



HAL
open science

Polygenic Adaptation: Integrating Population Genetics and Gene Regulatory Networks

Maud Fagny, Frederic Austerlitz

► **To cite this version:**

Maud Fagny, Frederic Austerlitz. Polygenic Adaptation: Integrating Population Genetics and Gene Regulatory Networks. Trends in Genetics, 2021, 37 (7), pp.631-638. 10.1016/j.tig.2021.03.005 . hal-03100982

HAL Id: hal-03100982

<https://hal.science/hal-03100982>

Submitted on 6 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Understanding the adaptation of polygenic traits: the importance of gene regulatory networks

Maud Fagny^a, Frédéric Austerlitz^a

^a*UMR7206 Eco-Anthropologie, Muséum National d'Histoire Naturelle, Centre National de la Recherche Scientifique, Université de Paris, Paris, France*

Abstract

The adaptation of populations to local environments often relies on the selection of beneficial versions of polygenic traits. At the molecular level, this process corresponds to polygenic selection models, in which many independent loci, often located in regulatory regions, experience subtle shifts in frequency. Here, we first review the different approaches developed to detect signatures of polygenic selection, their power and limitations. We then highlight the contribution of systems biology to the understanding of the molecular bases of polygenic traits and the evolution of gene regulatory networks involved in these traits. Finally, we discuss the need for a unifying framework merging the fields of population genetics, quantitative genetics and systems biology to better understand the molecular bases of polygenic traits adaptation.

Keywords: Quantitative traits, Polygenic selection, Gene regulatory networks, Systems Biology, Adaptive evolution, Population genetics modelling

1 Highlights

- 2 ● In many species, adaptation of populations to local environment relies on traits that have a
3 polygenic architecture.
- 4 ● Statistical approaches have been developed recently to specifically detect different aspects of
5 the weak genomic signatures left by directional polygenic selection.
- 6 ● Genome-wide scans for selection have revealed that polygenic selection is a driving force of
7 local adaptation and disproportionately targets regulatory regions, hinting for an important
8 role of gene regulatory networks in evolution.
- 9 ● Gene regulatory network inference helps identifying and grouping together genes and regu-
10 latory elements that participate to the same biological processes.

- 11 • The topology of gene regulatory elements put constraints on which genes and regulatory
12 regions can be leveraged by polygenic selection.

13 **Local adaptation of polygenic traits: the importance of regulatory regions**

14 Many adaptive phenotypes – *i.e.* phenotypes involved in adaptation to local environments – are
15 polygenic (see Box 1). Also called quantitative or complex traits, these phenotypes are determined
16 by multiple genes and regulatory loci. This complex genetic architecture has strong implications
17 on how these traits evolve. If these traits are under selection (either stabilising or directional), the
18 underlying genetic model will be a model of polygenic selection, *i.e.* subtle shifts in frequencies at
19 many independent loci coding for each adaptive trait [1–3]. This is for example the case for height,
20 immune response, and metabolism in humans [4, 5], flowering time, and stress response in plants,
21 [6–8], or pathogenicity in yeast [9]. Moreover, in humans, a lot of disorders such as autoimmune
22 diseases and metabolic disorders are considered to be consequences of past adaptation of polygenic
23 traits to changes in local environments, such as pathogens or lifestyle transition [10, 11]. Being
24 able to detect polygenic adaptation is thus important not only to understand the evolution of
25 polygenic traits, but also the origin of many human diseases.

26 Our apprehending of the effects of directional selection on polygenic traits at the molecular
27 level has strongly improved in the past ten years. QTL and genome-wide association studies
28 (GWAS), have led to the identification of genetic variants associated with these polygenic traits,
29 which were located overwhelmingly in non-coding regions [12, 13] – up to 90% in some species
30 – highlighting the importance of regulatory regions. Population genetics studies, leveraging the
31 ever growing number of genomic data sets available and the improvements in genomic functional
32 annotation have confirmed that local adaptation often relies on selection of mutations located in
33 regulatory regions that impact gene expression levels [9, 14]. This has spurred efforts to develop
34 population genetic approaches to detect signatures of selection beyond coding genomic regions.
35 Independently, these results helped creating new theoretical models, in which polygenic traits
36 result from a complex interplay of regulatory relationships between genes and different molecular
37 actors at the cellular level. This led to the development of systems biology models aiming at
38 understanding how these regulatory relationship evolve under directional selection, which brought
39 new clues to understand the molecular bases of polygenic traits adaptation to local environments.

40 **Detecting genomic signatures of polygenic selection: an ongoing quest**

41 During the past few years, population geneticists have attempted to develop approaches to
42 detect signatures of polygenic selection on genomic data. The first difficulty is to identify the type
43 of signatures that one must look for. For this purpose, models have been developed to investigate
44 the expected impact of polygenic adaptation on population genomic data [3, 15–19] (see Box 2).
45 They assume additive and/or epistatic interactions among loci. These studies have shown mainly
46 that polygenic adaptation is likely to yield subtle frequency shifts at the genes involved in the
47 polygenic trait: polygenic adaptation events target several mutations simultaneously and leave
48 weak molecular signatures on the genome [20]. Recently, Hollinger and colleagues have however
49 shown that the type of genomic signature observed ranges from sweeps (either total or partial) to
50 subtle frequency shifts, depending on the population-scaled mutation rate [3].

51 Methods to detect genomic signatures of polygenic selection events have mainly focused on
52 combining signatures of positive selection across sets of mutations that are defined *a priori* as likely
53 candidates for adaptive pressures. In particular, some studies have proposed to group mutations for
54 which allele frequencies correlate with specific ecological variables [4, 21]. These approaches, while
55 powerful, require access to past and present environmental data for the populations under study,
56 which are often difficult to obtain. Other approaches use gene pathways and information about
57 biological functions or GWAS results to group mutations together [5, 8, 22]. The first approach
58 consists in investigating whether known pathways or gene sets are particularly enriched in selection
59 signals [5], with many loci exhibiting a high level of genetic differentiation among populations. A
60 recently developed related approach looks for local subnetworks under strong differentiation [23].
61 The second approach consists in testing whether a known trait has been submitted to divergent
62 selection [22, 24–26]. These methods require that previously performed genome-wide association
63 studies (GWAS) are available, which have detected loci associated with the traits and estimated
64 their additive effect on these traits. This allows computing the polygenic score of the genotyped
65 individuals in the studied populations, which can then be compared to what is expected under
66 genetic drift only. Because they rely on already existing GWAS studies, these methods cannot be
67 used blindly on the genome.

68 On the other hand, other methods have focused on scanning the genome in order to detect loci
69 with high level of differences in allele frequencies among populations, without *a priori* grouping
70 mutations. This is the case for example of the SDS approach [27], OutFLANK [28] and PCAdapt

71 [29]. These genome-scan methods can be used blindly on the genome, however they were mainly
72 developed in a context of strong selection on a single locus and simulation studies have shown that
73 their efficiency for detecting loci evolving under polygenic selection is rather limited [30, 31].

74 **Shortcomings of polygenic selection detection approaches**

75 The shortcomings of existing approaches to detect polygenic selection have been extensively
76 discussed in recent papers [32–35] that have raised several caveats concerning the results of these
77 polygenic selection tests. The first set of caveats are technical and relate primarily to GWAS-based
78 results: if not correctly taken into account, a hidden population structure in the samples used for
79 GWAS can lead to the false detection of directional polygenic selection signal on the trait under
80 study. This has been demonstrated for height in European human populations [34, 35]. Similarly,
81 GWAS ascertainment biases can confuse selection signals [36]. Transferring GWAS results across
82 time and space can also lead to misidentification of causal mutations, the impact of which on the
83 selection tests is uncertain.

84 The last set of caveats relates to the interpretation of test results. They include the potential
85 existence of gene x environment interactions and the poorly known interactions between traits
86 [32], both of which can hide the selection signals or inverse the direction of selection. Finally,
87 an important part of the selection signals might be lost in current studies because a fundamental
88 component of the genetic architecture of polygenic traits is often minimized or totally ignored:
89 epistatic interactions [37]. On this topic, Payne and colleagues recently highlighted the impor-
90 tance of studying the topography of the adaptive landscape and, in particular, the exceptional
91 characteristics of gene expression regulatory networks in terms of epistatic interactions, pleiotropy
92 and robustness [38].

93 To this list of caveats, we can add another one: while carrying the majority of the genetic
94 variation associated with polygenic traits, the regulatory regions are a blind spot of the methods
95 designed to detect polygenic selection described above. These approaches, which combine signa-
96 tures of positive selection across loci associated to a given trait, often only target genes, completely
97 ignoring regulatory regions. The recently developed approaches based on GWAS results include
98 some genetic variants from regulatory regions. However, because they only include the variants
99 that are statistically significantly associated to the phenotype at the genome-wide level, they miss a
100 good part of the genetic bases of polygenic phenotypes – the famous “missing heritability” [39, 40].

101 On the other hand, blind genome-wide scans have detected the signatures of polygenic selection
102 in regulatory regions. Several studies have indeed revealed an enrichment in positive selection sig-
103 nals among cis-regulatory elements and expression QTLs [41–43]. However, the lack of functional
104 annotation of these regions often prevents identification of targeted phenotypes, beyond general
105 results of Gene Ontology enrichment assays.

106 In short, despite recent advances in polygenic selection signature testing, several difficulties
107 persist. While the technical caveats point to the need to carefully consider potential confounding
108 effects and to assess the sensitivity and specificity of polygenic selection tests, most of the difficul-
109 ties in interpreting polygenic selection tests can be explained by the “genomic only” approaches
110 used in these tests. These difficulties indeed arise from our limited knowledge of the true genetic
111 architecture and molecular bases of polygenic traits. These traits are determined by various reg-
112 ulatory regions, and depend on a complex interplay between different actors that are expressed
113 following a rigid spatio-temporal framework. Ignoring these features makes it difficult to interpret
114 results of polygenic selection tests, and to identify the biological functions and traits potentially
115 under polygenic selection.

116 **Embracing the Complexity: Contributions from Systems Biology**

117 Organizing and integrating information from diverse sources and understanding how polygenic
118 phenotypes are generated from the genome at different levels, from the cell to the organism, is the
119 core of systems biology [44]. The rise of this research field initiated a switch in the perception of
120 the molecular bases of polygenic traits, from a gene-first to an interaction-first model. Quantita-
121 tive genomics apprehends indeed the molecular bases of polygenic phenotypes as a collection of
122 independent genes, each coding for a fraction of the phenotype, with potentially some interactions.
123 Conversely, systems biology focuses primarily on these interactions, modelling the regulatory rela-
124 tionship between genes, their regulators and their products. During the past twenty years, buoyed
125 by the availability of high-throughput data, numerous approaches have been developed to model
126 these regulatory relationships and to efficiently deal with their complexity. In particular, the
127 representation of these relationships as Gene Regulatory Networks (GRNs), which gather tens to
128 thousands of genes, has allowed better understanding the molecular bases of polygenic traits [45].

129 From the systems biology point of view, understanding the molecular bases of polygenic se-
130 lection consists, therefore, in deciphering how regulatory interactions among genes evolve under

131 constraints. Models have been developed to study how GRNs evolve, which explicitly simulates
132 these interactions. Initially developed by Wagner, they model gene networks through a matrix
133 whose coefficients determines how each gene regulates each other and itself [46, 47]. These co-
134 efficients vary among individuals and can evolve through mutations that change the regulatory
135 relationships between genes. Each matrix leads to a given level of expression of each locus. Selec-
136 tion will occur through the assumption that there is an optimal level of expression for each gene.
137 This model was first developed to study the evolution of gene duplications [46].

138 It allowed then discovering some properties of GRN evolution. One of the major results is
139 the demonstration that, under stabilizing selection, network evolves toward a state where a single
140 mutation becomes less likely to reduce the fitness of the individuals [47], a phenomenon called
141 canalisation [48]. This insensitivity to mutations becomes stronger when the number of connections
142 in the network increases [49] or when phenotypic optima become more extreme [50]. Conversely,
143 this robustness of the network promotes the accumulation of hidden loss-of-function mutations,
144 which can then accelerate adaptation in a changing environment [51].

145 GRN evolution models have also helped better understanding how the structure of these GRNs
146 impacts the individual response of each gene to selection, as a function of selection intensity and
147 distance between the current phenotype and the optimal one. Genes that are more upstream in
148 the network (i.e. strongly regulating the other genes and/or less regulated by them) are more
149 affected by selection [52]. This could explain the empirical observation that transcription factors
150 are evolving faster than the other genes [53].

151 **Impact of GRN structure on architecture of polygenic selection**

152 Understanding the evolution of GRNs that underlies polygenic phenotypes is currently a hot
153 topic in evolutionary biology. The various scenarii of GRN evolution observed have been summa-
154 rized for animals, insects [54] and plants [55]. Modifications of GRNs seems to be frequent during
155 speciation [56] and domestication [57, 58], and can lead to dramatic changes in developmental
156 patterns and other polygenic phenotypes. Moreover, there are many examples of directional and
157 polygenic selection events affecting the gene regulatory networks [8, 9, 43, 59]. The analysis of
158 cellular GRN and their evolution has led to hypotheses about how polygenic selection can act,
159 and how gene regulatory network structure might facilitate or constrain it. Three GRN features
160 are worthy of attention in this context: redundancy, modularity, and their scale-free property (see

161 Figure 1 and Box 3).

162 Redundancy is a characteristics of GRNs that ensures stability of gene expression, buffers noise
163 and ensures phenotype robustness to random mutations that may cause gene inactivation [50, 60].
164 It has been shown also to facilitate polygenic adaptation. For example, an evolution experiment in
165 *Drosophila* showed phenotypic convergence in response to a temperature change in ten populations,
166 while the genetic analysis revealed that different pathways were targeted by polygenic selection
167 in each replicate [61]. Redundancy of GRNs thus facilitated polygenic selection by increasing
168 the number of available targets to reach the optimum phenotype (see Figure 1B). However, this
169 polygenic selection will be difficult to detect, as it will potentially affect different pathways among
170 populations confronted to the same constraints.

171 GRNs are also highly modular, i.e., made of small groups of genes strongly interconnected with
172 each other but only loosely connected to the rest of the network, thus constituting a module. It
173 has been proposed, in the omnigenic model, that as all genes are seemingly interconnected within
174 the GRN, each trait might be influenced by all genes [40]. Theoretically, such extreme generalized
175 pleiotropy – property of a gene that influences several, apparently unrelated traits – would severely
176 constrain the possibility for polygenic selection to act, because it could not modify a given trait
177 without influencing all the others. However, the modularity of GRNs can mitigate this effect.
178 Indeed, in each module, only a few genes are connected with genes from other modules. Most
179 genes in a module are thus isolated from the rest of the network and can be leveraged by polygenic
180 selection (see Figure 1A).

181 Finally, GRNs tend to be scale-free, i.e. there are many more genes that are connected to
182 only a few other genes than there are genes connected to lots of other genes [62] (see Figure 1B).
183 This has been shown on various types of networks [45, 62, 63]. As confirmed by observations on
184 molecular data (transcriptomics or proteomics), the level of connection of a gene in a network is
185 inversely correlated with its level of genetic diversity [64–66]. This means that the handful of highly
186 pleiotropic genes at the center of these networks often evolve under strong negative selection and
187 present few variations of expression among individuals. Conversely, the vast majority of genes are
188 at the periphery of the network, involved in only one or few pathways. Changes in their regulation
189 will, therefore, only very slightly affect the other pathways (see Figure 1A and B). This leaves
190 open a large field full of candidates for polygenic selection to act on, both at the level of the gene
191 sequence and the gene regulatory elements, while limiting the potential side effects of selection on

192 other pathways.

193 **Polygenic selection and GRNs: advantages and limits**

194 Theoretical models of GRNs combined with results on real data sets allow better understanding
195 the molecular bases of polygenic phenotypes and their evolution. They also help understanding
196 where and how polygenic selection can act, and where to look for genomic signatures of selection.
197 However, the models often focus on the evolution of unipartite networks of genes interacting
198 with each other, thus overlooking important features that would be useful to get a complete
199 picture of how polygenic selection can impact GRNs. First, despite the evidences for pervasive
200 polygenic selection targeting regulatory mutations [9], these models do not include explicitly
201 how modifications in non-coding regulatory regions, such as enhancers, can impact the expression
202 level of several genes at once, while not modifying their interactions. An important source of
203 polygenic trait variation and evolution is thus not taken into account. Some current approaches
204 integrate heterogeneous data using multipartite networks, in order to identify not only gene-gene
205 regulatory interactions but also interactions between regulatory regions and genes [45, 67, 68].
206 These approaches could serve as bases for the development of new GRN evolution models. Second,
207 most studies on GRN evolution do not model individual nucleotides. Consequently, they allow to
208 study questions such as redundancy, pleiotropy and canalisation, but not to dissect the genetic
209 bases of evolution, and to understand the genomic signatures that may be expected in different
210 polygenic selection scenarios.

211 **Concluding Remarks and Future Perspectives: towards systems population genetics**

212 In order to detect polygenic selection and to disentangle the impact of several confounding fac-
213 tors, we need to better understand not only the genetic architecture of polygenic traits, but also
214 their molecular architecture. A new integrative evolutionary framework becomes quite indispens-
215 able, which would take this architecture into account, in particular its redundancy and the fact that
216 two selective events in the same population might not lead to the same selection scenario at the
217 molecular level [69]. Here, we argue that, in order to achieve such a unifying framework, we need
218 not only to identify which genes and pathways contribute to polygenic traits, and how redundant
219 they are, but also investigate the complex regulatory interactions that govern their expression (see
220 Outstanding Questions). In particular, we need to take into account how the genomic regulatory

221 regions impact the expression of these traits. We should focus on the crucial post-transcriptional
222 regulatory steps, on the gene-gene interactions, and maybe more importantly, on the position of
223 these genes in the cellular gene regulatory network. All of this could be achieved by combining
224 population genomics and system biology approaches.

225 **Outstanding Questions**

- 226 • Redundancy allows for the selection of the same version of a trait through the selection of
227 genetic variants in different pathways. Do we observe this phenomenon at the scale of the
228 population? In that case, which molecular signatures do such evolutionary processes leaves
229 on the genome and how can we detect them?
- 230 • Pleiotropy is a major property of gene regulatory networks actors that constrains how poly-
231 genic selection acts. Genes are often involved in several distinct pathways. What is the level
232 of pleiotropy of the regulatory elements, in particular of the distal ones such as enhancers
233 and how does it constrain their evolution?
- 234 • Classical population genetics approaches model the evolution of genetic sequences but ignore
235 their complex interactions in gene regulatory networks. On the other hand, gene regulatory
236 networks model the evolution of these regulatory relationships but not their genetic bases.
237 Will more integrated approaches that explicitly model the molecular evolution of regulatory
238 regions in the context of gene regulatory networks improve our understanding of the molecular
239 bases of local adaptation?
- 240 • Selection on polygenic traits yields linkage disequilibrium among genes involved in the traits,
241 at the within- and among-population levels [15]. How do regulatory interactions affect this
242 linkage disequilibrium? Can it be used to detect polygenic selection?
- 243 • GWAS-based method aiming at detecting polygenic selection use the additive effect of each
244 gene on the trait. Could the interaction (dominance, epistasy) between these genes also be
245 integrated in these approaches?

246 **Box 1. Genetic architecture of traits: from the Mendelian to the omnigenic model**

247 Polygenic traits have been the focus of many theoretical and empirical studies since Fisher's
248 seminal model of polygenic inheritance, which bridged the gap between the Mendelian inheritance
249 model and the observation of continuous traits. This model suggests that the apparent continuity
250 of many phenotypic traits can be explained if multiple genes contribute additively to these traits
251 [70]. Classic genetic studies, recent advances in quantitative genetics and genome-wide association
252 studies (GWAS) have demonstrated the existence of a continuum of genetic architectures for
253 traits, in many organisms including yeast, insects, worms, plants and mammals [39, 71–73]. These
254 architectures range from the Mendelian model, where a trait is determined by a single gene (e.g.
255 ABO blood groups in primates, Mendel's peas color and shape) to the infinitesimal model [74]
256 where a seemingly infinite number of loci are involved in the determinism of a trait (e.g. height,
257 in both humans and plants [24, 75]). This last concept has been highlighted in a recent paper
258 [40] that labelled it the 'omnigenic' model. In between, we find a myriad of traits that present
259 an oligogenic or polygenic determinism, such as flowering time in plants, tolerance to ethanol in
260 yeasts or susceptibility to diabetes in humans [75–78].

261 **Box 2. Polygenic selection**

262 Polygenic selection consists in natural or artificial selection occurring on a trait coded by several
263 loci. Under this framework, even when the phenotypic optimum differs strongly from the current
264 average phenotype, populations can evolve toward this optimum without substantial changes in
265 allelic frequencies [15]. Indeed, each locus can undergo only swift changes in these frequencies,
266 while allelic associations will appear among the different loci, allowing to reach the phenotypic
267 optimum. In this case, methods aiming at detecting signals of selective sweeps such as *iHS* [79]
268 or *nSL* [80] will probably not detect any signal, as no locus is undergoing a strong shift in allelic
269 frequencies. Note, however, that this will depend on the number of loci involved in the trait: if
270 this number is low, each locus will respond more strongly in terms of allelic frequencies changes
271 and some might be detected. Moreover, if among the loci coding for the trait, some are major loci
272 with a high contribution to genetic variance, these loci are more likely to be affected by a strong
273 shift in allelic frequencies [16, 17], and thus to be detected by selective sweeps tests.

274 Other approaches modeled the fate of a locus coding for a given trait, while all other loci
275 coding for this trait were treated as background genetic variance [18]. They find that a selective

276 sweep will be observed in some cases at the focal locus, depending upon its relative variance as
277 compared to the background loci and also upon the shape of the selection function. In a model that
278 considered also epistatic interactions, Hollinger and colleagues then showed that the probability of
279 observing a sweep at a locus involved in a quantitative trait will depend mainly on a parameter
280 called the population-scaled background mutation rate Θ_{bg} , i.e. the product of the effective size of
281 the population by the total mutation rate of all background loci [3]. Low values for Θ_{bg} will lead
282 to the occurrence of selective sweeps at the focal locus, while high values will yield small polygenic
283 allele frequency shifts. Several partial selective sweeps will be observed for intermediate values.

284 **Box 3. Gene Regulatory Networks**

285 Gene Regulatory Networks (GRNs) describe a collection of molecules that interact together to
286 regulate the level of expression of genes, both as mRNA and proteins. These regulatory relation-
287 ships can be direct or indirect, for example when a transcription factor increases another protein
288 level through the activation of its coding gene transcription.

289 GRNs are often represented as graphs, mathematical objects that represent the elements
290 (nodes) and the interactions between elements (edges). They can include one type of elements
291 such as proteins (unipartite graphs), or several types of elements such as transcription factors, reg-
292 ulatory mutations and mRNAs (multipartite graphs). The study of the topology of these graphs
293 allows better understanding how the elements interact together and how GRNs evolve.

294 GRNs, whether in yeasts, insects, plants or mammals, show some shared topological charac-
295 teristics. Most importantly, they are scale-free. This means that very few of the nodes within the
296 network are connected to many others — the key or master regulators — while most of the nodes
297 are connected only to a few other nodes (see Figure Figure 1B). GRNs are also highly modular,
298 i.e. organized into groups of nodes that are strongly interconnected but only loosely connected to
299 the rest of the network, thus constituting a module. The modules often correspond to biological
300 pathways.

301 The topology of the network also seems to constrain the evolutionary history of each node. In
302 GRNs, the handful of high-degree genes, i.e. connected to many other genes, are often known to
303 be pleiotropic, involved in many pathways, and to be very constrained [64], evolve under strong
304 negative selection and present few variations of expression among individuals. Conversely, periph-
305 erical genes, which are only connected to few other genes, are often included in only one pathway.

³⁰⁶ They are also more likely to evolve under positive selection.

307 **References**

- 308 [1] Pritchard, J. K. and Di Rienzo, A. (2010) Adaptation – not by sweeps alone. *Nat Rev Genet*
309 11, 665–667.
- 310 [2] Fraser, H. B. (2013) Gene expression drives local adaptation in humans. *Genome Research*
311 23, 1089–1096.
- 312 [3] Höllinger, I. et al. (2019) Polygenic adaptation: From sweeps to subtle frequency shifts.
313 *PLOS Genetics* 15, e1008035.
- 314 [4] Hancock, A. M. et al. (2011) Adaptations to climate-mediated selective pressures in humans.
315 *PLoS genetics* 7, e1001375.
- 316 [5] Daub, J. T. et al. (2013) Evidence for polygenic adaptation to pathogens in the human
317 genome. *Mol Biol Evol* 30, 1544–58.
- 318 [6] He, F. et al. (2016) The Footprint of Polygenic Adaptation on Stress-Responsive *Cis* -
319 Regulatory Divergence in the *Arabidopsis Genus*. *Molecular Biology and Evolution* 33, 2088–
320 2101.
- 321 [7] Zan, Y. and Carlborg, O. (2019) A Polygenic Genetic Architecture of Flowering Time in the
322 Worldwide *Arabidopsis thaliana* Population. *Molecular Biology and Evolution* 36, 141–154.
- 323 [8] Hämälä, T. et al. (2019) Gene Expression Modularity Reveals Footprints of Polygenic Adap-
324 tation in *Theobroma cacao*. *Molecular Biology and Evolution* 37, 110–123.
- 325 [9] Fraser, H. B. et al. (2011) Systematic Detection of Polygenic cis-Regulatory Evolution. *PLOS*
326 *Genetics* 7, e1002023.
- 327 [10] Barreiro, L. B. and Quintana-Murci, L. (2010) From evolutionary genetics to human im-
328 munology: how selection shapes host defence genes. *Nat Rev Genet* 11, 17–30.
- 329 [11] Chouchani, E. T. and Kajimura, S. (2019) Metabolic adaptation and maladaptation in
330 adipose tissue. *Nature Metabolism* 1, 189–200.
- 331 [12] Edwards, S. L. et al. (2013) Beyond GWASs: illuminating the dark road from association to
332 function. *Am J Hum Genet* 93.

- 333 [13] Rodgers-Melnick, E. et al. (2016) Open chromatin reveals the functional maize genome.
334 Proceedings of the National Academy of Sciences 113, E3177–E3184.
- 335 [14] Enard, D. et al. (2014) Genome-wide signals of positive selection in human evolution. *Genome*
336 *Research* 24, 885–895.
- 337 [15] Latta, R. G. (1998) Differentiation of allelic frequencies at quantitative trait loci affecting
338 locally adaptive traits. *Am Nat* 151, 283–292.
- 339 [16] Le Corre, V. and Kremer, A. (2003) Genetic variability at neutral markers, quantitative
340 trait loci and trait in a subdivided population under selection. *Genetics* 164, 1205–1219.
- 341 [17] Kremer, A. and Le Corre, V. (2012) Decoupling of differentiation between traits and their
342 underlying genes in response to divergent selection. *Heredity* 108, 375–85.
- 343 [18] Chevin, L.-M. and Hospital, F. (2008) Selective sweep at a quantitative trait locus in the
344 presence of background genetic variation. *Genetics* 180, 1645–1660.
- 345 [19] Stephan, W. (2016) Signatures of positive selection: from selective sweeps at individual loci
346 to subtle allele frequency changes in polygenic adaptation. *Molecular Ecology* 25, 79–88.
- 347 [20] Pritchard, J. K. et al. (2010) The genetics of human adaptation: hard sweeps, soft sweeps,
348 and polygenic adaptation. *Current biology: CB* 20, R208–215.
- 349 [21] Haasl, R. J. and Payseur, B. A. (2016) Fifteen years of genomewide scans for selection:
350 trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology* 25, 5–23.
- 351 [22] Berg, J. J. and Coop, G. (2014) A Population Genetic Signal of Polygenic Adaptation. *PLoS*
352 *Genetics* 10, e1004412.
- 353 [23] Gouy, A. et al. (2017) Detecting gene subnetworks under selection in biological pathways.
354 *Nucleic Acids Res* 45, e149.
- 355 [24] Turchin, M. C. et al. (2012) Evidence of widespread selection on standing variation in Europe
356 at height-associated SNPs. *Nat Genet* 44, 1015–1019.
- 357 [25] Racimo, F. et al. (2018) Detecting polygenic adaptation in admixture graphs. *Genetics* 208,
358 1565–1584.

- 359 [26] Robinson, M. R. et al. (2015) Population genetic differentiation of height and body mass
360 index across Europe. *Nature Genetics* 47, 1357–1362.
- 361 [27] Field, Y. et al. (2016) Detection of human adaptation during the past 2000 years. *Science*
362 354, 760.
- 363 [28] Whitlock, M. C. and Lotterhos, K. E. (2015) Reliable detection of loci responsible for local
364 adaptation: Inference of a null model through trimming the distribution of F_{st} . *Am Nat* 186
365 Suppl 1, S24–S36.
- 366 [29] Duforet-Frebourg, N. et al. (2015) Detecting Genomic Signatures of Natural Selection with
367 Principal Component Analysis: Application to the 1000 Genomes Data. *Molecular Biology*
368 *and Evolution* 4, 1082–93.
- 369 [30] de Villemereuil, P. et al. (2014) Genome scan methods against more complex models: when
370 and how much should we trust them? *Molecular Ecology* 23, 2006–2019.
- 371 [31] Vilas, A. et al. (2012) A simulation study on the performance of differentiation-based methods
372 to detect selected loci using linked neutral markers. *J Evol Biol* 25, 1364–76.
- 373 [32] Novembre, J. and Barton, N. H. (2018) Tread Lightly Interpreting Polygenic Tests of
374 Selection. *Genetics* 208, 1351–1355.
- 375 [33] Rosenberg, N. A. et al. (2019) Interpreting polygenic scores, polygenic adaptation, and human
376 phenotypic differences. *Evolution, Medicine, and Public Health* 2019, 26–34.
- 377 [34] Berg, J. J. et al. (2019) Reduced signal for polygenic adaptation of height in UK Biobank.
378 *eLife* 8, e39725.
- 379 [35] Sohail, M. et al. (2019) Polygenic adaptation on height is overestimated due to uncorrected
380 stratification in genome-wide association studies. *eLife* 8, e39702.
- 381 [36] Uricchio, L. H. (2020) Evolutionary perspectives on polygenic selection, missing heritability,
382 and GWAS. *Human Genetics* 139, 5–21.
- 383 [37] Csillery, K. et al. (2018) Detecting the genomic signal of polygenic adaptation and the role
384 of epistasis in evolution. *Molecular Ecology* 27, 606–612.

- 385 [38] Payne, J. L. and Wagner, A. (2019) The causes of evolvability and their evolution. *Nature*
386 *Reviews Genetics* 20, 24–38.
- 387 [39] Manolio, T. A. et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461,
388 747–753.
- 389 [40] Boyle, E. A. et al. (2017) An Expanded View of Complex Traits: From Polygenic to Omni-
390 genic. *Cell* 169, 1177–1186.
- 391 [41] Kudaravalli, S. et al. (2009) Gene expression levels are a target of recent natural selection in
392 the human genome. *Molecular Biology and Evolution* 26, 649–658.
- 393 [42] Torgerson, D. G. et al. (2009) Evolutionary processes acting on candidate cis-regulatory
394 regions in humans inferred from patterns of polymorphism and divergence. *PLoS genetics* 5,
395 e1000592.
- 396 [43] Rougeux, C. et al. (2019) Polygenic selection drives the evolution of convergent transcriptomic
397 landscapes across continents within a Nearctic sister species complex. *Molecular Ecology* 28,
398 4388–4403.
- 399 [44] Kirschner, M. W. (2005) The Meaning of Systems Biology. *Cell* 121, 503–504.
- 400 [45] Sonawane, A. R. et al. (2019) Network Medicine in the Age of Biomedical Big Data. *Frontiers*
401 *in Genetics* 10.
- 402 [46] Wagner, A. (1994) Evolution of gene networks by gene duplications: a mathematical model
403 and its implications on genome organization. *Proceedings of the National Academy of Sciences*
404 91, 4387–4391.
- 405 [47] Wagner, A. (1996) Does evolutionary plasticity evolve? *Evolution* 50, 1008–1023.
- 406 [48] Waddington, C. H. (1942) Canalization of development and the inheritance of acquired
407 characters. *Nature* 150, 563–565.
- 408 [49] Siegal, M. L. and Bergman, A. (2002) Waddington’s canalization revisited: Developmental
409 stability and evolution. *Proceedings of the National Academy of Sciences* 99, 10528–10532.
- 410 [50] Runneburger, E. and Le Rouzic, A. (2016) Why and how genetic canalization evolves in
411 gene regulatory networks. *Bmc Evolutionary Biology* 16.

- 412 [51] Bergman, A. and Siegal, M. L. (2003) Evolutionary capacitance as a general feature of
413 complex gene networks. *Nature* 424, 549–552.
- 414 [52] Rhone, B. et al. (2011) Impact of selection on genes involved in regulatory network: a
415 modelling study. *J Evol Biol* 24, 2087–98.
- 416 [53] Bustamante, C. D. et al. (2005) Natural selection on protein-coding genes in the human
417 genome. *Nature* 437, 1153–1157.
- 418 [54] Halfon, M. S. (2017) Perspectives on Gene Regulatory Network Evolution. *Trends in Genetics*
419 33, 436–447.
- 420 [55] Das Gupta, M. and Tsiantis, M. (2018) Gene networks and the evolution of plant morphology.
421 *Current Opinion in Plant Biology* 45, 82–87.
- 422 [56] ten Tusscher, K. H. and Hogeweg, P. (2009) The role of genome and gene regulatory network
423 canalization in the evolution of multi-trait polymorphisms and sympatric speciation. *BMC*
424 *Evolutionary Biology* 9, 159.
- 425 [57] Doebley, J. F. et al. (2006) The Molecular Genetics of Crop Domestication. *Cell* 127, 1309–
426 1321.
- 427 [58] Swanson-Wagner, R. et al. (2012) Reshaping of the maize transcriptome by domestication.
428 *Proceedings of the National Academy of Sciences* 109, 11878–11883.
- 429 [59] Bullard, J. H. et al. (2010) Polygenic and directional regulatory evolution across pathways in
430 *Saccharomyces*. *Proceedings of the National Academy of Sciences* 107, 5058–5063.
- 431 [60] MacNeil, L. T. and Walhout, A. J. M. (2011) Gene regulatory networks and the role of
432 robustness and stochasticity in the control of gene expression. *Genome Research* 21, 645–657.
- 433 [61] Barghi, N. et al. (2019) Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLOS*
434 *Biology* 17, e3000128.
- 435 [62] Albert, R. (2005) Scale-free networks in cell biology. *Journal of Cell Science* 118, 4947–4957.
- 436 [63] Platig, J. et al. (2016) Bipartite Community Structure of eQTLs. *PLOS Computational*
437 *Biology* 12, e1005033.

- 438 [64] Jordan, I. K. et al. (2004) Conservation and Coevolution in the Scale-Free Human Gene
439 Coexpression Network. *Molecular Biology and Evolution* 21, 2058–2070.
- 440 [65] Josephs, E. B. et al. (2017) The Relationship between Selection, Network Connectivity,
441 and Regulatory Variation within a Population of *Capsella grandiflora*. *Genome Biology and*
442 *Evolution* 9, 1099–1109.
- 443 [66] Luisi, P. et al. (2015) Recent Positive Selection Has Acted On Genes Encoding Proteins With
444 More Interactions Within the Whole Human Interactome. *Genome Biology and Evolution* 4,
445 1141–1154.
- 446 [67] Fagny, M. et al. (2017) Exploring regulation in tissues with eQTL networks. *PNAS* 114,
447 E7841–E7850.
- 448 [68] Baker, R. L. et al. (2019) Integrating transcriptomic network reconstruction and eQTL
449 analyses reveals mechanistic connections between genomic architecture and *Brassica rapa*
450 development. *PLOS Genetics* 15, e1008367.
- 451 [69] Barghi, N. et al. (2020) Polygenic adaptation: a unifying framework to understand positive
452 selection. *Nature Reviews Genetics* 21, 769–781.
- 453 [70] Fisher, R. A. (1919) XV.—The Correlation between Relatives on the Supposition of
454 Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52, 399–433.
- 455 [71] Mackay, T. F. C. et al. (2009) The genetics of quantitative traits: challenges and prospects.
456 *Nature Reviews Genetics* 10, 565–577.
- 457 [72] Holland, J. B. (2007) Genetic architecture of complex traits in plants. *Genome Studies and*
458 *Molecular Genetics* 10, 156–161.
- 459 [73] Timpson, N. J. et al. (2018) Genetic architecture: the shape of the genetic contribution to
460 human traits and disease. *Nature Reviews Genetics* 19, 110–124.
- 461 [74] Barton, N. H. et al. (2017) The infinitesimal model: Definition, derivation, and implications.
462 *Theoretical Population Biology* 118, 50 – 73.
- 463 [75] Peiffer, J. A. et al. (2014) The Genetic Architecture Of Maize Height. *Genetics* 196, 1337–
464 1356.

- 465 [76] Pais, T. M. et al. (2013) Comparative Polygenic Analysis of Maximal Ethanol Accumulation
466 Capacity and Tolerance to High Ethanol Levels of Cell Proliferation in Yeast. PLOS Genetics
467 9, e1003548.
- 468 [77] Salome, P. A. et al. (2011) Genetic Architecture of Flowering-Time Variation in *Arabidopsis*
469 *thaliana*. Genetics 188, 421.
- 470 [78] Fuchsberger, C. et al. (2016) The genetic architecture of type 2 diabetes. Nature advance
471 online publication.
- 472 [79] Voight, B. F. et al. (2006) A map of recent positive selection in the human genome. PLoS
473 Biol 4, e72.
- 474 [80] Ferrer-Admetlla, A. et al. (2014) On detecting incomplete soft or hard selective sweeps using
475 haplotype structure. Mol Biol Evol 31, 1275–91.

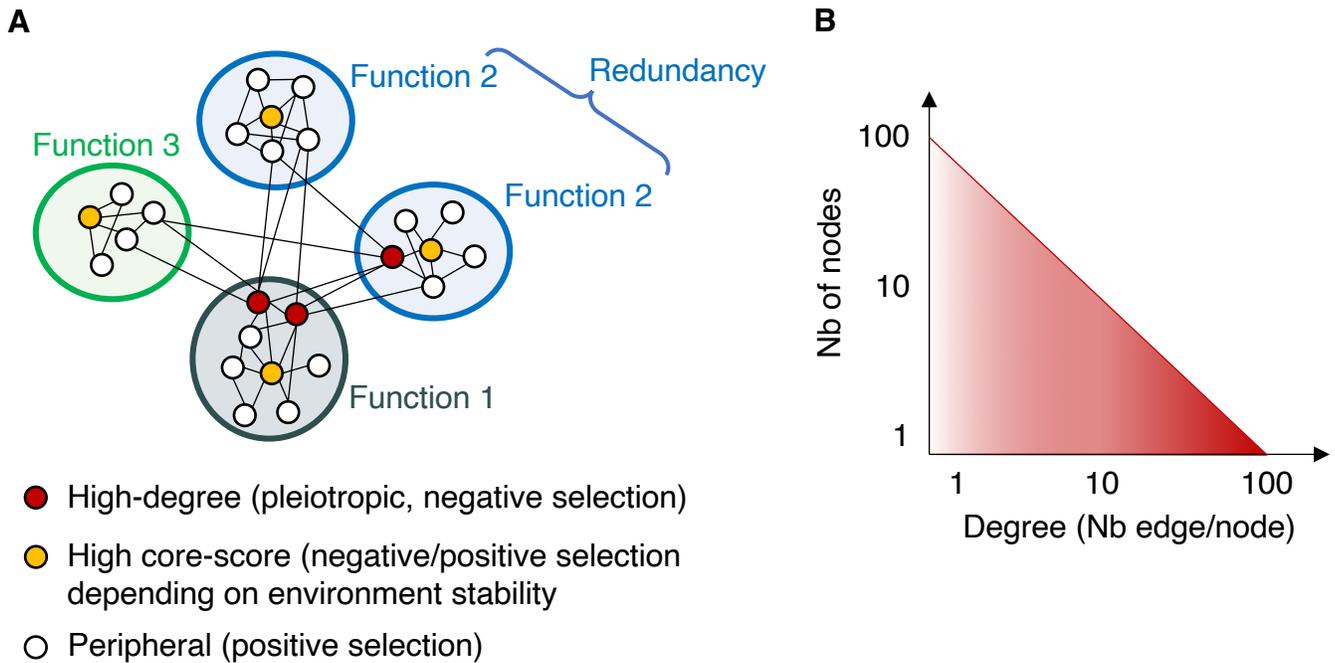


Figure 1: Impact of gene regulatory network structure on selection action. **(A)** Modular structure of a gene regulatory network. Genes are represented as circles and edges represent regulatory relationship between gene products. Genes are colored by topological characteristics: high degree genes in red, high core-score genes in yellow and others in white. The network is structured into modules, i.e. groups of genes that regulate each others. These modules often correspond to biological functions. Topological characteristics of gene regulatory networks such as redundancy (several modules involved in the regulation of the same biological functions) and pleiotropy (genes involved in the regulation of several biological functions) are highlighted. **(B)** Expected power law distribution of the node degree in q gene regulatory network.

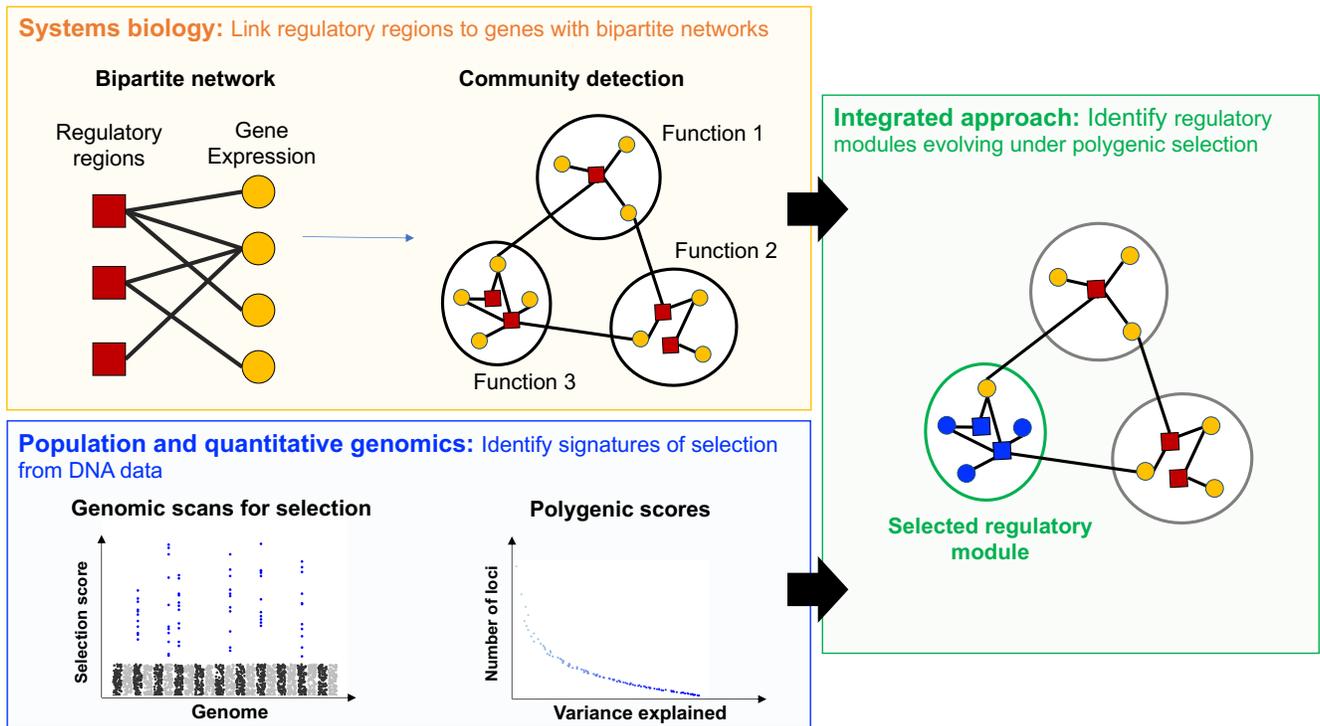


Figure 2: Proposition for a new integrated framework to detect signatures of polygenic selection. Genes and regulatory regions in blue present signatures of selection.