while stronger *cis*-regulators accumulate on one of the homologs, other deleterious mutations will occur. Because of the partial silencing of one of the gene copies, they will persist longer or will be more likely to fix when they occur on that partially silenced copy. This self-reinforcing process will eventually lead to haploidization (i.e., total silencing and degeneration of one gene copy) when starting from a newly arising clonal lineage. Selection for favorable genetic associations (which is also at the basis of Enhancer Runaway), in clonal lineages, leads to the accumulation of stronger enhancer (more generally *cis*-regulator) associated with a purged gene copy on one chromosome, and weaker enhancers and deleterious gene mutations on the homologous chromosome. We refer to this process as the Enhancer Divergence process. The homologous copy of the gene that becomes silenced and degenerated entirely depends on the initial stochasticity of occurrence of mutations on the gene and its regulatory region. It is expected to vary from one gene to another. This process thus does not necessarily lead to the degeneration of one chromosome as the two homologs each preserve a distinct subset of active genes in diploid clonal lineages.

Enhancer Divergence has similarities with the degeneration of redundant gene duplicates (Innan and Kondrashov 2010). In this case, deleterious mutations are free to accumulate on one duplicate because they are hidden by the presence of the other functional duplicate. In this view, "degenerative" mutations can be diverse and are not distinguished (e.g., a change in amino acid, the occurrence of a stop codon, a frameshift, a large deletion, or a silencing mutation in the *cis*-regulatory region). Degeneration of the redundant duplicate simply results from the relaxation of purifying selection and the fixation of an almost neutral mutation. In contrast, we explicitly model *cis*-regulatory sequences to analyze the interplay between accumulation of mutations in the coding and regulatory region. Enhancer Divergence occurs through the simultaneous accumulation of mutations and the build-up of genetic associations in coding and regulatory mutations. This is a process of selective silencing, where regulatory mutations play the role of compensatory mutations and mitigates the fitness effects of deleterious mutations on coding sequences. It would certainly be interesting to use this model for studying duplicate silencing given the data available (Li et al. 2005; Ganko et al. 2007). Finally, it is important to note that degeneration of a redundant gene copy could occur without regulatory evolution (by a loss-of-function mutation) and that selective silencing could occur in a second, independent step. The ED process shows that these two

steps can occur simultaneously and combine. However, adding these processes and type of mutation would lead to even faster haploidization.

CIS-REGULATOR EVOLUTION IN NON-STRICTLY CLONAL LINEAGES

The dynamic process of Enhancer Divergence, like all models of allele divergence in asexuals, relies on the fact that homologous chromosomes remain "isolated" from each other. This isolation can be modified by several processes (as for Meselson effect, see discussion in, e.g., Butlin 2002). We studied different mechanisms that can cause a departure from strict clonal reproduction: partial outcrossing, RWC, and gene conversion. Again, RWC covers cases of recombination with the centromere that occur with meiotic recombination in central fusion automicts or more generally with mitotic recombination. These different mechanisms reduce heterozygosity compared to strict-cloning. This reduction can prevent Enhancer Divergence to take place. Enhancer Divergence occur only with rates of outcrossing/RWC/conversion divergence below 10^{-3}–10^{-4} on Fig. 5). Enhancer Divergence is gradually replaced by Enhancer Runaway as the rate of outcrossing increases. This change of regime is not observed with RWC and conversion. Contrary to outcrossing, these two processes lead to a rapid and near complete loss of heterozygosity when their rate exceeds c.a. 10^{-3}, which suppresses both Enhancer Runaway and Divergence. Enhancer Runaway can, however, occur in an intermediate regime with intermediate rates of RWC (around 10^{-4} on Fig. 5; simulations not shown indicate that this intermediate rate is of the same order as the mutation rate). In this regime, there is enough heterozygosity for Enhancer Runaway to occur. Homologs are not fully isolated and the favorable association between the (strong) *cis*-regulator and the (purged) gene can fix on both homologues. These fixations can occur repeatedly, which causes the Runaway. With RWC, the favorable association of the *cis*-regulator and the gene is preserved and not altered by recombination with the centromere. This favorable combination is preserved and can segregate. In contrast, this combination is lost with gene conversion: a conversion event arises on only one of these two regions independently, which necessarily breaks the genetic association. This difference explains why there is no intermediate Enhancer Runaway regime with gene conversion. The effect of conversion on Enhancer Divergence is particularly strong when it occurs on the regulatory region, since it directly reduces divergence. When conversion occurs on the gene, it only removes the cause of divergence (heterozygosity for deleterious mutations) and not divergence itself.

The biological range of these parameter values in natural population is difficult to estimate. Rare outcrossing can occur in asexuals to different levels (and ease of detection) (Schurko et al. 2009). The rate of mitotic recombination depends on the position of the focal locus to the centromere, it varies, e.g., in the range 10^{-4}–10^{-6} in for example yeast (see, e.g., Mandegar and Otto 2007). Rates of gene conversion are usually quite low, c.a. 10^{-5} (Gay et al. 2007), but also heterogeneous within genomes, and possibly more frequent in regulatory regions than in genes for many eukaryotes, depending on hotspot location (Lenormand et al. 2016).

LIMITS OF THE MODEL AND POSSIBLE EXTENSIONS

As in Fyon et al. (2015), we did not add an upper bound or a physical limit to cis-regulator "strength." As explained in the introduction, there are many ways that *cis*-regulatory "strength" could be increased for a given cellular environment (binding sequence identity, number of binding sequences, chromatin state etc.), and it is not clear whether a physical limit would be easily reached, especially if other regulators (e.g., transcription factors) evolve as well. In any case, adding such limit would simply add a trivial modification of the results: runaway or divergence will simply hit this limit and stop.

Considering a multilocus model may quantitatively alter some of the results mentioned above, especially in models involving RWC. Indeed, when there is recombination with the centromere, all loci that are distal to the crossing-over position (i.e., between this position and the telomere) become homozygotes. Some of these loci will become homozygote for the beneficial association (stronger enhancer and purged gene copy), but some will become homozygote for the other association between weaker enhancers and deleterious gene copies. Hence, the closer the crossing-over position is to the centromere, the less likely it is that it will lead to viable offspring (given that it will cause homozygosity at more loci). In other words, lineages that survive are those that experienced few recombination events as the recombination events are located distally from the diverging enhancer-gene pairs. This low effective recombination rate is then likely to promote divergence at larger genetic distance (from the centromere) than in the single locus model. We may even expect divergence to progressively spread from the centromere through time. Independently, recombination suppression may also evolve simply to prevent loss-of-heterozygosity and the exposition of deleterious mutations in homozygous state in central fusion automicts (Haag et al. 2017).

The various reproductive modes that we considered cover all major cases of parthenogenesis and asexuality. Yet, we did not explicitly model all possible life cycles and combinations of reproductive modes. Our results indicate that, as far as Enhancer Runaway is concerned, it is sufficient to understand how levels of heterozygosity will vary to figure out what will be going on. Also, we did not consider selection and mutation during episodes of somatic development, vegetative growth, or other kinds of mitotic divisions in the life cycle. Similarly, we did not include haploid

selection. Depending on parameters, adding these processes could certainly alter quantitatively the mutation selection-balance and levels of heterozygosity. However, as in similar models (Otto and Orive 1995), this is captured by varying selection intensity in our model, which is unlikely to cause qualitative changes in our results.

TESTING ENHANCER RUNAWAY

The different outcomes of Enhancer Divergence and Runaway lead to different predictions that could be empirically investigated. Concerning Runaway, we predict that enhancers will be generally stronger in populations with higher heterozygosity levels. This may lead to a variety of tests. For example, this could be investigated by measuring allele-specific expression in F1 hybrids between species differing by their reproductive systems. In such tests, as all *trans*-acting regulators from each parent are shared in the hybrid, any difference in transcript abundance between genes coming from each parents are due to differences in *cis*-acting regulators. Because the rate of the ER process increases with heterozygosity levels, we expect genes inherited from the outcrossing species to be on average more expressed in the hybrid. Expression patterns in this kind of hybrid have been investigated by He *et al.* (2012). They crossed outcrossing plants *Arabidopsis lyrata* (as male) with selfing *Arabidopsis thaliana* (as female), and found that, in 90% of the genes displaying allele-specific expression, the *lyrata* genome was preferentially expressed. However, this may also be caused by a sex-of-origin (rather than a species-of-origin) effect as the study lacks reciprocal crosses due to technical difficulties. A similar study was performed by Steige et al. (2015) using outcrossing *Capsella grandiflora* (as female) and selfing *Capsella rubella* (as male). In this case, results were in the opposite direction (*C. rubella* genes tended to be preferentially expressed in the hybrid). But again, sex-of-origin effects cannot be ruled out. Finding a model system where the two different directions of the cross could be performed would be particularly insightful.

More generally, this prediction could be tested comparing strength of enhancers between populations with different heterozygosity levels. Reproductive systems are not the only factor influencing heterozygosity. Other factors include mutation, migration, nonrandom mating, selection, recombination, gene conversion, and population size. Such factors act in concert. Rather than disentangling the effects of each factor on heterozygosity and then on enhancer strength, one may simply assess heterozygosity and enhancer strengths in different populations and see if they correlate positively. The detection of such a correlation would be an important clue towards confirmation of the occurrence of Enhancer Runaway and its impact on expression regulation evolution. To circumvent the problem of obtaining hybrids, it may be possible to use models calibrating enhancer strength from interactions with different *trans*-acting regulators as with *STARR*-seq methods (Muerdter et al. 2015).

TESTING ENHANCER DIVERGENCE

In some asexual species displaying very high levels of heterozygosity (apomixis, premeiotic doubling, central fusion without recombination with the centromeres), haploidization of expression should evolve. This prediction may be modulated depending on the degree of genetic isolation between homolog chromosomes. With central fusion automixis, recombination, and thus loss-of-heterozygosity really depends on the distance from the centromere. Hence, we predict that regions close to the centromeres should exhibit a stronger bias toward haploid expression. On the contrary, such divergence should not occur away from centromeres and the rate of Enhancer Runaway should also be small, owing to the low level of heterozygosity. Such genomes would therefore offer a very strong and simultaneous test of both Enhancer Runaway and Divergence if hybrids can be somehow obtained between a parthenogenetic and a related sexual species. More generally, in parthenogenetic lineages, haploidization should evolve whenever chromosomes become sufficiently isolated (i.e., with low enough rates of outcrossing, RWC, or conversion). Haploid expression in diploid species can be revealed using allele-specific expression assessing techniques, like RNA-seq or micro/oligo arrays. Enhancer Divergence should create a signal of large and widespread allele-specific expression. In central fusion automicts, allele-specific expression should decrease with increased genetic distance from the centromeres. Contrasting automictic species where conversion occurs preferentially on regulatory regions or not might also be a possibility, although such variability needs to be characterized first.

Another important empirical implication of Enhancer Divergence is that mutation accumulation in asexuals may not have a strong fitness impact overall if most of these mutations are completely or at least partially silenced. The load accumulated in asexuals may therefore largely depend on the coevolution of regulated and regulatory regions. For this reason, existing load models (e.g., Haag and Roze 2007) may require extension in this direction. Several studies have found elevated rates of nonsynonymous versus synonymous mutations (dN/dS) in asexuals, suggestive of the presence of a fitness load in asexuals (see review in Hartfield 2016). Most of these studies focus on mitochondrial genes, where the theory developed here is not directly relevant, but some also focus on nuclear genes (Henry et al. 2012; Hollister et al. 2015) where it can matter. Finding an elevated rate of accumulation of heterozygous deleterious mutations may not translate into a decline in fitness if those mutations are underexpressed. This could be investigated by also looking at the evolution of regulatory regions or by investigating whether alleles with lower expression (or more simply lower transcript abundance) are tied to a greater

burden of deleterious mutations. Silencing of the degenerated allele may also cause an ascertainment bias against finding divergence in, for example transcriptomics studies, which adds to the difficulties of studying asexual genomes.

Overall, we showed that the evolution of regulatory regions strongly depends on reproductive systems, beyond the stochastic effects (mutation, drift, selective interference) considered in classical population genetics models. These selective effects result from competition for expression of *cis*-regulatory region and lead to a blend of runaway and divergence, depending on genetic isolation of homologous chromosomes. These selective effects are not related to the optimization of expression levels due to organismal-level selection. Yet, they can largely influence the genome evolution in diploids and leave genome-wide footprints that may be revealed using large-scale genomic data.

AUTHOR CONTRIBUTIONS

F.F. and T.L. set up the project. F.F. implemented and ran the simulations. T.L. and F.F. interpreted the results and wrote the article.

ACKNOWLEDGMENTS

We thank R. Han and two anonymous reviewers for comments on the manuscript. This work was supported by a Ph.D. grant from French ministry of research to F.F., and the ANR grant GenAsex.

DATA ARCHIVING

The code is available on GitHub (https://github.com/FredFyon/Cis-regulator-runaway-and-divergence-in-asexuals).

LITERATURE CITED

Abzhanov, A., M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. Science 305:1462–1465.

Asher, J. H. 1970. Parthenogenesis and genetic variability. II. One-locus models for various diploid populations. Genetics 66:369–391.

Bell, G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. California Univ. Press, California.

Butlin, R. 2002. The costs and benefits of sex: new insights from old asexual lineages. Nat. Rev. Genet. 3:311–317.

Carroll, S. B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134:25–36.

———. 2005. Evolution at two levels: on genes and form. PLoS Biol. 3:e245.

Coolon, J. D., C. J. McManus, K. R. Stevenson, B. R. Graveley, and P. J. Wittkopp. 2014. Tempo and mode of regulatory evolution in *Drosophila*. Genome Res. 24:797–808.

Cooper, T. F., D. E. Rozen, and R. E. Lenski. 2003. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 100:1072–1077.

Denver, D. R., K. Morris, J. T. Streelman, S. K. Kim, M. Lynch, and W. K. Thomas. 2005. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. Nat. Genet. 37:544–548.

Fay, J. C., and P. J. Wittkopp. 2008. Evaluating the role of natural selection in the evolution of gene regulation. Heredity 100:191–199.

Ferea, T. L., D. Botstein, P. O. Brown, and R. F. Rosenzweig. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. Proc. Natl. Acad. Sci. USA 96:9721–9726.

Force, A., W. A. Cresko, F. B. Pickett, S. R. Proulx, C. Amemiya, and M. Lynch. 2005. The origin of subfunctions and modular gene regulation. Genetics 170:433–446.

Fyon, F., A. Cailleau, and T. Lenormand. 2015. Enhancer runaway and the evolution of diploid gene expression. PLoS Genet. 11:e1005665.

Ganko, E. W., B. C. Meyers, and T. J. Vision. 2007. Divergence in expression between duplicated genes in arabidopsis. Mol. Biol. Evol. 24:2298–2309.

Gay, J., S. Myers, and G. McVean. 2007. Estimating meiotic gene conversion rates from population genetic data. Genetics 177:881–894.

Gilad, Y., A. Oshlack, G. K. Smyth, T. P. Speed, and K. P. White. 2006. Expression profiling in primates reveals a rapid evolution of human transcription factors. Nature 440:242–245.

Haag, C. R., and D. Roze. 2007. Genetic load in sexual and asexual diploids: segregation, dominance and genetic drift. Genetics 176:1663–1678.

Haag, C., L. Theodosiou, R. Jabbour-Zahab, and T. Lenormand. 2017. Low recombination rates in sexual species and sex-asex transitions. Philos. Trans. R Soc. Lond. 372: https://doi.org/10.1098/rstb. 2016–0461.

Hartfield, M. 2016. Evolutionary genetic consequences of facultative sex and outcrossing. J. Evol. Biol. 29:5–22.

He, F., X. Zhang, J. Hu, F. Turck, X. Dong, U. Goebel, J. Borevitz, and J. de Meaux. 2012. Genome-wide analysis of cis-regulatory divergence between species in the *Arabidopsis* genus. Mol. Biol. Evol. 29:3385–3395.

Henry, L., T. Schwander, and B. J. Crespi. 2012. Deleterious mutation accumulation in asexual timema stick insects. Mol. Biol. Evol. 29:401–408.

Hoekstra, H. E., and J. A. Coyne. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995–1016.

Hollister, J. D., S. Greiner, W. Wang, J. Wang, Y. Zhang, G. K. S. Wong, S. I. Wright, and M. T. J. Johnsony. 2015. Recurrent loss of sex is associated with accumulation of deleterious mutations in oenothera. Mol. Biol. Evol. 32:896–905.

Innan, H., and F. Kondrashov. 2010. The evolution of gene duplications: classifying and distinguishing between models. Nat. Rev. Genet. 11:97–108.

Khaitovich P., G. Weiss, M. Lachmann, I. Hellmann, W. Enard, B. Muetzel, U. Wirkner, W. Ansorge, and S. Paabo. 2004. A neutral model of transcriptome evolution. PLoS Biol. 2:e132.

Kirkpatrick, M., and C. D. Jenkins. 1989. Genetic segregation and the maintenance of sexual reproduction. Nature 339:300–301.

Kuo, D., K. Licon, S. Bandyopadhyay, R. Chuang, C. Luo, J. Catalana, T. Ravasi, K. Tan, and T. Ideker. 2010. Coevolution within a transcriptional network by compensatory trans and cis mutations. Genome Res. 20:1672–1678.

Lenormand, T., J. Engelstädter, S. E. Johnston, E. Wijnker, and C. R. Haag. 2016. Evolutionary mysteries in meiosis. Philos. Trans. R Soc. B Biol. Sci. 371:20160001.

Lenormand, T., D. Roze, and F. Rousset. 2009. Stochasticity in evolution. Trends Ecol. Evol. 24:157–165.

Li, W.-H., J. Yang, and X. Gu. 2005. Expression divergence between duplicate genes. Trends Genet. 21:602–607.

Ludwig, M. Z., C. Bergman, N. H. Patel, and M. Kreitman. 2000. Evidence for stabilizing selection in a eukaryotic enhancer element. Nature 403:564–567.

Lynch, M. 2007. The evolution of genetic networks by non-adaptive processes. Nat. Rev. Genet. 8:803–813.

Lynch, V. J., and G. P. Wagner. 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution 62:2131–2154.

Mandegar, M. A., and S. P. Otto. 2007. Mitotic recombination counteracts the benefits of genetic segregation. Proc. R Soc. B Biol. Sci. 274:1301–1307.

Manna, F., R. Gallet, G. Martin, and T. Lenormand. 2012. The high-throughput yeast deletion fitness data and the theories of dominance. J. Evol. Biol. 25:892–903.

Mittwoch, U. 1978. Parthenogenesis. J. Med. Genet. 15:165–181.

Muerdter, F., Ł. M. Boryń, and C. D. Arnold. 2015. STARR-seq—principles and applications.Genomics 106:145–150.

Nougué, O., N. O. Rode, R. Jabbour-Zahab, A. Ségard, L.-M. Chevin, C. R. Haag, and T. Lenormand. 2015. Automixis in Artemia: solving a century-old controversy. J. Evol. Biol. 28:2337–2348.

Oleksiak, M. F., G. A. Churchill, and D. L. Crawford. 2002. Variation in gene expression within and among natural populations. Nat. Genet. 32:261–266.

Otto, S. P., and M. E. Orive. 1995. Evolutionary consequences of mutation and selection within an individual. Genetics 141:1173–1187.

Raymond, M., C. Chevillon, T. Guillemaud, T. Lenormand, and N. Pasteur. 1998. An overview of the evolution of overproduced esterases in the mosquito *Culex pipiens*. Philos. Trans. R Soc. Lond. Ser. B Biol. Sci. 353:1707–1711.

Schön, I., K. Martens, and P. van Dijk. 2009. Lost sex the evolutionary biology of parthenogenesis. Springer, Dordrecht.

Schurko, A. M., M. Neiman, and J. M. Logsdon. 2009. Signs of sex: what we know and how we know it. Trends Ecol. Evol. 24:208–217.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jónsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717–723.

Steige, K. A., J. Reimegård, D. Koenig, D. G. Scofield, and T. Slotte. 2015. Cis-regulatory changes associated with a recent mating system shift and floral adaptation in capsella. Mol. Biol. Evol. 32:2501–2514.

Suomalainen, E., A. Saura, and J. Lokki. 1987. Cytology and evolution in parthenogenesis. CRC Press, Boca Raton, FL.

Tautz, D. 2000. Evolution of transcriptional regulation. Curr. Opin. Genet. Dev. 10:575–579.

Weirauch, M. T., and T. R. Hughes. 2010. Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. Trends Genet. 26:66–74.

Whitehead, A., and D. L. Crawford. 2006. Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. 103:5425–5430.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85–58.

Wittkopp, P. J., and G. Kalay. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. 13:59–69.

Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8:206–216.

Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano. 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20:1377–1419.

Yanai, I., D. Graur, and R. Ophir. 2004. Incongruent expression profiles between human and mouse orthologous genes suggest widespread neutral evolution of transcription control. Omi. AJ. Integr. Biol. 8:15–24.

Associate Editor: M. W Sayres
Handling Editor: P. Tiffin

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website: