Infection, Genetics and Evolution 17 (2013) 123–131

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Purifying selection after episodes of recurrent adaptive diversification in fungal pathogens

Pierre Gladieux [a,b,c,*,1], Benjamin Devier [a,b,c], Gabriela Aguileta [a,b,c,2], Corinne Cruaud [d], Tatiana Giraud [a,b,c]

[a] *Laboratoire Ecologie, Systématique et Evolution, Univ. Paris-Sud, F-91405 Orsay cedex, France*
[b] *CNRS, Orsay, France*
[c] *AgroParisTech, Orsay, France*
[d] *Genoscope, Centre National de Sequençage, UMR CNRS 8030, 2 Gaston Crémieux, CP 5706, 91507 Evry, France*

ARTICLE INFO

Article history:
Received 7 November 2012
Received in revised form 26 February 2013
Accepted 7 March 2013
Available online 10 April 2013

Keywords:
Approximate Bayesian computation
Adaptive divergence
Speciation
Parasite
Anther smut
Microbotryum violaceum

ABSTRACT

Tests of selective neutrality based on comparative and population genetic data target different time scales of selection. Investigating both time scales on the same genes has rarely been done and yet can provide insights into histories of selection. In pathogens, such a comparison can elucidate whether the same genes experience recurrent positive selection across specialization events onto novel hosts and selection in the short term, *e.g.* by coevolution with a current host. Here, we investigated whether the genes showing signs of positive selection (based on d_N/d_S ratios) between pathogens specialized on different hosts also exhibited footprints of recent positive selection. We sequenced 58 DNA fragments in the model fungal plant pathogen *Microbotryum lychnidis-dioicae* infecting *Silene latifolia*. Eleven focal genes were chosen because they showed a significant signal of positive selection in a previous study and putative functions likely involved in host–parasite interactions. In addition, 47 control genes were randomly chosen among genes showing no evidence of positive selection. We used a population genetics approach to search for signatures of recent selection, controlling for population structure and demographic history. We found footprints of purifying selection in all focal genes and of recent positive selection in two of them. Signs of purifying selection were also found at the same eleven genes in a closely related species, *Microbotryum silenes-dioicae*, infecting *Silene dioica*. These results suggest that genes experiencing episodes of adaptive diversification during host shifts may subsequently be under strong functional constraint, although some can remain under positive selection in the short term.

1. Introduction

Pathogens are attractive systems for the study of adaptive divergence. Numerous pathogens appear as multiple recently diverged sibling species specialized on different hosts, allowing investigation of the early stages of specialization and speciation. Strong selection pressures on pathogens to exploit new hosts is prone to provide spectacular examples of discontinuous evolution, with rapid and major shifts in phenotypic traits caused by drastic molecular changes at relatively few loci but leading to the emergence of new specialized parasite species (Giraud et al., 2010). Within pathogen species, coevolution with hosts also offers unique opportunities to study adaptation. Coevolution can result in an 'arms race', *i.e.* the occurrence of recurrent selective sweeps because novel resistance and virulence alleles are favored, or in a 'trench warfare', where balancing selection maintains stable and long standing polymorphism in genes involved in host–pathogen recognition. Linking these two scales of adaptation, *i.e.* intraspecific coevolution with a host and long-term specialization and speciation on novel hosts, has very rarely been addressed, and is our aim in the present study. Of particular interest is to determine whether the genes that have been under diversifying selection between species specialized onto different hosts also undergo positive selection in the short term, *e.g.* by being involved in the recurrent coevolution of a pathogen with its current host (Giraud et al., 2010; Hörger et al., 2012; Schulze-Lefert and Panstruga, 2011).

Comparative studies involving divergence data among multiple different species (*e.g.* based on the ratio of nonsynonymous to synonymous substitution rates, d_N/d_S) have provided unambiguous evidence for the occurrence of positive selection (Nielsen, 2005),

---

* Corresponding author at: Laboratoire Ecologie, Systématique et Evolution, Univ. Paris-Sud, F-91405 Orsay cedex, France. Tel.: +33 169157049; fax: +33 169154697.

E-mail addresses: pierre.gladieux@u-psud.fr (P. Gladieux), bendevier@gmail.com (B. Devier), gabriela.aguileta@crg.eu (G. Aguileta), cruaud@genoscope.cns.fr (C. Cruaud), tatiana.giraud@u-psud.fr (T. Giraud).

[1] Current address: Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102, United States.

[2] Current address: Comparative Bioinformatics Center for Genomic Regulation, Parc de Recerca Biomèdica, C/Dr. Aiguader, 88, Barcelona 08003, Spain.

1567-1348/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.meegid.2013.03.012

in particular in pathogens (Aguileta et al., 2009; Stukenbrock and McDonald, 2009). In pathogens, most studies have focused on specific likely candidate genes, such as those coding for effector molecules, antigenic proteins, or drug resistance. More recently, the flood of genomic data has allowed a surge in 'reverse ecology' studies searching for genes under positive selection in large datasets (Ellison et al., 2011; Li et al., 2008). The advantage of these approaches is that they can identify genes without a priori expectations, and thus generate new hypotheses for further investigations about the role of certain molecular processes in pathogen adaptation. Such comparative approaches are best suited for detecting selection that operates recurrently, generating repeated adaptive changes occurring in the same gene or even at the same site in different variants (Nielsen, 2005).

Inferences regarding recent selection require polymorphism data. Population genetics/genomics approaches often ask whether patterns of polymorphism within species and differences among species are compatible with expectations under a standard equilibrium neutral model in which molecular variation is mostly governed by genetic drift. These methods are more powerful than comparative genomics for detecting episodes of recent selection, when the excess of non-synonymous to synonymous differences is not great enough to be significant using d_N/d_S. Because of their differences in scope, there is no reason to expect that the genes showing signs of selection would be the same in comparative and population genetics surveys. Nonetheless, some genes involved in the coevolution between hosts and pathogens may in fact experience both short-term and long-term selection and thus also be involved in specialization onto novel hosts (Giraud et al., 2010; Schulze-Lefert and Panstruga, 2011). This may be the case in particular for effectors in pathogens that may be under selection for ongoing changes to avoid recognition by the host or to maintain recognition of the host susceptibility target (Hörger et al., 2012; Stukenbrock and McDonald, 2009). Such genes would be expected to show signs of selection both with comparative and population genetics approaches. In contrast, genes experiencing repeated but transitory events of selection, for instance solely during specialization onto novel hosts, should show signatures of positive selection only in the comparative approach, while appearing under purifying selection within species. Despite these clear distinguishing predictions there has been little effort expended to date to compare the targets of selection identified using these alternative approaches, in pathogens and in general (but see Begun et al. (2007)). Yet, such comparisons could grant insight into the impact of selection at the two scales on different types of genes as well as identifying genes that are important at only one or at the two scales.

Here, we use a population genetic approach controlling for the impact of population structure and demographic history on allelic frequency spectrums to test whether genes identified as showing clear signs of adaptive divergence between pathogen species adapted to different hosts also show signatures of selection at the within-species level. The model organism is the pathogen *Microbotryum*, a fungal species complex causing anther-smut disease on many plant species in the Caryophyllaceae (Hood et al., 2010). Diploid teliospores of *Microbotryum* are produced in anthers of diseased plants and are dispersed by pollinators. *Microbotryum violaceum sensu lato* is composed of several sibling species highly specialized on their host plants (Le Gac et al., 2007). *Microbotryum* fungi are model organisms for a wide range of questions in evolution and ecology, such as speciation, mating systems, disease transmission, metapopulation dynamics, mating type and sex chromosome evolution, local adaptation and virulence evolution (Bernasconi et al., 2009).

In a previous study, Aguileta et al. (2010) used comparative genomics to identify genes under positive selection in Expressed Sequence Tag (EST) sequences from each of four *Microbotryum* species specialized on different host plants: *Microbotryum lychnidis-dioicae* on *Silene latifolia*, *Microbotryum lagerheimii* on *Silene vulgaris*, *Microbotryum dianthorum* on *Dianthus carthusianorum*, and *Microbotryum violaceum sensu stricto* on *Silene nutans*. Forty-two of the 372 studied clusters of orthologs showed a significant signal of positive selection, and the sequencing of 16 of these genes in 9 additional *Microbotryum* species confirmed that they have indeed been rapidly evolving between species specialized on different hosts. The 42 genes were putatively involved in secondary metabolite synthesis and secretion, respiration under stressful conditions and stress response, hyphal growth and differentiation, regulation of expression by other genes, and many had transmembrane domains. These putative functions are consistent with the idea that the genes showing significant signals of diversifying selection between pathogen species specialized on different hosts are likely involved in host specialization.

In the present study, we focused on the species *M. lychnidis-dioicae* (Denchev et al., 2009); called MvSl in Le Gac et al. (2007) parasitizing the plant *S. latifolia*. This species is highly selfing, undergoing mostly intra-tetrad mating in nature (*i.e.* automixis, Giraud et al., 2008; Gladieux et al., 2011). We sequenced 11 of the genes showing significant signals of positive selection in Aguileta et al. (2010) in 35 individuals of *M. lychnidis-dioicae*. These 11 focal genes were chosen because they showed a significant signal of positive selection in d_N/d_S analyses and putative functions likely involved in host–parasite interactions. We also analyzed polymorphism at 47 control genes chosen at random among the genes showing no evidence of positive selection in Aguileta et al. (2010). We also run similar analyses, albeit on a smaller dataset, on *Microbotryum silenes-dioicae* (Denchev et al., 2009); called MvSd in Le Gac et al. (2007) parasitizing the plant *Silene dioica* (the dataset included 17 control DNA fragments as control genes; and the same 11 focal genes as for *M. lychnidis-dioicae*). Our specific aims were (i) to elucidate the population structure and recent demographic history of the species, as these factors can bias inferences of natural selection; a clear phylogeographic structure has been previously detected in both species in Europe, with highly differentiated clusters corresponding to the footprints of glacial refugia (Gladieux et al., 2011; Vercken et al., 2010), but the demographic history of the clusters has not been investigated yet; (ii) to determine whether the genes that have experienced recurrent codon changes between species, as evidenced from the comparative analysis, also show signatures of recent positive selection, or in contrast exhibit footprints of purifying selection, using tests based on the site frequency spectrum of DNA polymorphisms within clusters, controlling for population structure and demographic history, and tests based on contrasts of the levels of polymorphism and divergence.

2. Materials and methods

2.1. Data generation

Analyses were performed on 35 individuals of *M. lychnidis-dioicae* (MvSl) and 12 individuals of *M. silenes-dioicae* (MvSd) (detailed sample information available in Supplementary Table S1). Diploid teliospores were isolated from infected anthers and cultured on synthetic medium for DNA extraction as previously described (Le Gac et al., 2007). For strains that could not be cultured, DNA was extracted as in Giraud (2004) directly from anthers full of diploid teliospores using Chelex (Biorad). The diploid teliospores collected in an infected anther will be considered to represent the fungal "individual". Individual plants of *Silene* can be infected by several *Microbotryum* strains, but individual flowers always carry a single genotype (López-Villavicencio et al., 2007).