

Ecological genomics

Isheng Jason Tsai

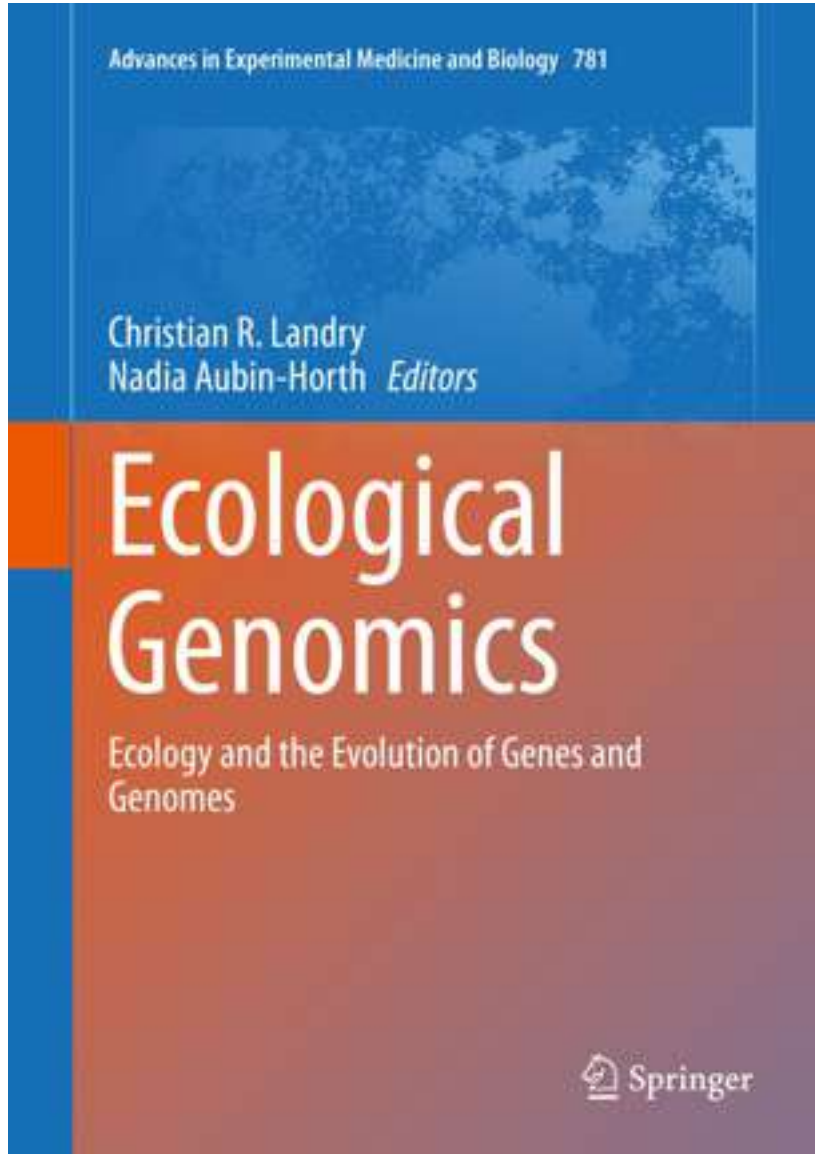
Advanced Microbiology
v2020



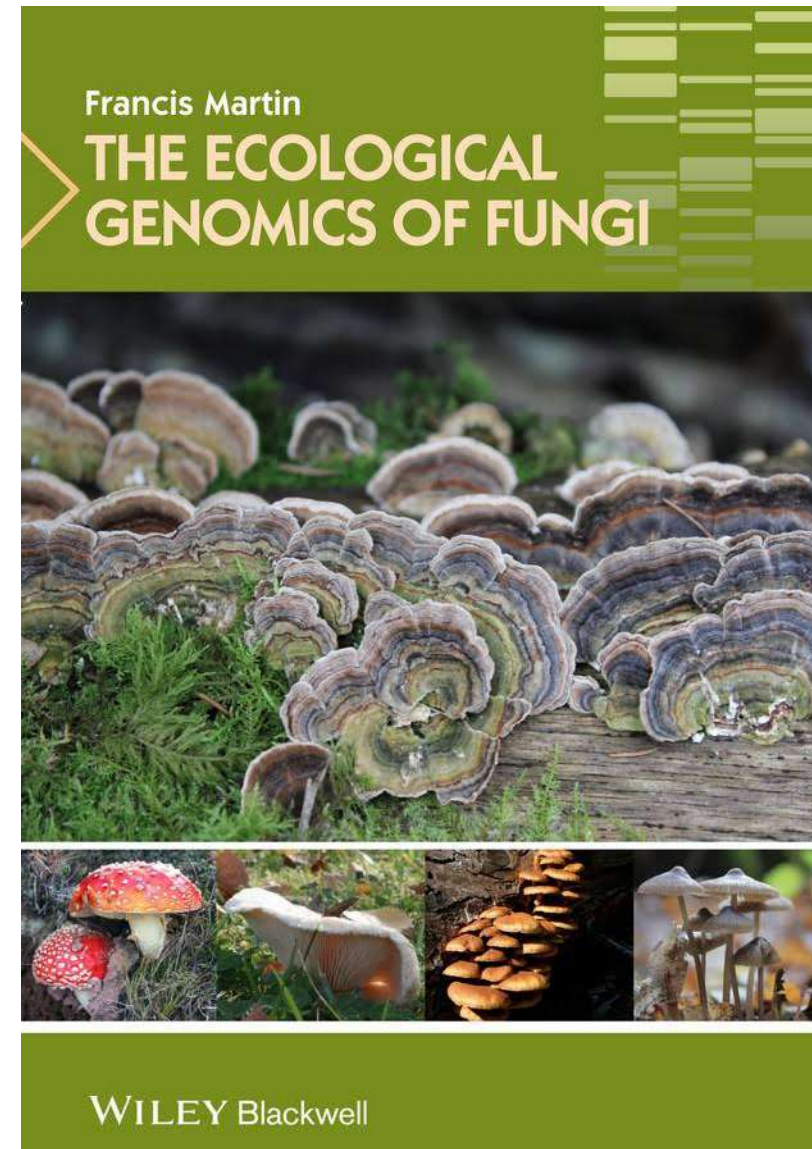
Lecture objective

- Ecological genomics
- Various approaches and disciplines
- Case studies

Recommended textbooks



(2014)



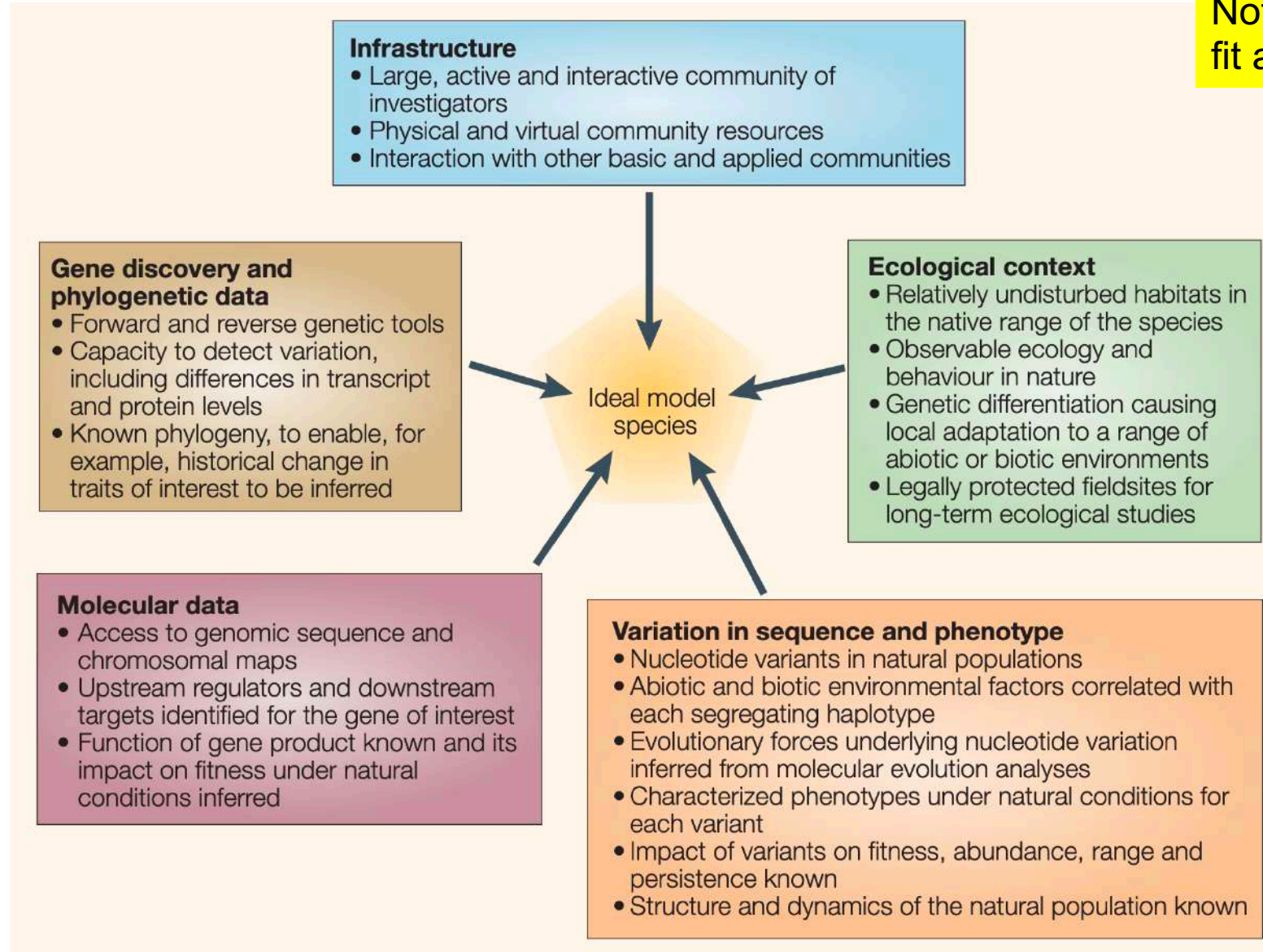
Ecological genomics (EG) - definition

“A unique combination of disciplines is emerging — evolutionary and ecological functional genomics — which **focuses on the genes that affect ecological success and evolutionary fitness in natural environments and populations**

- “the focus is on organisms that inhabit natural environments and the goal of researchers is to explain variation in DARWINIAN FITNESS in populations, and variation in size, range, longevity and diversity among populations, species and higher taxa
- Identify gene or genes of interest
- This is challenging and requires multiple disciplines (ecology, evolution, functional biology and genomics)

An ideal model organism for EG

Not many organisms fit all these criteria



Ecological genomics

Genotypes

- Genotype frequencies
- Genomic variations
- Population genomics
- Comparative genomics

Phenotypes

- Phenotype frequencies
- Phenotype plasticity
- Development

Traditional model
organisms

Ecology

- Abiotic
- Biotic
- Short term / long term

Ecologists

Ecological genomics

Genotypes

- Genotype frequencies
- Genomic variations
- Population genomics
- Comparative genomics

Advances in genomics really kick off this field; rather than choose a model species we can ask virtually any questions across all organisms

Phenotypes

- Phenotype frequencies
- Phenotype plasticity
- Development

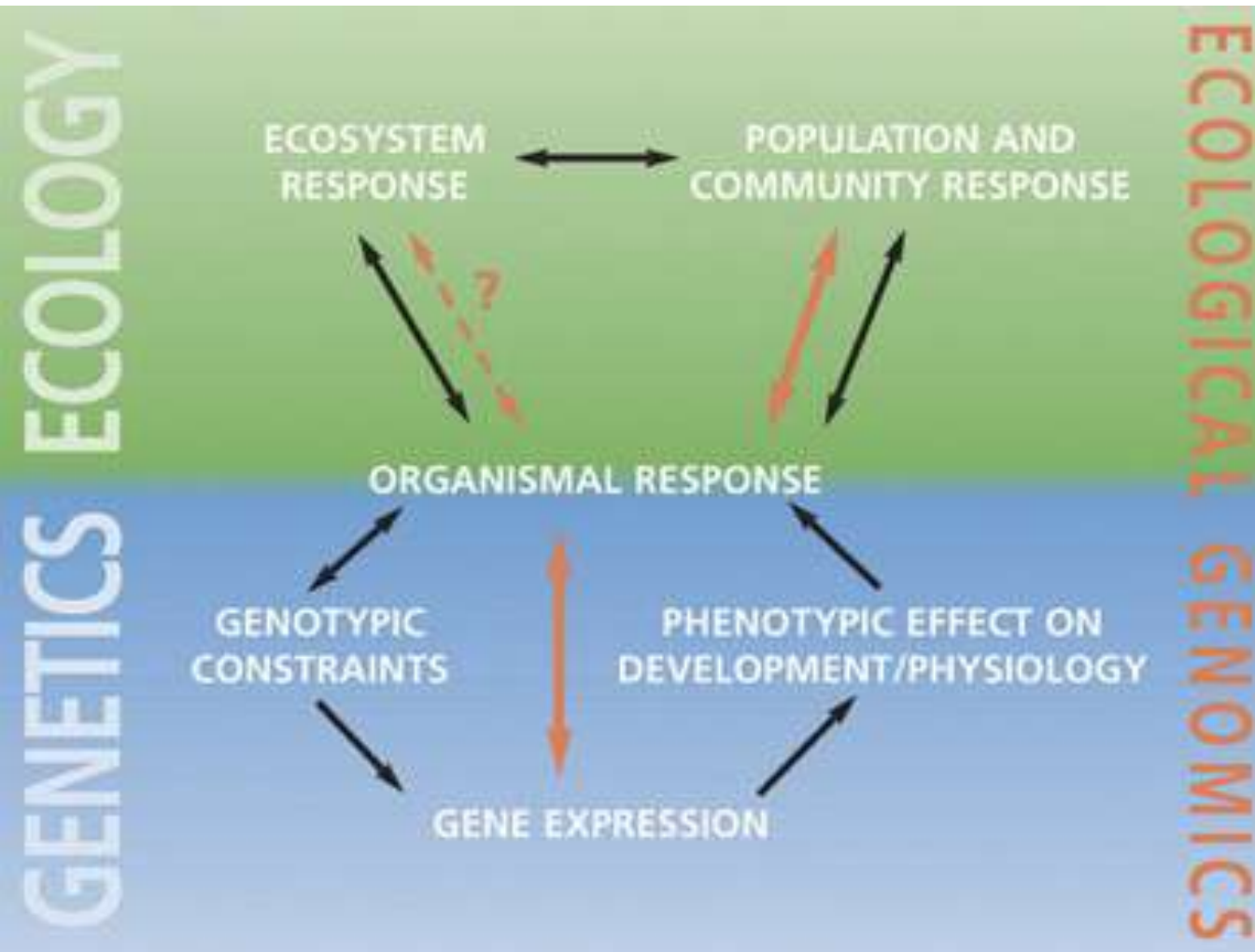
Traditional model organisms

Ecology

- Abiotic
- Biotic
- Short term / long term

Ecologists

Conceptual framework for eco genomics



Ecological interactions between the organism, the population and community levels and the ecosystem

Interactions between the levels, with organismal responses affecting and being affected by its genotype, which in turn affects what genes are expressed and at what levels, which in turn has effects on the phenotype of the organism, ultimately leading to its overall response.

Ecological genomic studies seek to integrate these disciplines (orange arrows) through the use of functional genomics approaches.

Some questions in EG

- What are the genes that underlie traits that mediate ecological responses in nature?
- How does environmental variation influence mechanisms underlying organismal response?
- Does adaptation and plasticity involve many genes of small effect or a few genes of large effect?
- How does ecological context influence the evolution of genome structure and function?
- How does genotypic variation within and among species influence evolutionary responses, and (or) population, community, or ecosystem dynamics?
- How do microbial communities shift when environments change and how do these shifts influence ecological processes?

Ecological genomics – a fungi perspective

Fungi – a definition

fungus (n.)

1520s, "a mushroom," from Latin *fungus* "a mushroom, fungus;" used in English at first as a learned alternative to *mushroom* (*funge* was used in this sense late 14c.). The Latin word is believed to be cognate with (or derived from) Greek *sphongos*, the Attic form of *spongos* "sponge" (see **sponge** (n.)). "Probably a loanword from a non-IE language, borrowed independently into Greek, Latin and Armenian in a form **sphong-*" [de Vaan]

"The concept of a "fungus" has developed over many years, and the historic definition of fungi as nonphotosynthetic plants has been shown to be both too simplistic and phylogenetically inaccurate"

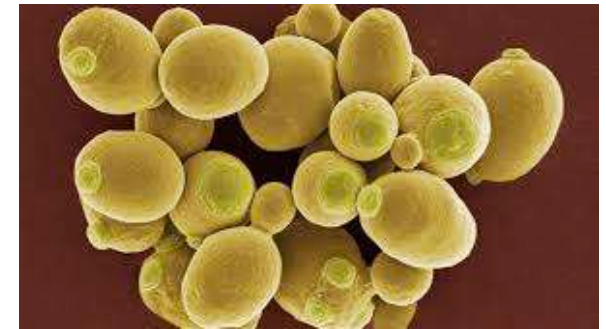
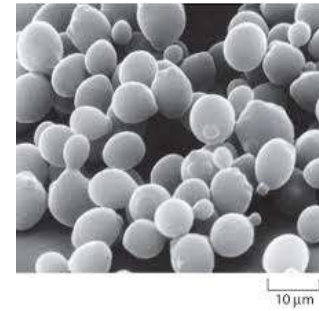
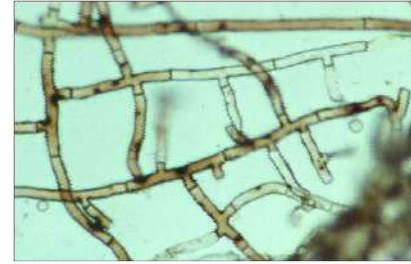
Bridge *et al* (2005) *Advances in Botanical Research*

<https://www.etymonline.com/word/fungus>

General Characteristics of Fungi

- Cell wall present, composed of cellulose and/or chitin.
- Food storage - generally in the form of lipids and glycogen.
- Eukaryotes - true nucleus and other organelles present.
- Most fungi require water and oxygen (no obligate anaerobes).
- Fungi grow in almost every habitat imaginable, as long as there is some type of organic matter present and the environment is not too extreme.
- Diverse group, number of described species is **~70,000** (estimated 5.1 million species total).

Fungi are eukaryotes



Fungi are heterotrophs ('other food')

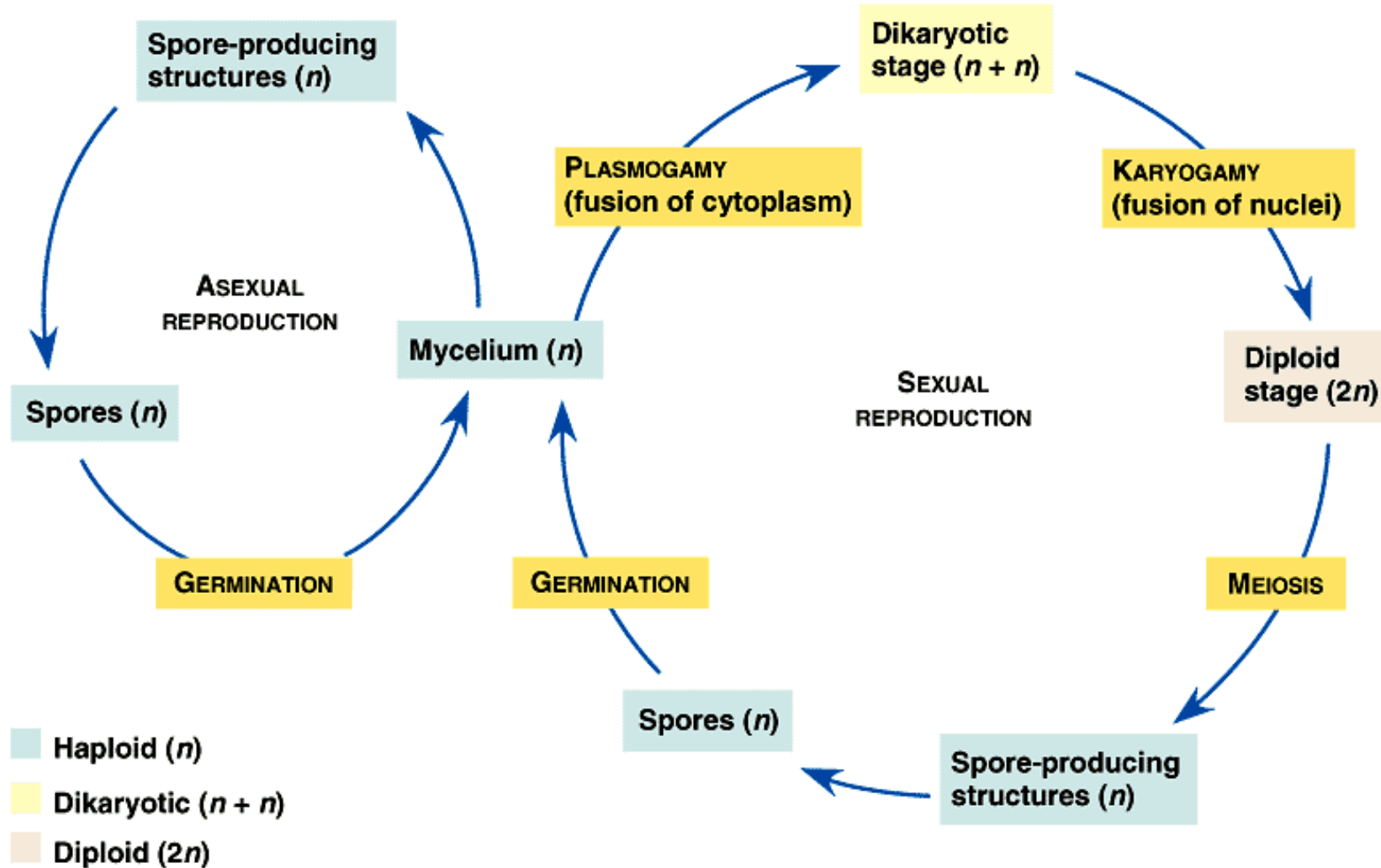
- i.e., acquire nutrients by absorption



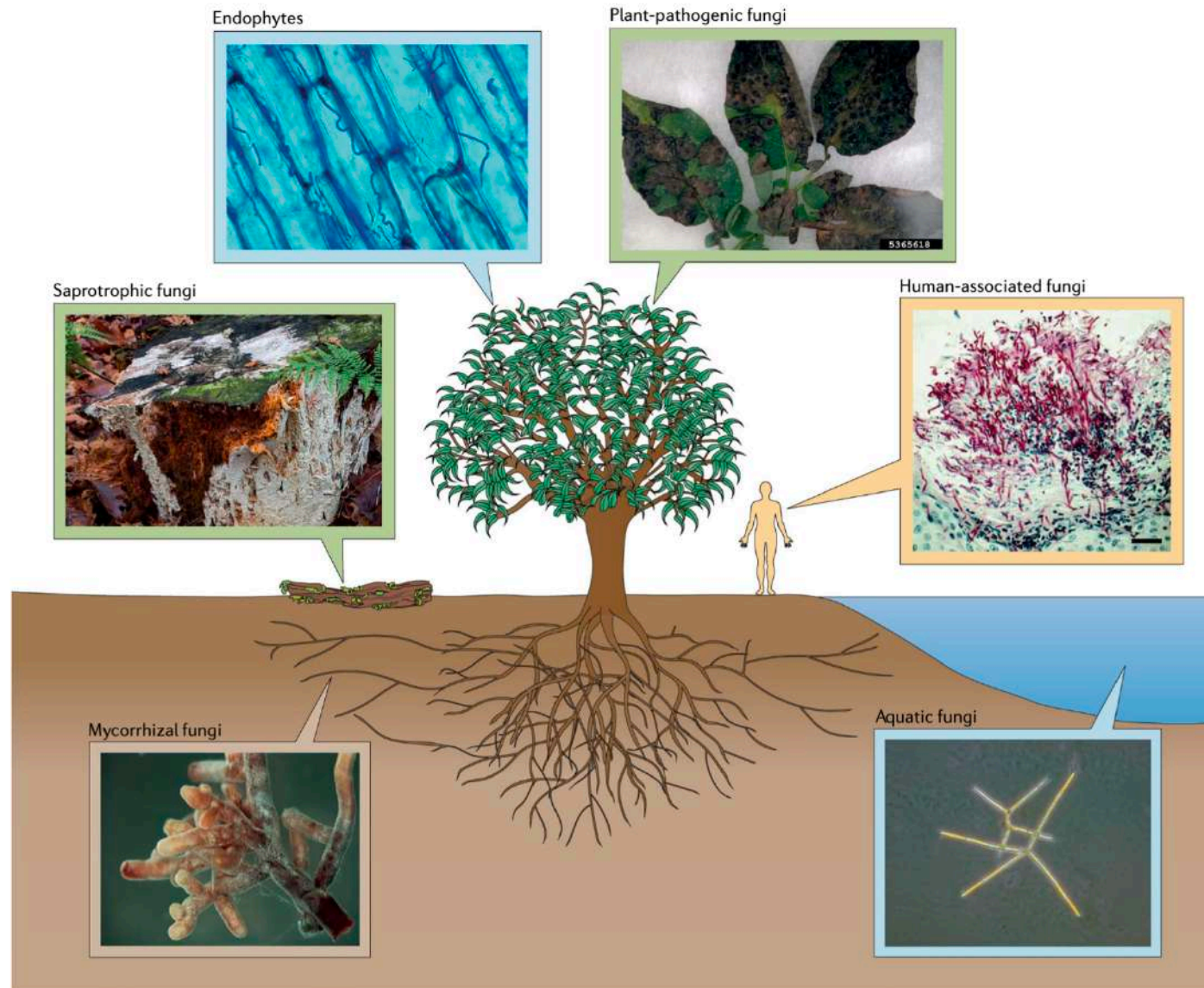
Credit: UWA 2005

- Three main types:
 - **Saprophytes** or saprobes - absorb nutrients from dead organic material
 - **Parasitic** fungi - absorb nutrients from cells of living hosts; some are pathogenic
 - **Mutualistic** fungi - absorb nutrients from a host, but reciprocate to benefit the host

Generalized Life Cycle of a Fungus



Fungal diversity in different environments



Ecological impacts of fungi

- **Beneficial Effects of Fungi**

- Many organisms depend on/utilise fungi
- Decomposition - nutrient and carbon recycling.
- Biosynthetic factories. Can be used to produce drugs, antibiotics, alcohol, acids, food (e.g., fermented products, mushrooms).
- Traffic network for microorganisms and host
- Model organisms for biochemical and genetic studies.

- **(Host) Harmful Effects of Fungi**

- Destruction of food, lumber, paper, and cloth.
- Animal and human diseases, including allergies.
- Toxins produced by poisonous mushrooms and within food (e.g., grain, cheese, etc.).
- Plant diseases.

How fungi are engaged in ecosystem processes

1. Primary production
 - a. Making nutrients available for plant growth
 - b. Enhancing nutrient uptake in the rhizosphere
2. Secondary production
 - a. Providing food for both vertebrate and invertebrate animals
 - b. Other fungal/faunal interactions
3. Population and community regulation by plant and animal pathogens
4. Interactions between fungi and human activities
 - a. Pollution
 - b. The built environment

Ecosystem Service		Fungal Functional Group
Soil formation	Rock dissolution	Lichens Saprotrophs Mycorrhizae
	Particle binding	Saprotrophs Mycorrhizae
Providing fertility for primary production	Decomposition of organic residues and nutrient mineralization	Saprotrophs (Ericoid and ectomycorrhizae)
	Soil stability (aggregates)	Saprotrophs Arbuscular mycorrhizae
Primary production	Direct production	Lichens
	Nutrient accessibility	Mycorrhizae Endophytes
Plant yield	Plant yield	Mycorrhizae Pathogens
	Defense against pathogens	Mycorrhizae Endophytes Saprotrophs
Plant community structure	Defense against herbivory	Endophytes
	Plant–plant interactions	Mycorrhizae Pathogens
Secondary production	As a food source	Saprotrophs Mycorrhizae
	Population/biomass regulation	Pathogens
Modification of pollutants		Saprotrophs Mycorrhizae
Carbon sequestration and storage		Mycorrhizae (Saprotrophs)
Decay of human structures and artifacts		Saprotrophs

Note: Fungal groups in parentheses are regarded as being of lesser importance in that function.

PLANT-FUNGAL MUTUALISMS

1. Endophytes can be beneficial to plant growth, especially in harsh environments

2. Mycorrhizas. key benefits to terrestrial ecosystems, including: (i) enhanced nutrient uptake, (ii) soil structure; and (iii) carbon sequestration,

PARASITOIDS

Parasitoids are attracted by the elevated production of volatile organic compounds in mycorrhizal plants, increasing the rate of parasitism of herbivores.

HERBIVORES

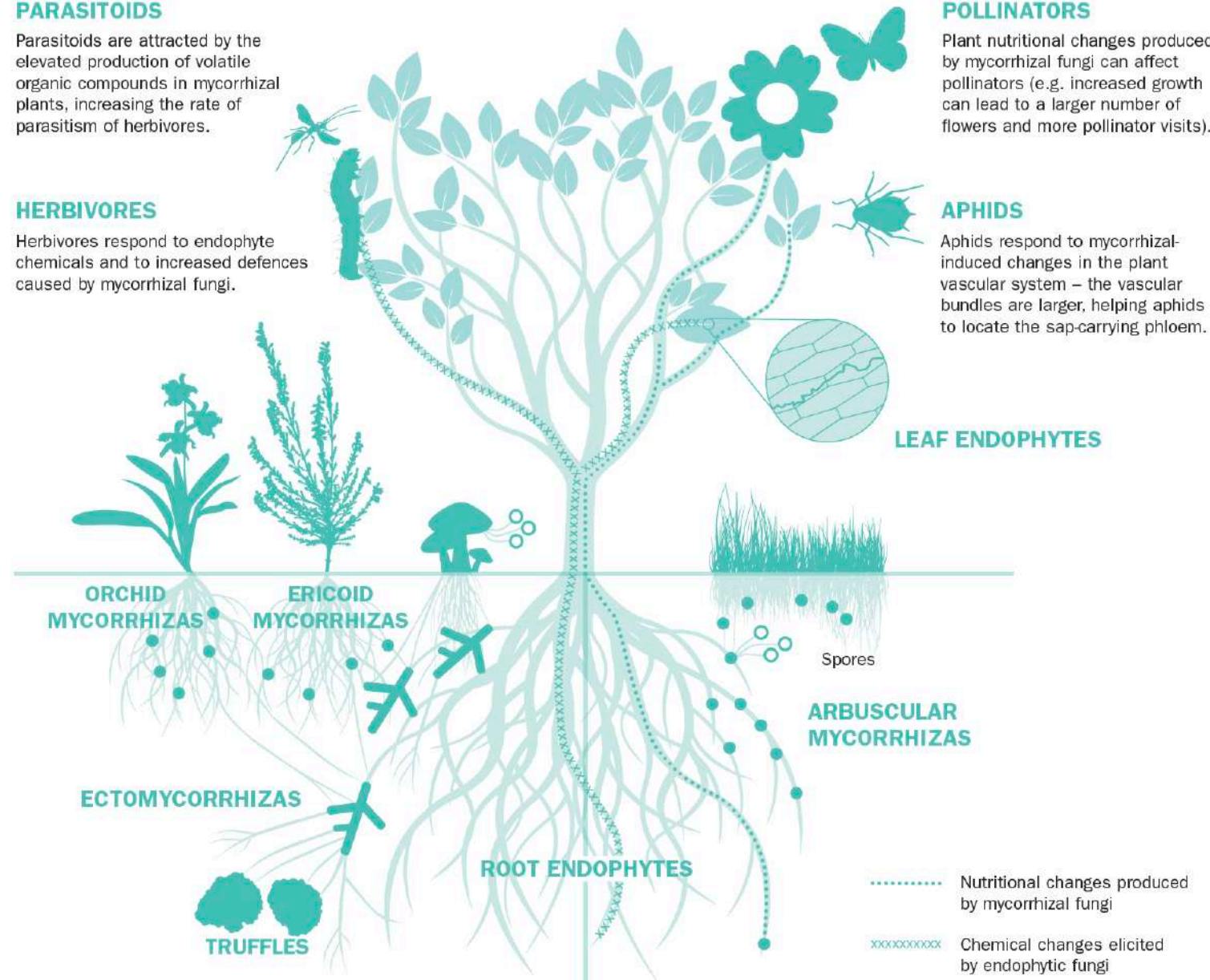
Herbivores respond to endophyte chemicals and to increased defences caused by mycorrhizal fungi.

POLLINATORS

Plant nutritional changes produced by mycorrhizal fungi can affect pollinators (e.g. increased growth can lead to a larger number of flowers and more pollinator visits).

APHIDS

Aphids respond to mycorrhizal-induced changes in the plant vascular system – the vascular bundles are larger, helping aphids to locate the sap-carrying phloem.

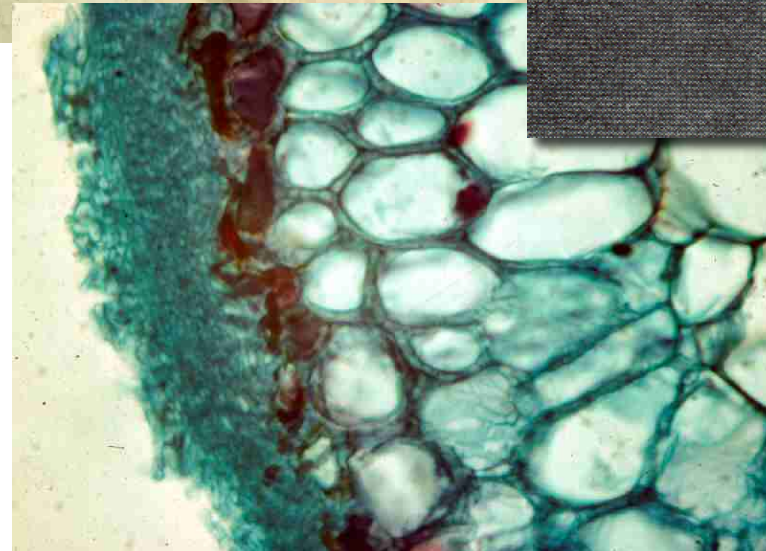
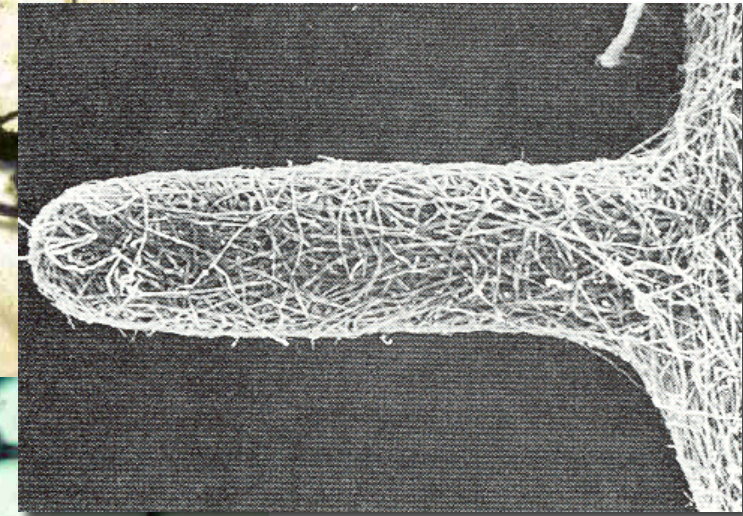


Mycorrhizae

- “Fungus roots”
- Mutualism between:
 - Fungus (nutrient & water uptake for plant)
 - Plant (carbohydrate for fungus)
- Several kinds
 - Zygomycota – hyphae invade root cells
 - Ascomycota & Basidiomycota – hyphae invade root but don’t penetrate cells
- **Extremely** important ecological role of fungi!



Russula
mushroom
mycorrhizas on
Western
Hemlock root



Main types of Mycorrhizae and their plant partners

MYCORRHIZAL TYPE	PLANT PARTNER	FUNGAL PARTNER	MAIN ECOSYSTEMS
Ectomycorrhizal fungi	2.2% of plant species, especially woody species; Pinaceae (e.g. pine, spruce, fir, larch) and angiosperms (e.g. beech, oak, chestnut, hazelnut, rockrose)	20,000 Basidiomycota and Ascomycota species	Temperate, boreal, mediterranean, and some tropical forests
Arbuscular mycorrhizal fungi	78% of plant species; Herbs, shrubs, trees, liverworts, hornworts, lycopods and ferns	300–1,600 Mucoromycota (Glomeromycotina) species	Tropical and temperate forests, grasslands, savannas, shrublands, deserts, agricultural crops
Ericoid mycorrhizal fungi	1.5% of plant species; Ericaceae (e.g. heather, rhododendron, blueberry) and liverworts	>150 Ascomycota and some Basidiomycota species	Heathlands, tundra, boreal and temperate forests
Orchid mycorrhizal fungi	10% of plant species; Orchidaceae (orchids)	25,000 Basidiomycota and some Ascomycota species	Tropical, temperate, mediterranean

REVIEW

doi:10.1038/nature10947

Emerging fungal threats to animal, plant and ecosystem health

Matthew C. Fisher¹, Daniel A. Henk¹, Cheryl J. Briggs², John S. Brownstein³, Lawrence C. Madoff⁴, Sarah L. McCraw⁵ & Sarah J. Gurr⁵

Plant pathogens

<i>Sporisorium scitamineum</i> (Sscl8)	Sugar cane	Sugar cane	Biotroph	Sexual	20	6.7	Repeat-rich gene clusters that encode effector candidates	48
<i>Ustilago maydis</i> (521)	Maize	Maize	Biotroph	Sexual	20	6.7	Repeat-rich gene clusters that encode effector candidates	48, 66
<i>Microbotryum lychnidis-dioicae</i> (p1A1 Lamole)	Red campion	Red campion	Biotroph	Sexual	33	14	Repeat-rich gene clusters that encode effector candidates	103, 104
<i>Melampsora larici-populina</i> (98AG31)	Poplar and larch	Poplar and larch	Biotroph	Sexual and asexual	101	45	ND [#]	73
<i>Puccinia graminis</i> f. sp. <i>tritici</i> (CDL75-36-700-3, race SCCL)	Wheat	Wheat and barley	Biotroph	Sexual and asexual	89	45	Highly polymorphic effector candidates	73
<i>Zymoseptoria tritici</i> (IPO323)	Wheat	Wheat	Hemi-biotroph	Sexual and asexual	40	<ul style="list-style-type: none"> • 18.6 (genome mean) • 16.6 (core) • 33.6 (accessory) 	<ul style="list-style-type: none"> • Orphan regions are enriched in <i>in planta</i>-expressed genes • Possible function of accessory chromosomes in virulence 	45, 57, 58
<i>Leptosphaeria biglobosa</i> 'canadensis' (J154)	Mustard	Crucifers	Necrotroph	Sexual and asexual	32	3.9	ND [#]	71
<i>Leptosphaeria maculans</i> 'brassicae' (v23.1.3)	Oilseed rape	Crucifers	Hemi-biotroph	Sexual and asexual	45	35.5 (99.8% of all repeats located in AT-isochores)	<ul style="list-style-type: none"> • Enrichment of effector candidates and chromatin-mediated effector candidate regulation in AT isochores • Conditionally dispensable chromosome contains avirulence-encoding gene 	29, 60

Plant pathogens

							effectors	
<i>Blumeria graminis</i> f. sp. <i>hordei</i> (DH14)	Barley	Various	Biotroph	Sexual and asexual	120	64	Repeat-rich accessory regions that encode infection-specific transcribed genes	142
<i>Magnaporthe oryzae</i> (70–15)	Rice	Various crops and wild grasses	Hemi-biotroph	Sexual and asexual	41	10	Highly polymorphic effector candidates and translocations of effector genes	22, 143
<i>Ophiostoma novo-ulmi</i> (H327)	Elm	Elm	Necrotroph	Sexual and asexual	32	3.4	ND [#]	144
<i>Verticillium dahliae</i> (VdLs17)	Lettuce	Various	Necrotroph	Asexual	37	12	Enrichment of <i>in planta</i> -expressed effector candidates in LS [#] regions	16, 44
<i>Fusarium solani</i> / <i>Nectria haematococca</i> MPVI (77-13-4)	Pea	Various	Hemi-biotroph	Sexual and asexual	54	• <5 (core) • >10–25 (supernumerary)	LS chromosomes confer host specificity and virulence	46
<i>Fusarium graminearum</i> (PH-1)	Wheat	Wheat and barley	Hemi-biotroph	Sexual and asexual	36	<3	Enrichment of <i>in planta</i> -expressed and species-specific genes in regions of high SNP [#] density	145
<i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> (4287)	Tomato	Various	Hemi-biotroph	Asexual	60	28 (~74% of TEs [#] located on LS [#] chromosomes)	LS chromosomes confer host specificity and virulence	43

LS, lineage-specific; ND, no data; TEs, transposable elements. *Genome size and repeat content refer to the respective reference isolate. Isolate-specific hosts refer to the host plant from which the reference isolate was collected. In some cases, other isolates of the same species infect other hosts. †Genome characteristics have been inferred from comparative genomics analyses.

Corn smut *Ustilago*





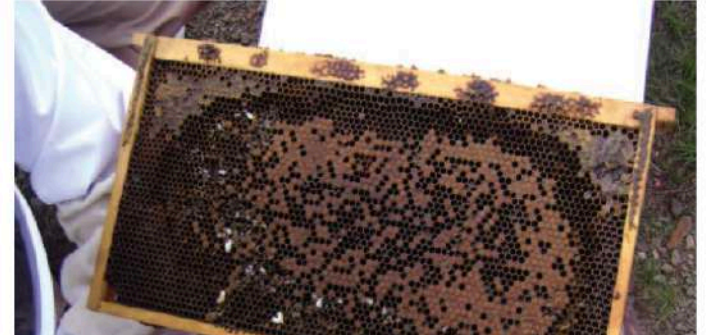
Photo by David Cohen/flickr/CC BY 2.0



Zizania latifolia infected by *Ustilago esculenta*



Animal pathogens

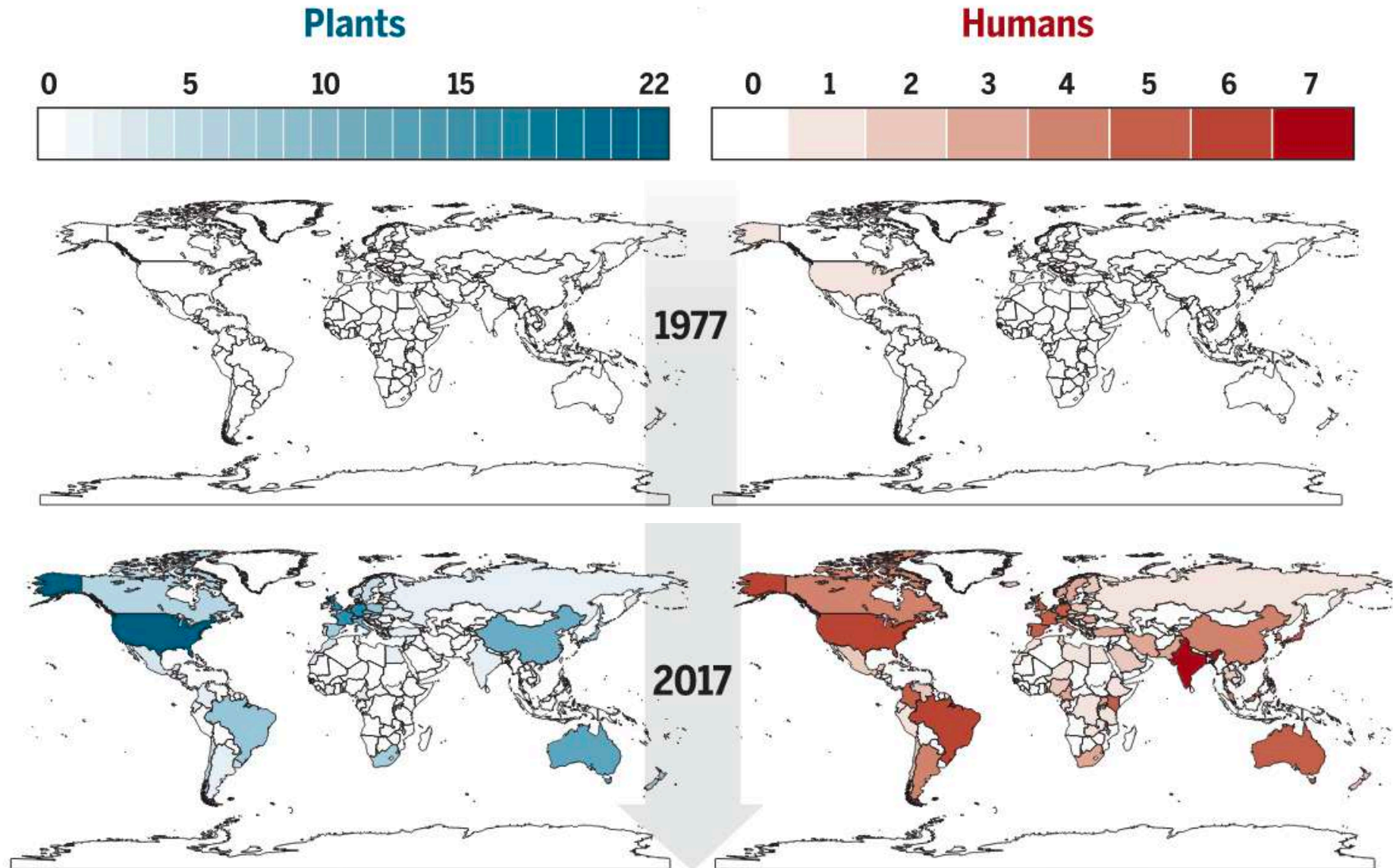
	Host	Pathogen (Phylum)	Disease dynamics leading to mass mortality in animal and plant hosts
	Amphibian species (for example, the common midwife toad, <i>Alytes obstetricans</i>)	<i>Batrachochytrium dendrobatidis</i> (Chytridiomycota)	Worldwide dispersal of a hypervirulent lineage by trade ⁶⁴ . Ultra-generalist pathogen manifesting spillover between tolerant/susceptible species. Extent of chytridiomycosis is dependent on biotic and abiotic context ^{15,82} .
	Coral species (for example, the sea fan, <i>Gorgonia ventalina</i>)	<i>Aspergillus sydowii</i> (Ascomycota)	Sea-fan aspergillosis caused by a common terrestrial soil fungus ^{21,86} . Epizootics are associated with warm-temperature anomalies. Coral immunosuppression is probably a factor causing decline.
	Bee species (for example, the hive of the domestic honeybee (<i>Apis mellifera</i>) suffering colony collapse disorder)	<i>Nosema</i> species (Microsporidia)	Microsporidian fungal infections are associated with colony collapse disorder and declining populations. Pathogen prevalence is probably a part of a multifactorial phenomenon that includes environmental stressors and polyparasitism ^{87,88} .

Worldwide emergence of resistance to antifungal drugs challenges human health and food security

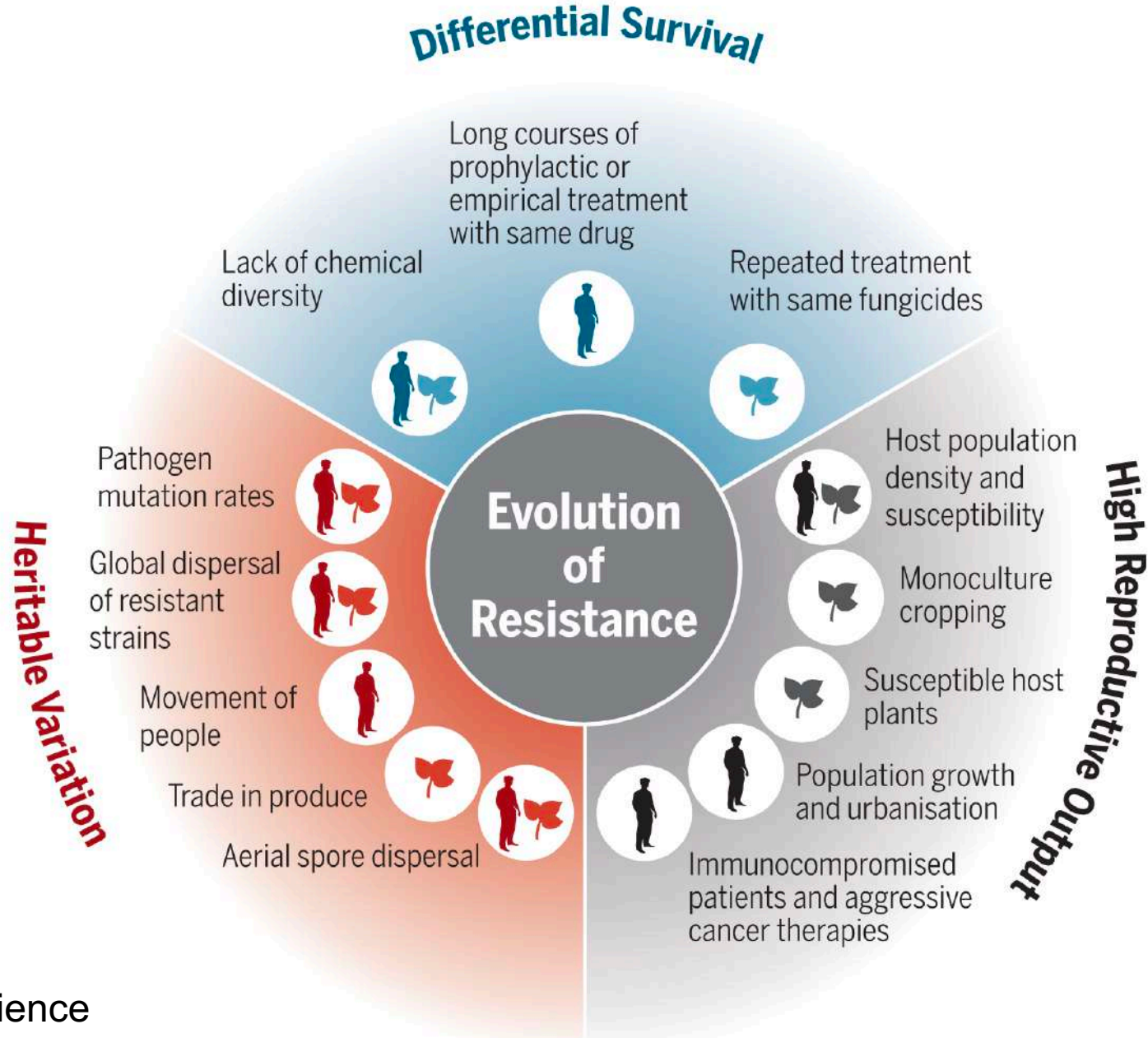
Matthew C. Fisher,^{1*} Nichola J. Hawkins,² Dominique Sanglard,³ Sarah J. Gurr^{4,5*}

The recent rate of emergence of pathogenic fungi that are resistant to the limited number of commonly used antifungal agents is unprecedented. The azoles, for example, are used not only for human and animal health care and crop protection but also in antifouling coatings and timber preservation. The ubiquity and multiple uses of azoles have hastened the independent evolution of resistance in many environments. One consequence is an increasing risk in human health care from naturally occurring opportunistic fungal pathogens that have acquired resistance to this broad class of chemicals. To avoid a global collapse in our ability to control fungal infections and to avoid critical failures in medicine and food security, we must improve our stewardship of extant chemicals, promote new antifungal discovery, and leverage emerging technologies for alternative solutions.

Fungal species with reported antifungal resistance, by country



Evolutionary drivers of antifungal resistance



So how do we study all of these traits at once?

Genome reference

- Genome assembly

Compare different species (inter-species comparisons)

- Phylogenomics
- Comparative genomics
 - Gene family evolution
 - Gain/loss
 - Expansion/Contraction

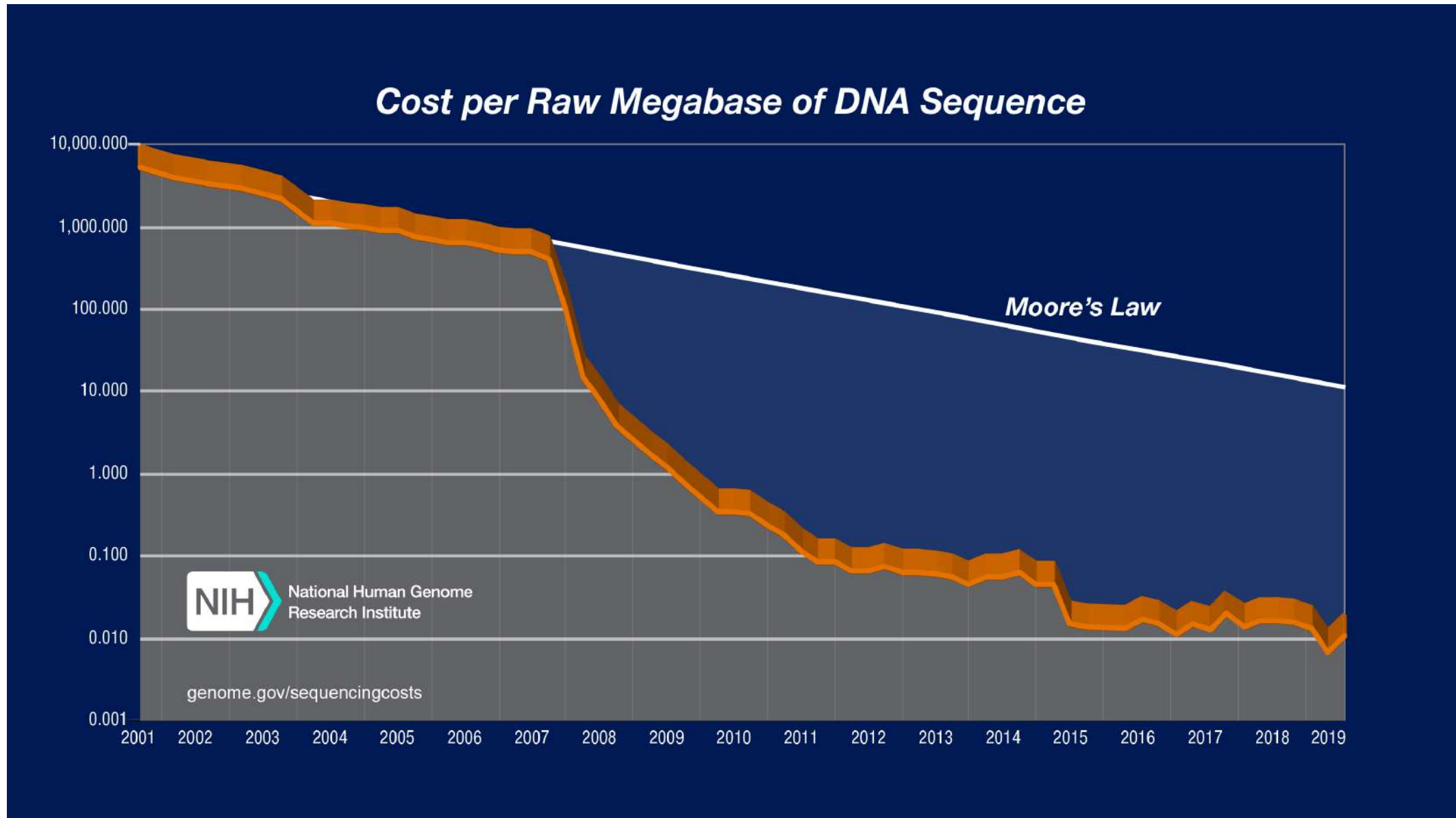
Compare different strains (intra-species comparisons)

- Population genomics
- QTL
- GWAS

Ecological modeling

First – you need to sequence genomes

Advances in sequencing



Fungal genomes



Goal: Sequencing 1000 fungal genomes from across the Fungal Tree of Life will provide references for research on plant-microbe interactions and environmental metagenomics.

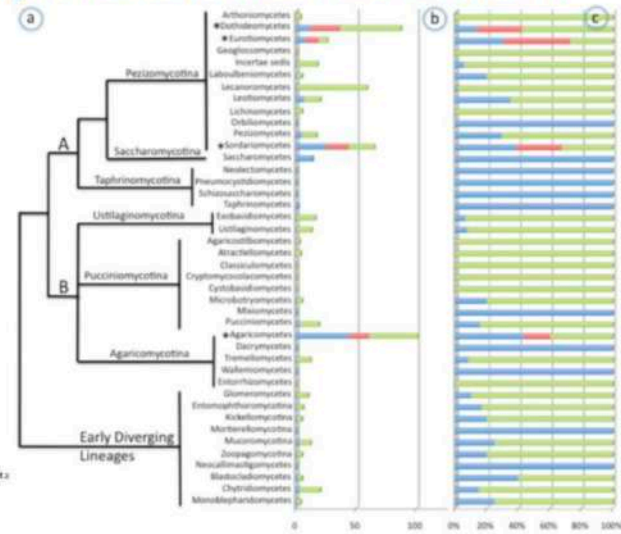
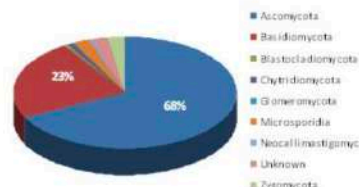
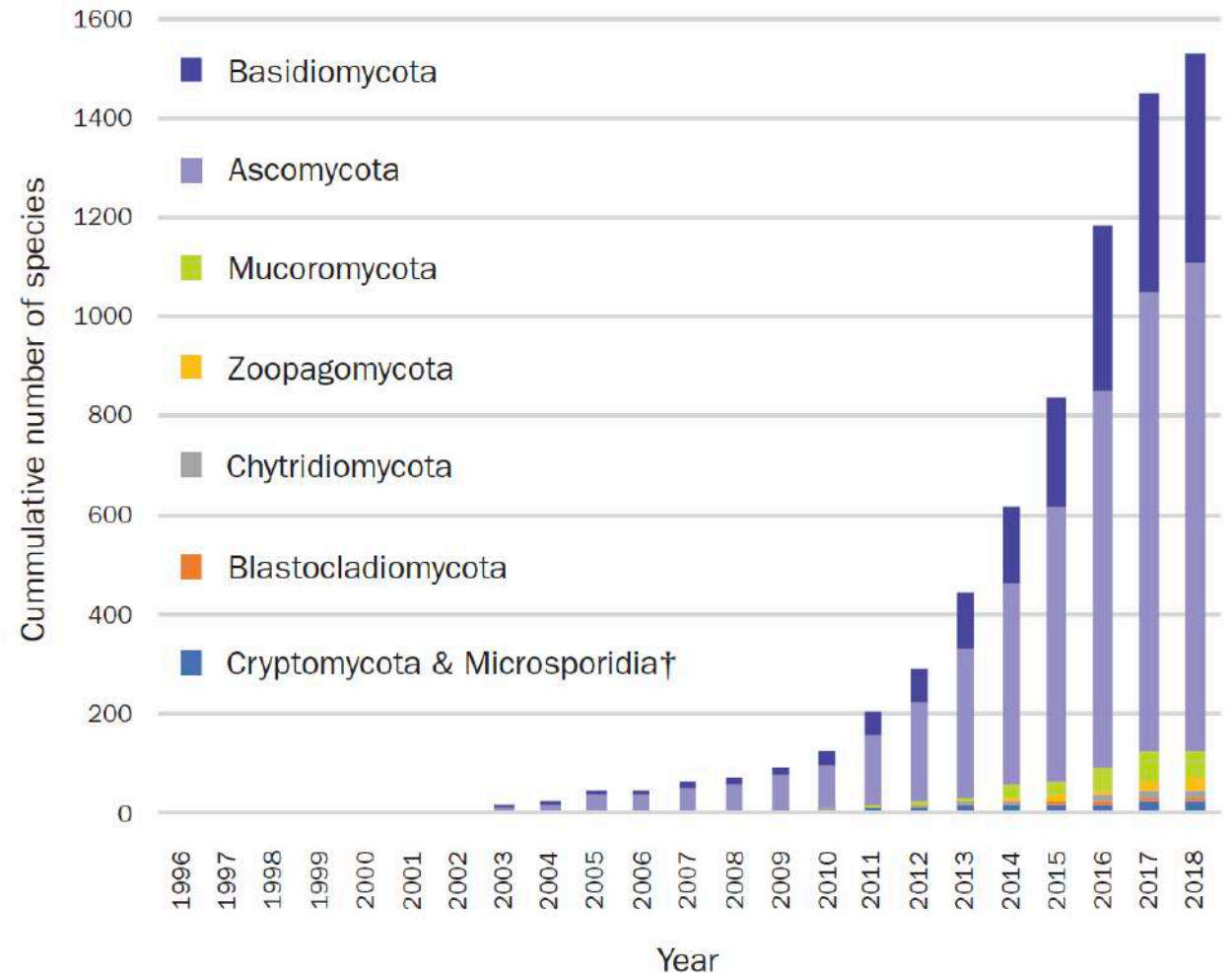


FIGURE 1: CUMULATIVE NUMBER OF FUNGAL SPECIES WITH WHOLE-GENOME SEQUENCES

[Data collated from online genome databases including NCBI Genome database (ncbi.nlm.nih.gov/genome), JGI Genome Portal: MycoCosm (genome.jgi.doe.gov/fungi), and EnsembleFungi (fungi.ensembl.org)]



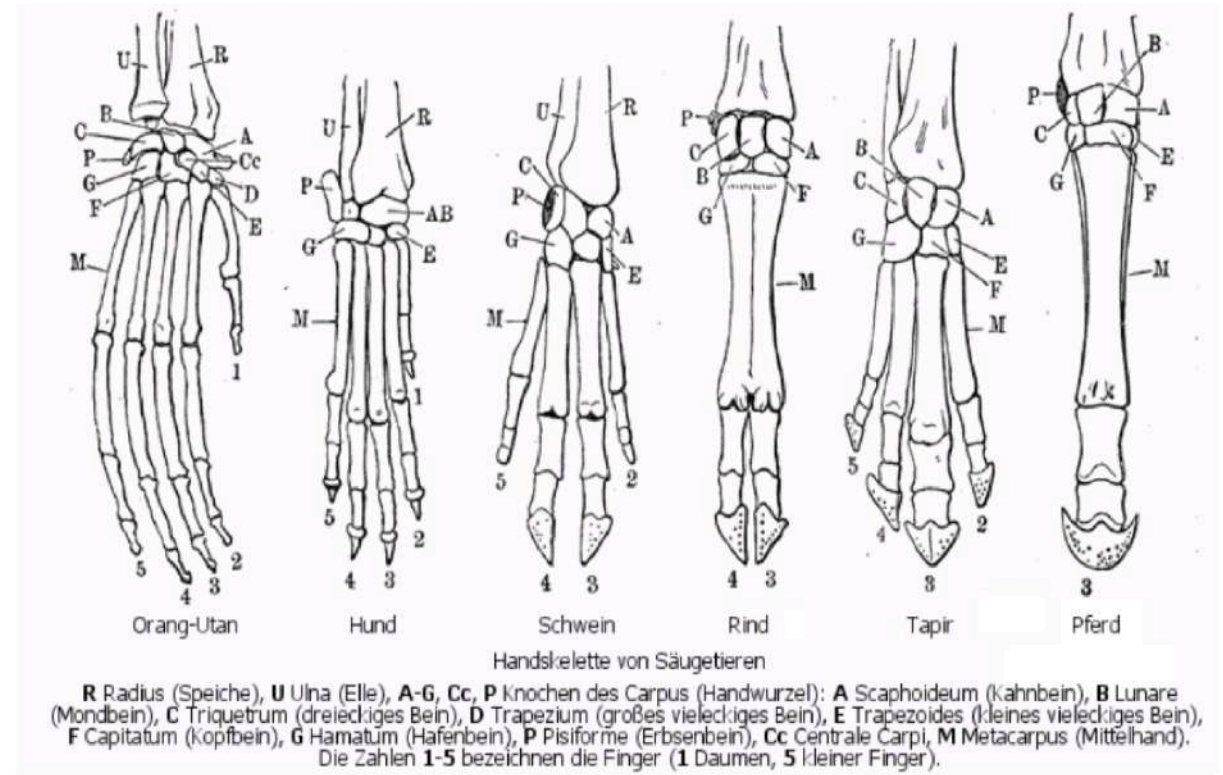
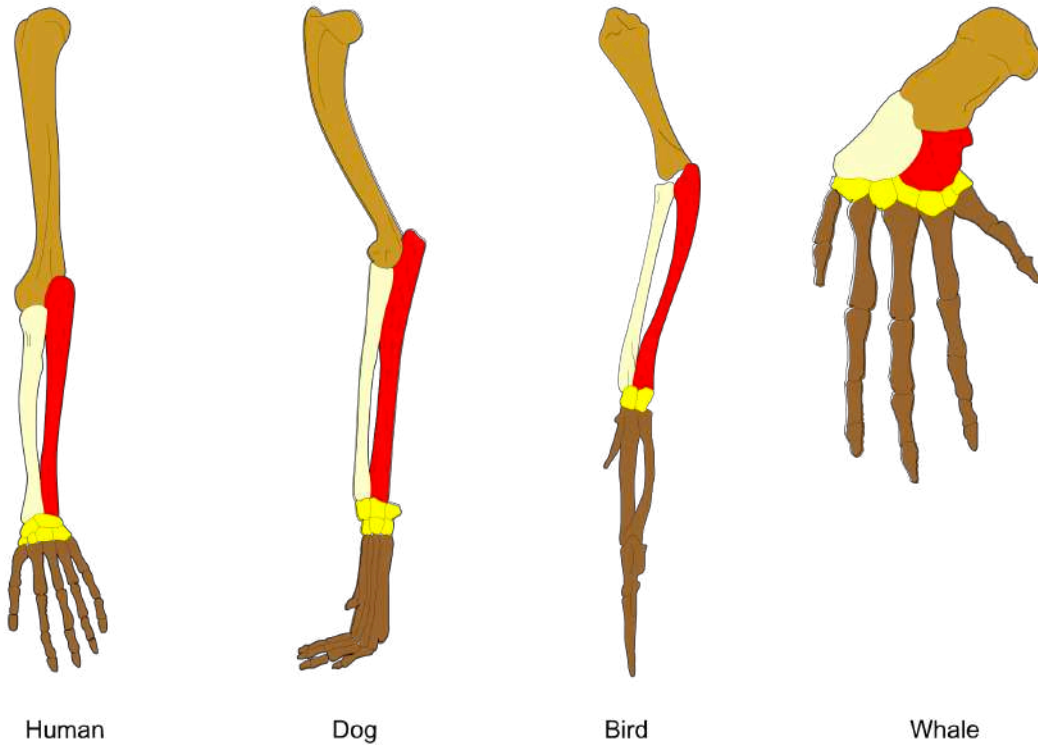
SCIENTIFIC REPORTS

OPEN The Genome Sequences of 90 Mushrooms

Huiying Li¹, Surui Wu³, Xiao Ma^{4,5}, Wei Chen^{2,4}, Jing Zhang⁶, Shengchang Duan⁶, Yun Gao⁶, Ling Kui^{7,8}, Wenli Huang¹², Peng Wu⁴, Ruoyu Shi^{4,5}, Yifan Li⁵, Yuanzhong Wang⁹, Jieqing Li⁹, Xiang Guo³, Xiaoli Luo³, Qiang Li¹², Chuan Xiong¹², Honggao Liu⁹, Mingying Gui³, Jun Sheng^{4,5} & Yang Dong^{2,10,11}

Second – assign orthology and phylogenetic position of the species

The term **homology** was first coined before Darwin

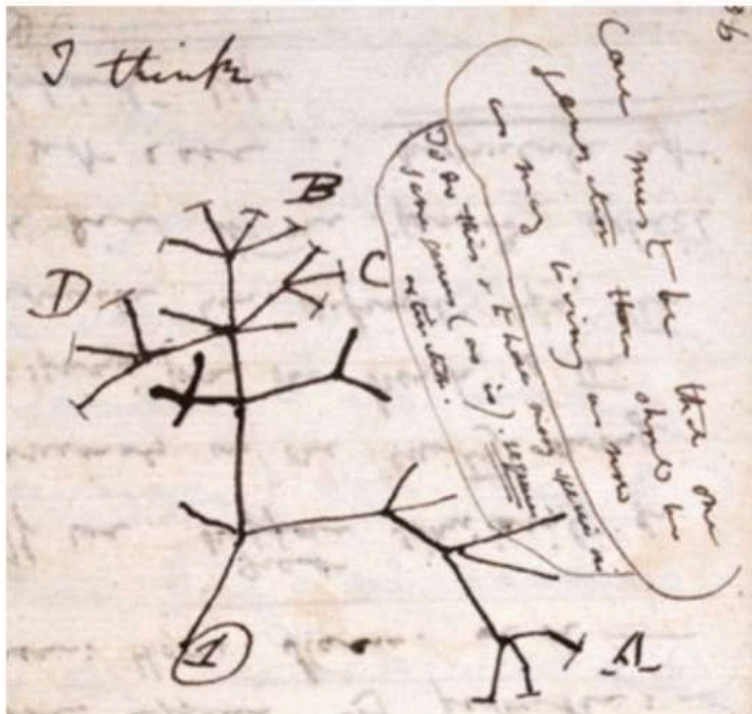


“the same organ in different animals under every variety of form and function” – Richard Owen

Owen 1843, p.379

[https://en.wikipedia.org/wiki/Homology_\(biology\)](https://en.wikipedia.org/wiki/Homology_(biology))

Darwin later reformulated homology as a result of “**descent with modification**”, i.e., share ancestry



CHAPTER VI.
DIFFICULTIES ON THEORY.
Difficulties on the theory of descent with modification—Transitions—Absence or rarity of transitional varieties—Transitions in habits of life—Diversified habits in the same species—Species with habits widely different from those of their allies—Organs of extreme perfection—Means of transition—Cases of difficulty—Natura non facit saltum—Organs of small importance—Organs not in all cases absolutely perfect—The law of Unity of Type and of the Conditions of Existence embraced by the theory of Natural Selection, 154

CHAPTER XIII.
MUTUAL AFFINITIES OF ORGANIC BEINGS: MORPHOLOGY: EMBRYOLOGY: RUDIMENTARY ORGANS.
CLASSIFICATION, groups subordinate to groups—Natural system—Rules and difficulties in classification, explained on the theory of descent with modification—Classi-

Refining *how* homologous genes are related

DISTINGUISHING HOMOLOGOUS FROM ANALOGOUS PROTEINS (1970)

WALTER M. FITCH



1929 - 2011

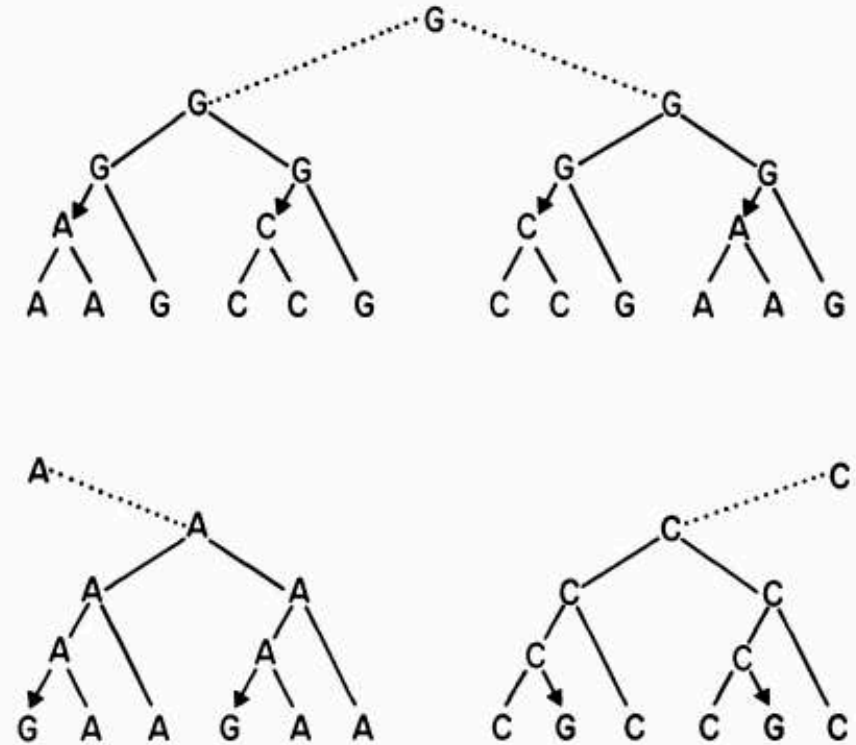


FIG. 1.—Distinguishing convergent from divergent types of nucleotide replacement patterns. Given are two groups of species (related within each group as shown by the solid lines) together with the nucleotide present at a specific position of the gene for each member species as shown at the branch tips. Given also the requirement that the ancestral nucleotide must permit the descendant nucleotides to be obtained in the minimum number of replacements, the ancestral nucleotide of the upper two groups must be set as G, with the required replacements indicated by the arrows. Were one to postulate a common ancestor for the two groups, no new mutations would need to be assumed; hence, this kind of pattern is called the divergent types. The lower two groups are identical except for rearranging the nucleotides at the branch tips, but now, in order to account for descendants in only four nucleotide replacements, the ancestral nucleotide of the lower two groups must be A and C. To postulate a common ancestor for these two groups would require, unlike the upper pair, an additional mutation. This situation shows different ancestral characters apparently converging toward the same descendant character, and hence is called the convergent type. One can calculate the frequency with which one might expect each type to be found in examining a large number of such nucleotide positions and compare that value to what is in fact found for a particular set of proteins. An abnormally large number of either type is evidence favoring that type of relation between the two groups examined.

From homology to orthology

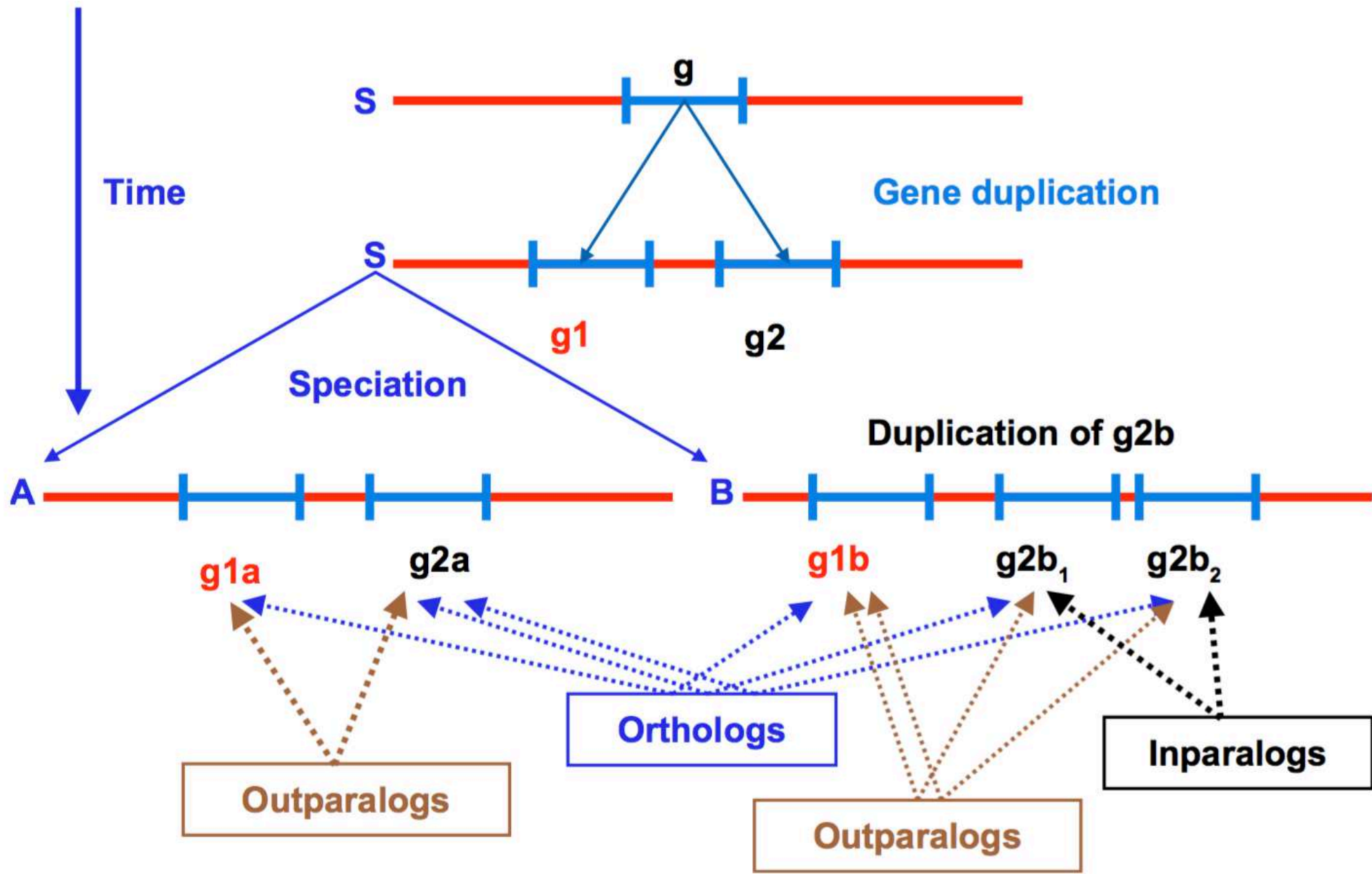
Homologues are sequences derived from a common ancestor...

- What are then orthologues? and paralogues?

Original definition of orthology and paralogy by Walter Fitch
(1970, Systematic Zoology 19:99-113):

*"Where the homology is **the result of gene duplication** so that both copies have descended side by side during the history of an organism, (for example, alpha and beta hemoglobin) the genes should be called **paralogous** (para = in parallel).*

*Where the homology is **the result of speciation** so that the history of the gene reflects the history of the species (for example alpha hemoglobin in man and mouse) the genes should be called **orthologous** (ortho = exact)."*

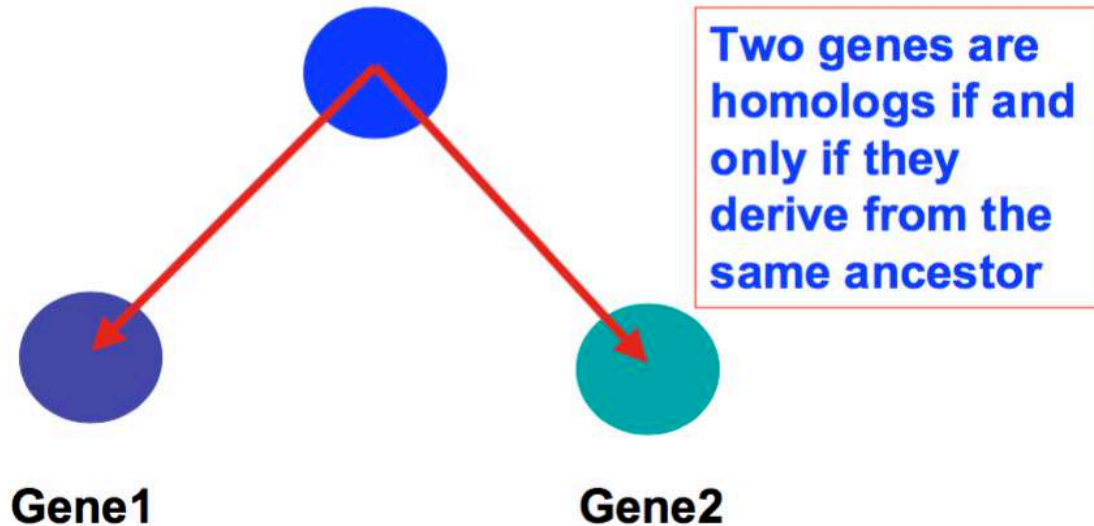


Search for similarity , collinearity, conservation of morphological characters

Search for similarity

One of the most frequent activity in Bioinformatics

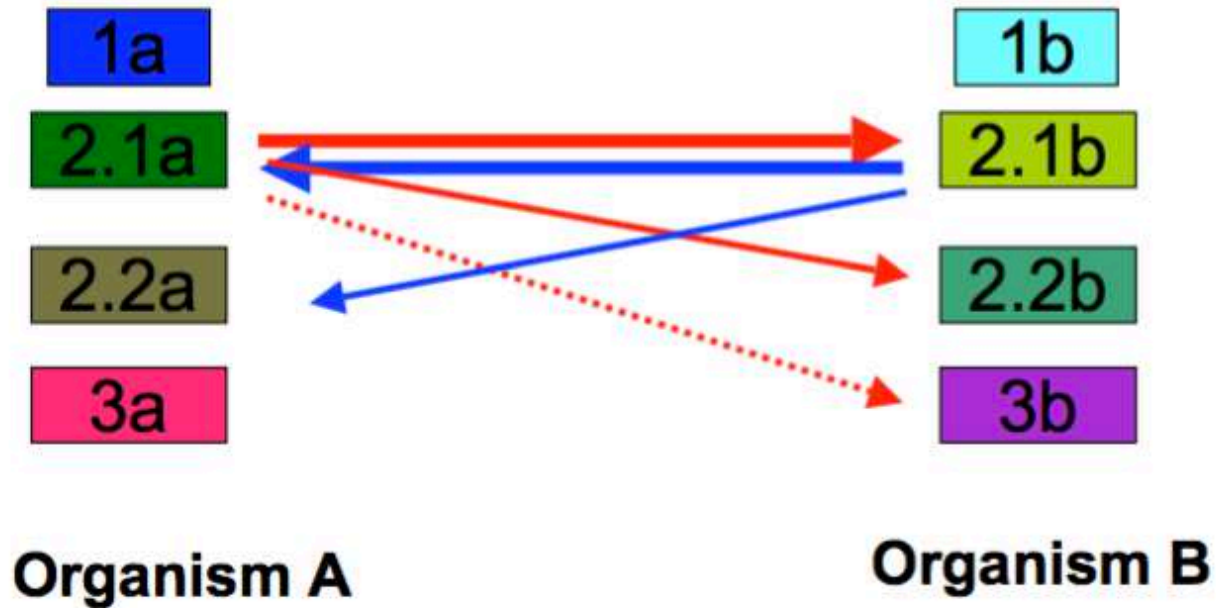
Common ancestor



Homology is almost uniquely inferred by sequence similarity

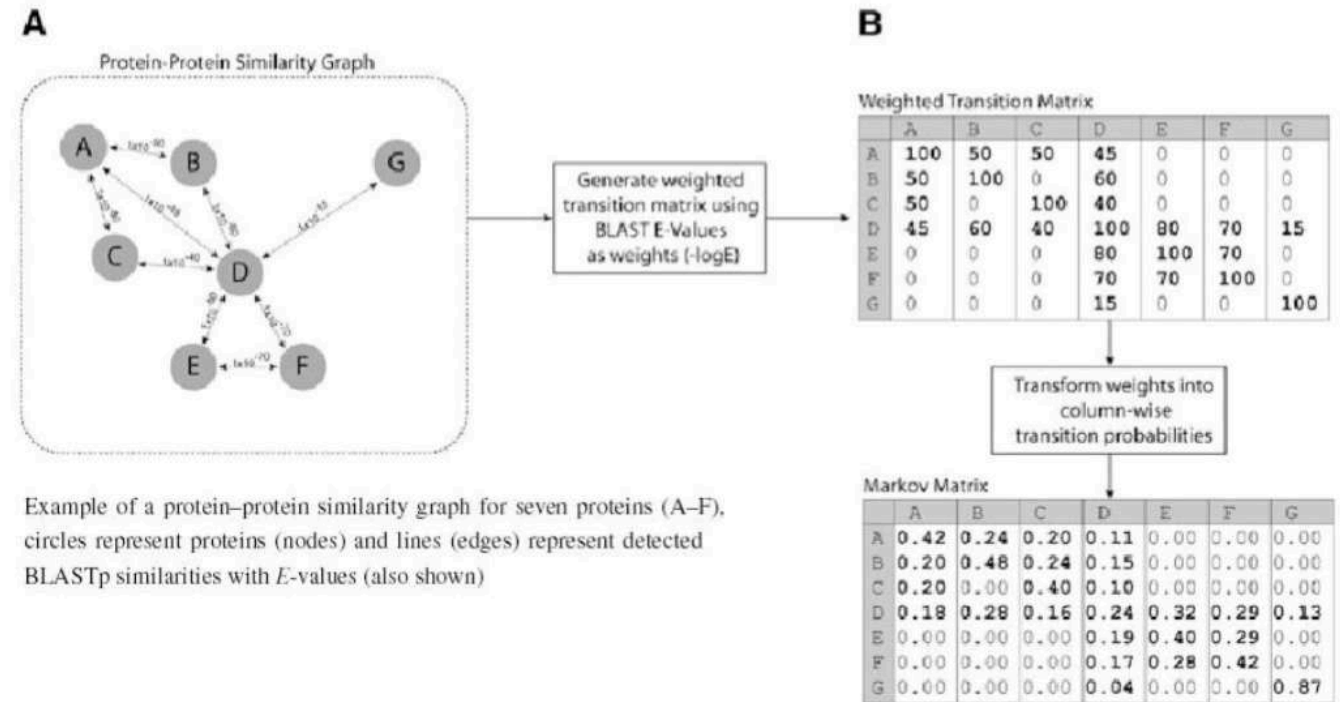
How to detect orthologous genes?

- The most intuitive way: **Best Reciprocal Hit (RBH)**



Sequence by clustering

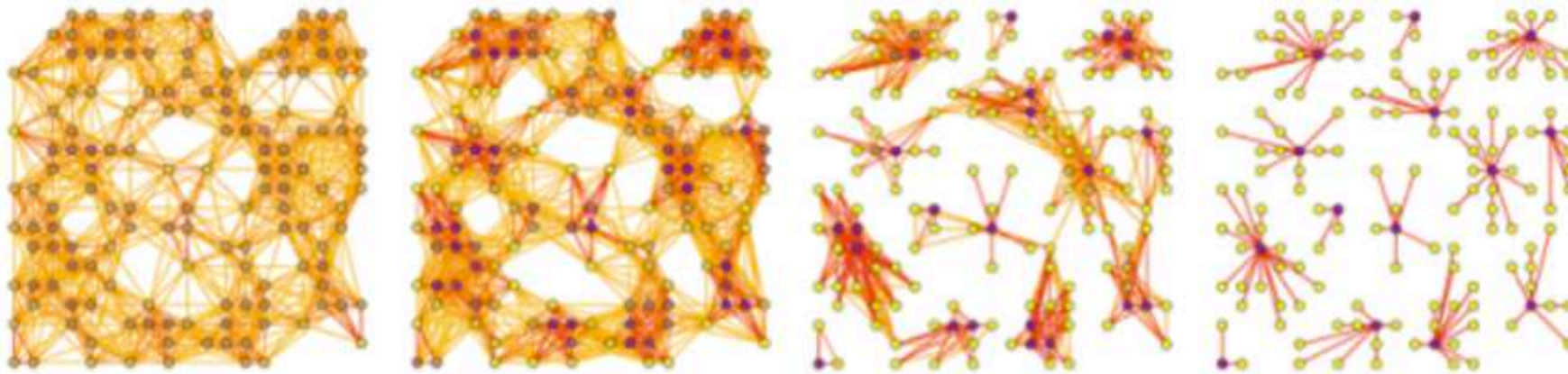
mcl: The Markov Cluster Algorithm <http://micans.org/mcl/> (Stijn Van Dongen)



Example of a protein-protein similarity graph for seven proteins (A-F), circles represent proteins (nodes) and lines (edges) represent detected BLASTp similarities with *E*-values (also shown)

Weighted transition matrix and associated column stochastic Markov matrix for the seven proteins shown in (A).

Produce clusters (gene families) using different inflation parameter



Orthology prediction methods

Similarity

Rely on genome comparisons and clustering of highly similar genes to identify orthologous groups **(suitable for large genome datasets)**

Phylogeny

use candidate gene families determined by similarity and then rely on the reconciliation of the phylogeny of these genes with their corresponding species phylogeny to determine the subset of orthologs

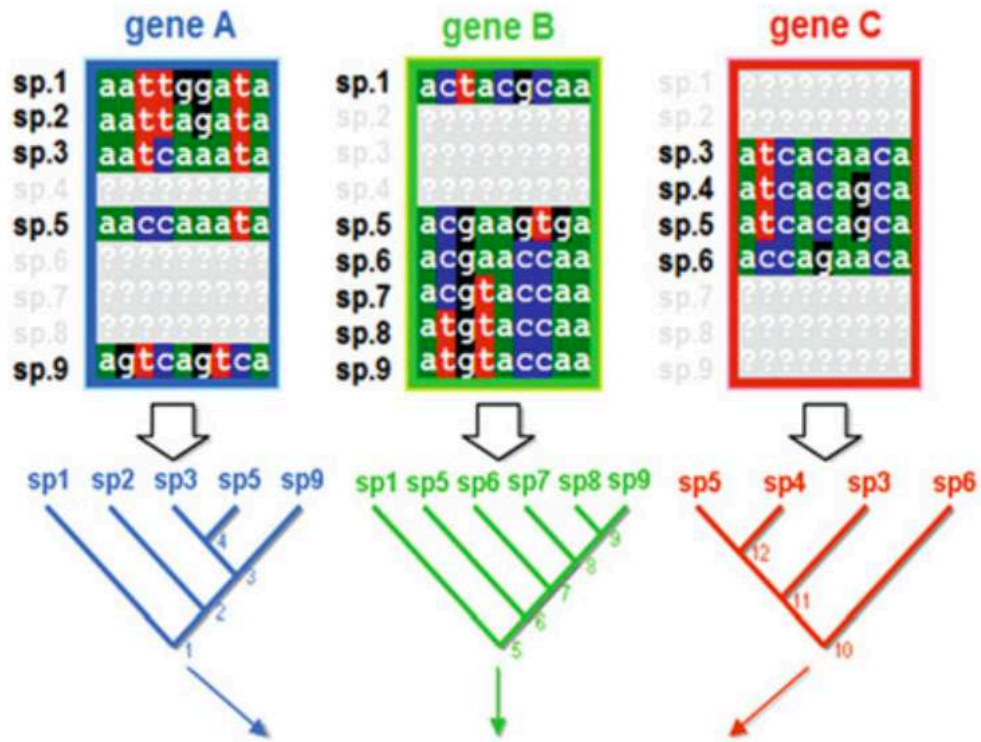
(Good and more interpretable for small set of genomes)

Others

Combination of (1) and (2)

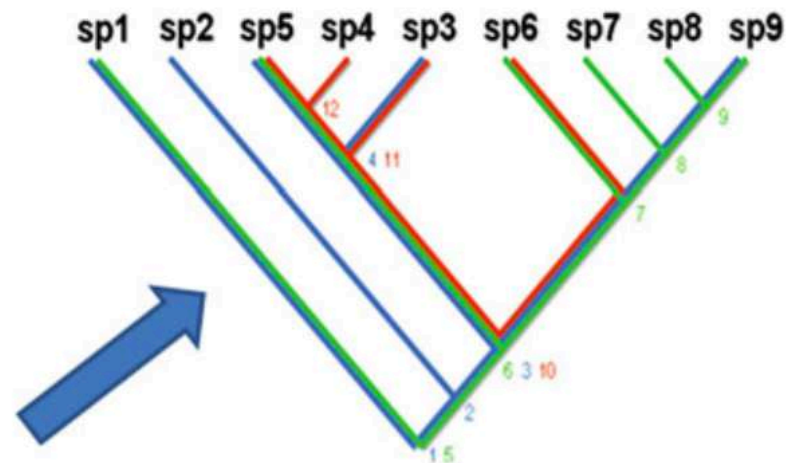
Some uses synteny

From gene trees to species tree

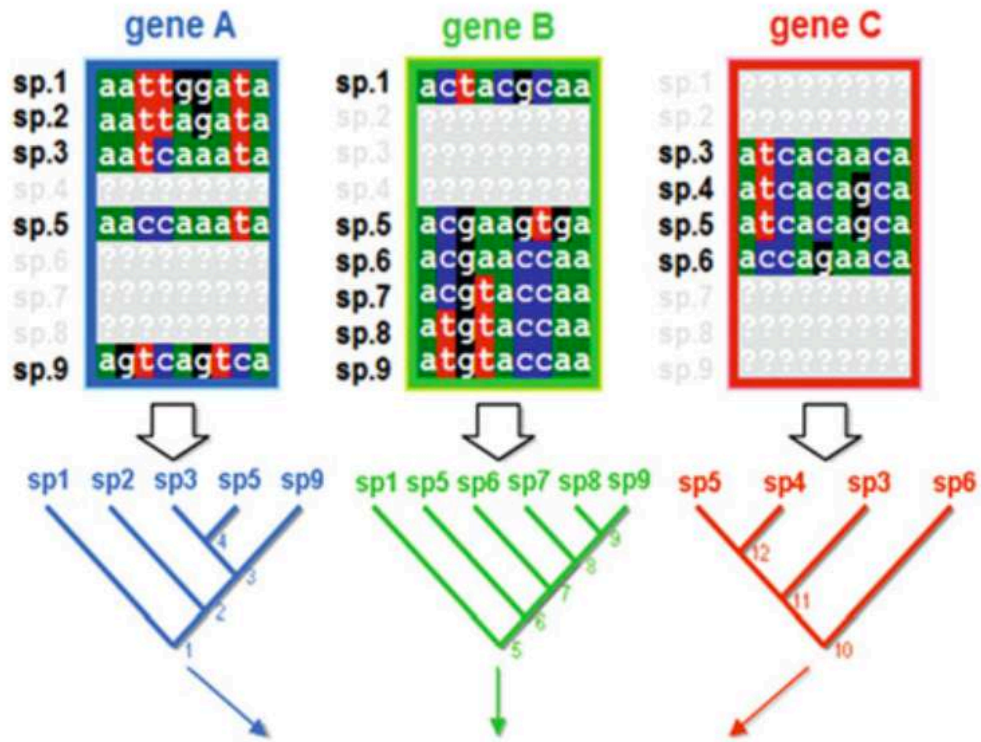


MRP	1	2	3	4	5	6	7	8	9	10	11	12
sp1	1	0	0	0	1	0	0	0	0	?	?	?
sp2	1	1	0	0	?	?	?	?	?	?	?	?
sp3	1	1	1	1	?	?	?	?	?	1	1	0
sp4	?	?	?	?	?	?	?	?	?	1	1	1
sp5	1	1	1	1	1	1	0	0	0	1	1	1
sp6	?	?	?	?	1	1	1	1	0	1	0	0
sp7	?	?	?	?	1	1	1	1	0	?	?	?
sp8	?	?	?	?	1	1	1	1	1	?	?	?
sp9	1	1	1	0	1	1	1	1	1	?	?	?

Supertree methods infer the best topology for each gene (using the same phylogenetic method for each), and then a topological consensus is obtained. Such methods are able to make consensus trees even if the number of leaves among gene trees differs but overlaps to some extent, for example when a gene has not been sequenced for some taxa

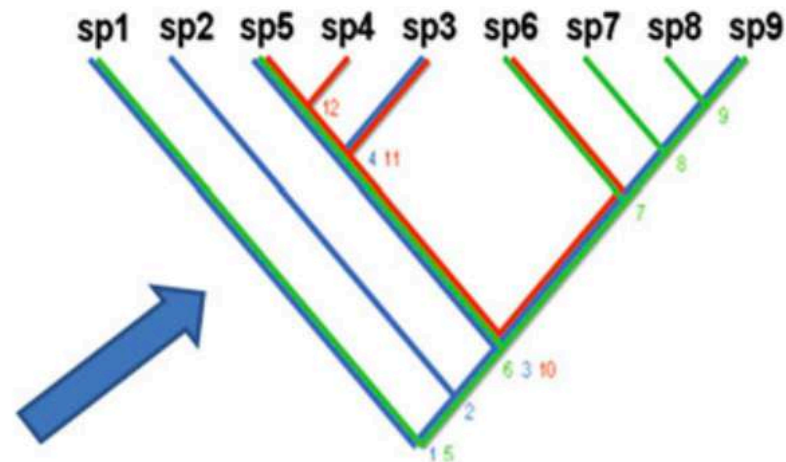


From gene trees to species tree



MRP	1	2	3	4	5	6	7	8	9	10	11	12
sp1	1	0	0	0	1	0	0	0	0	?	?	?
sp2	1	1	0	0	?	?	?	?	?	?	?	?
sp3	1	1	1	1	?	?	?	?	?	1	1	0
sp4	?	?	?	?	?	?	?	?	?	1	1	1
sp5	1	1	1	1	1	1	0	0	0	1	1	1
sp6	?	?	?	?	1	1	1	1	0	1	0	0
sp7	?	?	?	?	1	1	1	1	0	?	?	?
sp8	?	?	?	?	1	1	1	1	1	?	?	?
sp9	1	1	1	0	1	1	1	1	1	?	?	?

A step beyond supertrees is the use of methods that take into consideration specific evolutionary processes that may be responsible for differences in gene topologies, and then estimate the species tree which would most likely have generated such gene trees, under different scenarios

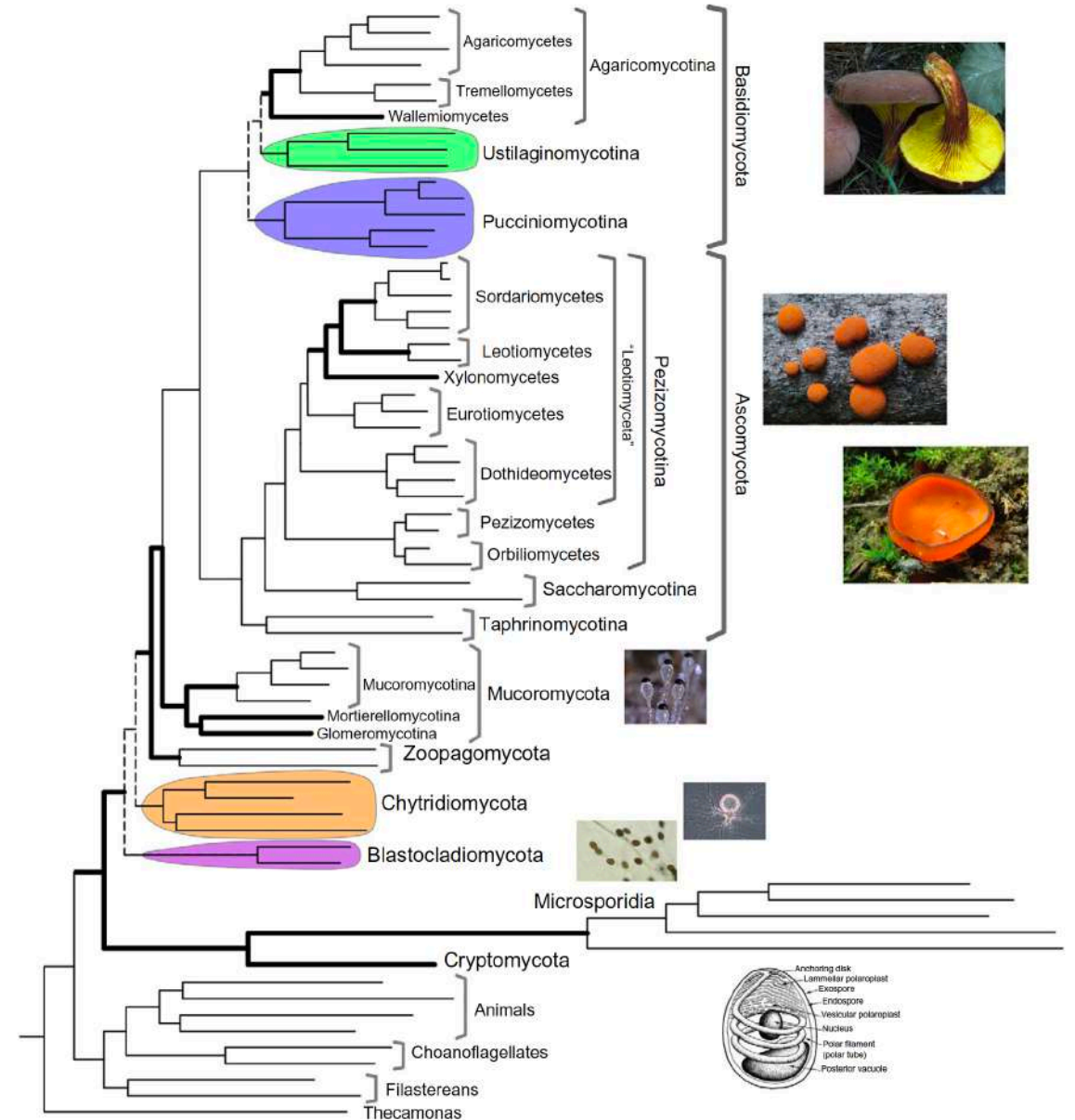


Importance of genomics in fungi - Phylogeny

The accelerating pace of fungal genome sequencing by a number of large-scale sequencing projects **paved the way for assembling larger and taxon-specific datasets that clarified some of the puzzling fungal relationships.**

- sister relationship between the Leotiomycetes and Sordariomycetes
- resolved the position of the Xylonomycetes, a small class of leaf endophytes, as a sister group to the Lecanoromycetes and Eurotiomycetes.

Resolving ancient divergences poses significant challenges even for phylogenomic datasets. For example, the definition of the fungal kingdom and the placement of the Microsporidia as fungi or nonfungal eukaryotes have been debated. (to be discussed in the next lecture)



3rd – Comparative genomics

Why comparative genomics?

Compare multiple genomes now a norm

Similarity and differences between genomes

Reveal the evolutionary relationships among species

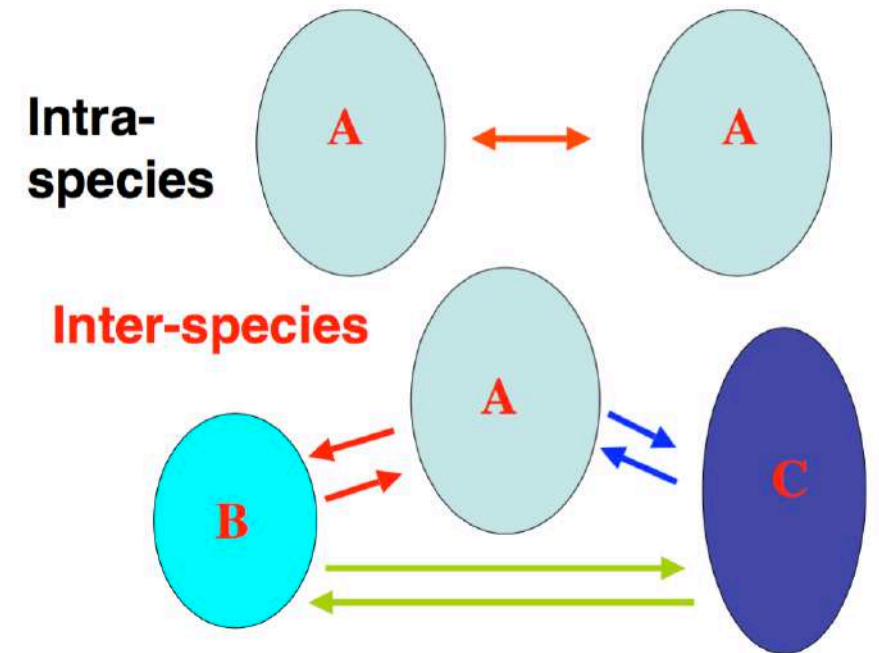
Link evolutionary processes with function

Use genomes to study evolution of these species:

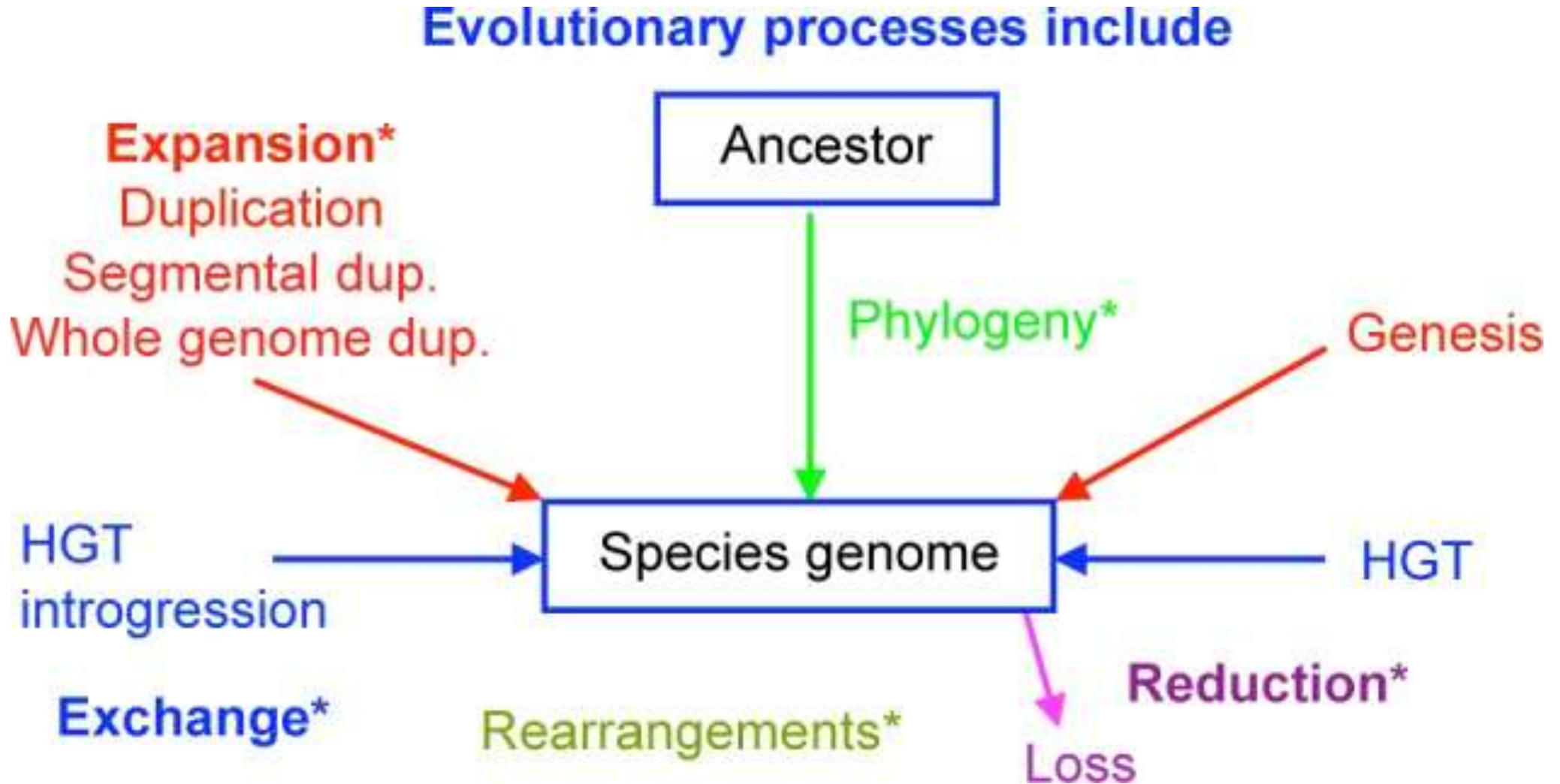
- At various resolution (whole genome, chromosomes, regions, genes, base pairs)
- Identify the genomic basis of key phenotypes

Comparing genomes

- Alignment of homologous regions
 - **Inter**-genomic: aligning genomic sequences from **different** species
 - **Intra**-genomic aligning genomic sequences from the **same** species
- Different levels of **resolution**
 - Comparative mapping (markers)
 - Synteny (~ gene content)
 - Colinearity (gene content + order conservation)
 - DNA-based alignments (base-to-base mapping)



Evolution process of a genome

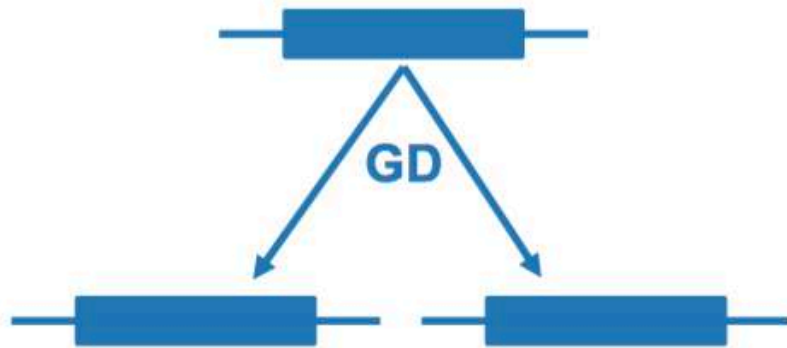


Sources of gene innovation

(Intuitive as genome gain genes of new functions)

Gene duplication (GD)

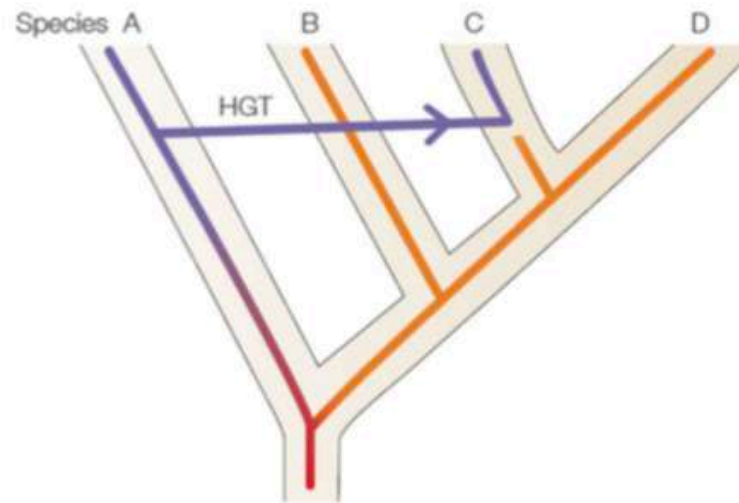
Any duplication of a region of DNA that contains a gene



- ❖ Plant organic material decay
- ❖ Starch catabolism
- ❖ Degradation of host tissues
- ❖ Toxin production

Horizontal gene transfer (HGT)

Exchange of genes between organisms other than through reproduction



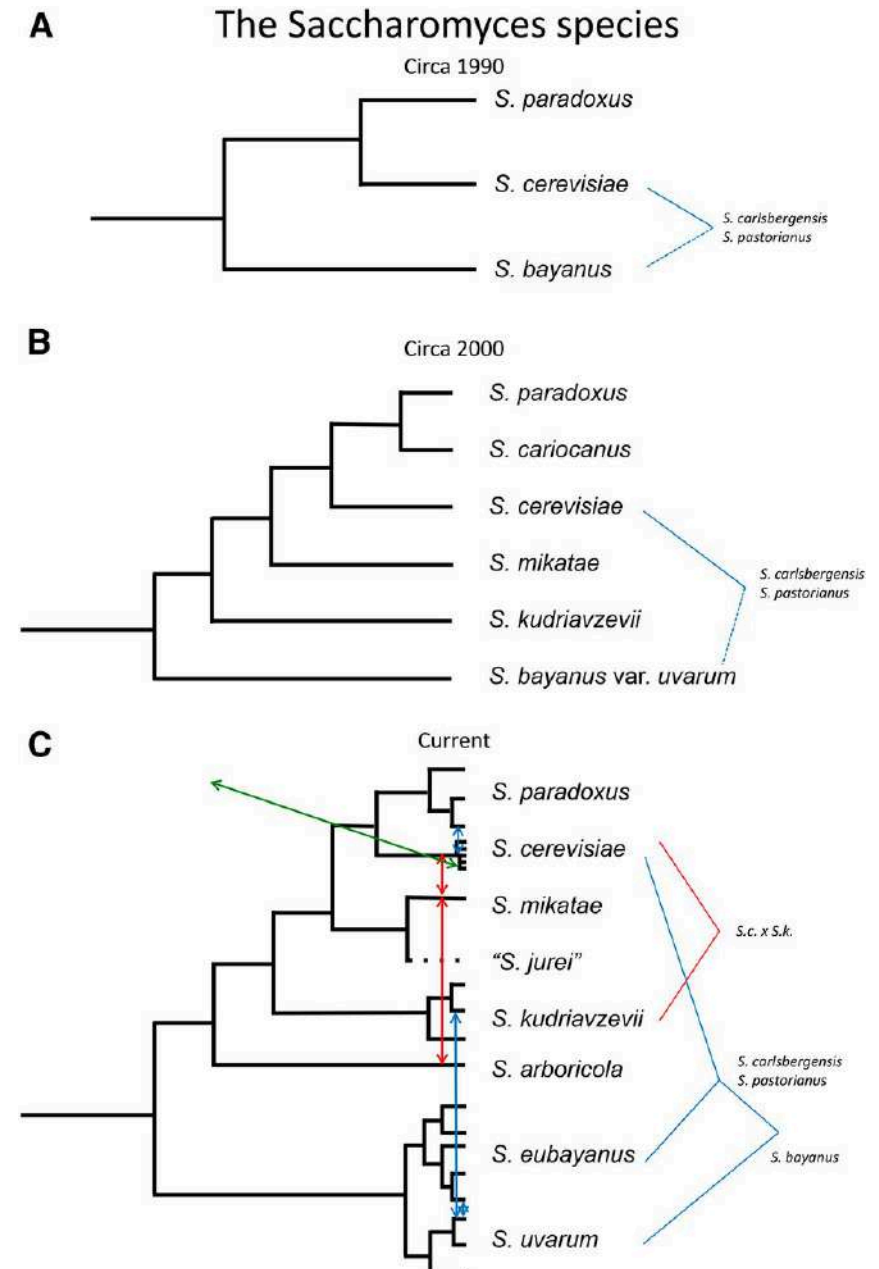
- ❖ Xenobiotic catabolism
- ❖ Toxin production
- ❖ Degradation of plant cell walls
- ❖ Wine fermentation

Importance of genomics in fungi - Comparative

- Comparative genomics analysis of *S. cerevisiae* and closely related species has contributed to our understanding of **how new species emerge and has shed light on the various mechanisms that contribute to reproductive isolation.**
- Genomic analysis of *Saccharomyces* yeasts has provided a better understanding of the mechanisms underlying large-scale genomic changes, such as polyploidy, and their consequences for genome evolution and cell physiology.
- Genomic approaches are increasingly contributing to our understanding of how budding yeasts adapt to natural environments by identifying the genes that are involved in adaptation within natural substrates.

Dujon and Louis (2017) Genetics

Marsit *et al* (2017) Nature Genetics Review



Importance of genomics in fungi - Comparative

- Fungi contain a remarkable diversity of both primary and secondary metabolic pathways involved in ecologically specialized or accessory functions.
- Genes in these pathways are frequently physically linked on fungal chromosomes, forming metabolic gene clusters (MGCs).
- Improved knowledge of the evolutionary life cycle of MGCs will advance our understanding of the ecology of specialized metabolism and of the interplay between the lifestyle of an organism and genome architecture.

a Standard configuration of fungal MGCs

Primary metabolism

Nitrate assimilation MGC in *Ogataea polymorpha*



Secondary metabolism

Gliotoxin MGC in *Aspergillus fumigatus*



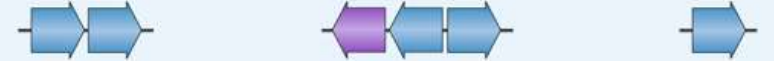
b Variation in gene content and structure among fungal MGCs

Primary metabolism

Galactose MGC and unclustered pathway genes in *Saccharomyces cerevisiae*

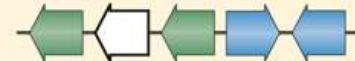


Biotin MGCs and unclustered pathway gene in *Saccharomyces cerevisiae*



Secondary metabolism

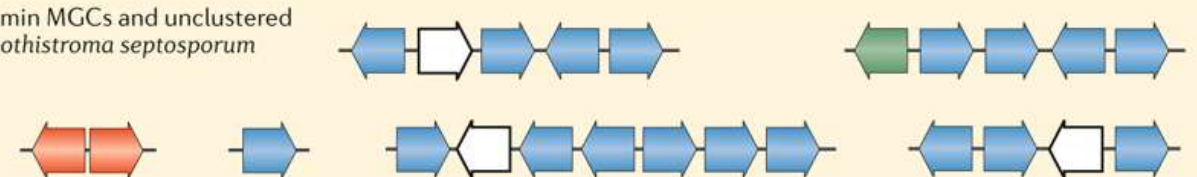
Emericellamide MGC in *Aspergillus nidulans*



Kojic acid MGC in *Aspergillus oryzae*



Dothistromin MGCs and unclustered gene in *Dothistroma septosporum*



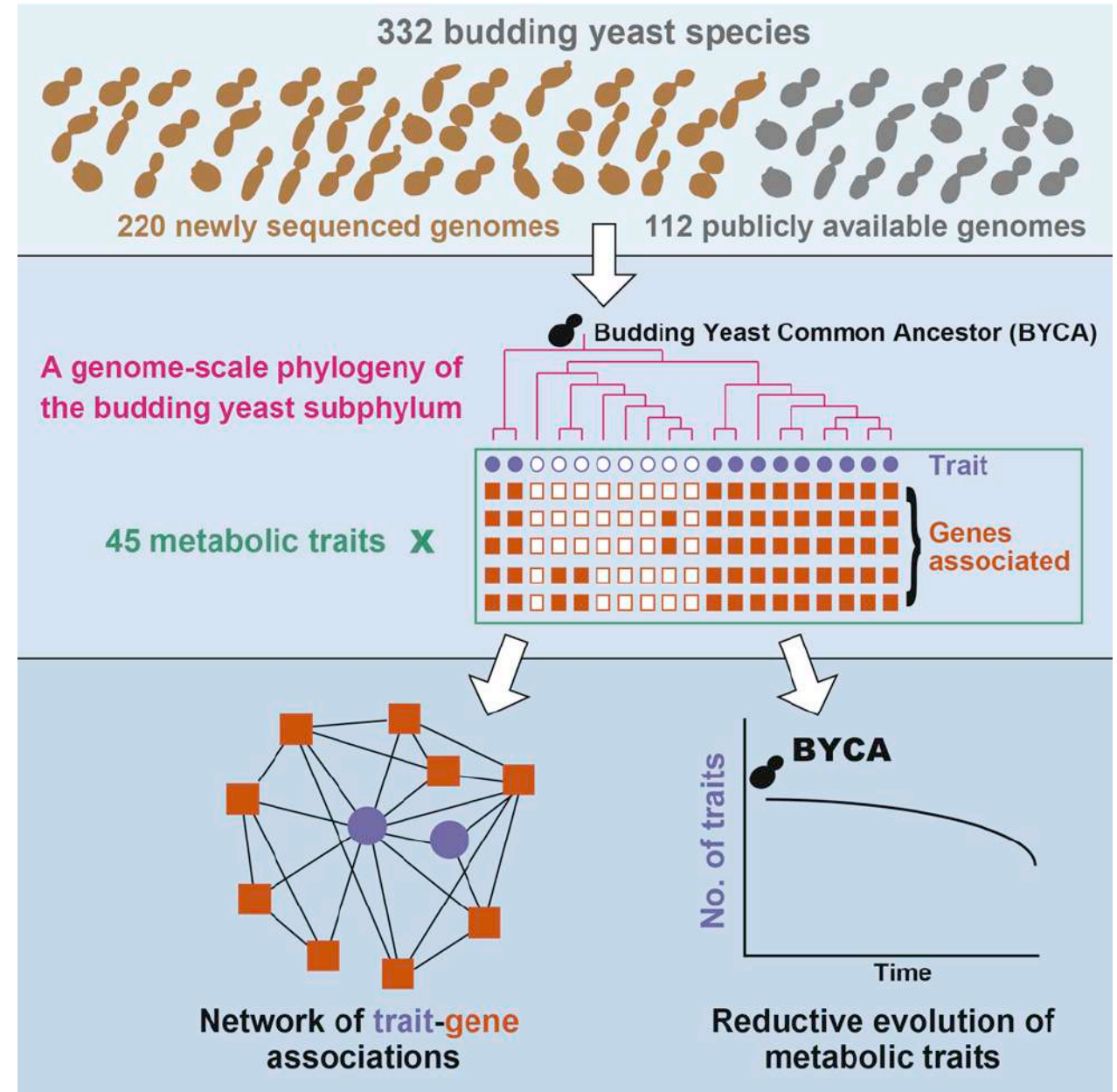
The intertwined fumagillin (F) and pseurotin (P) super MGC in *Aspergillus fumigatus*



Importance of genomics in fungi - Comparative (example)

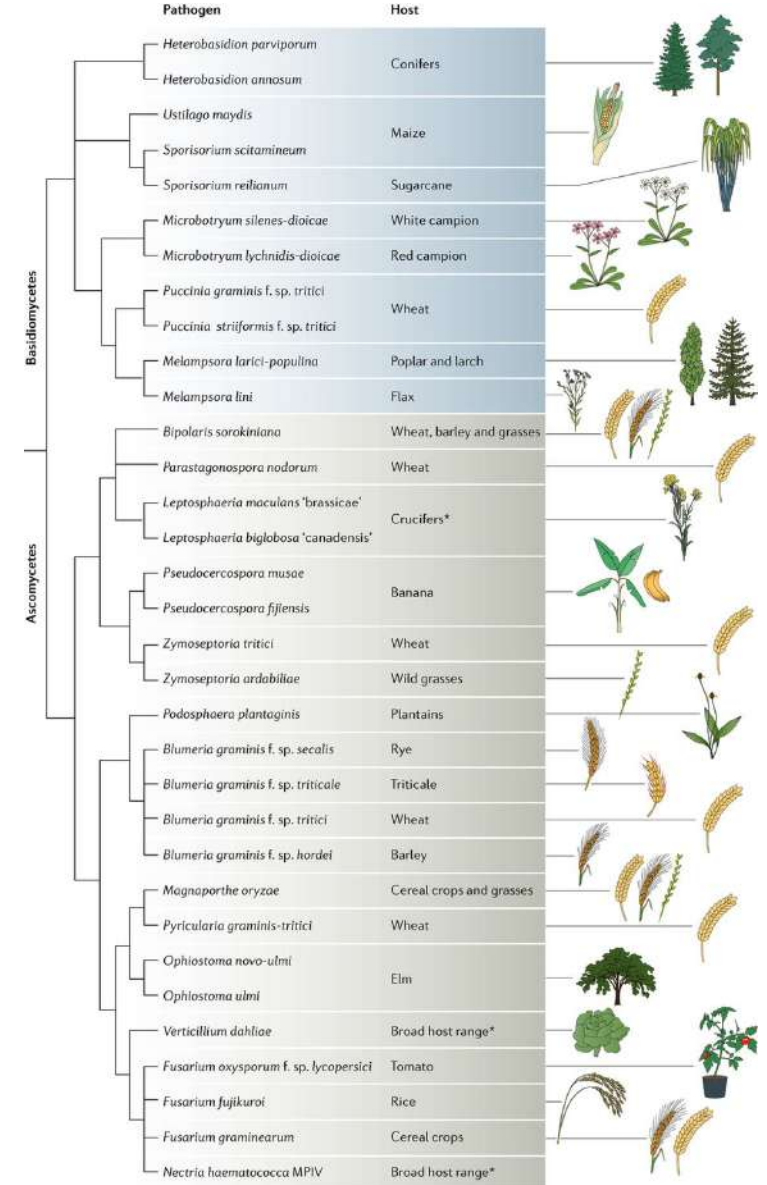
Highlights

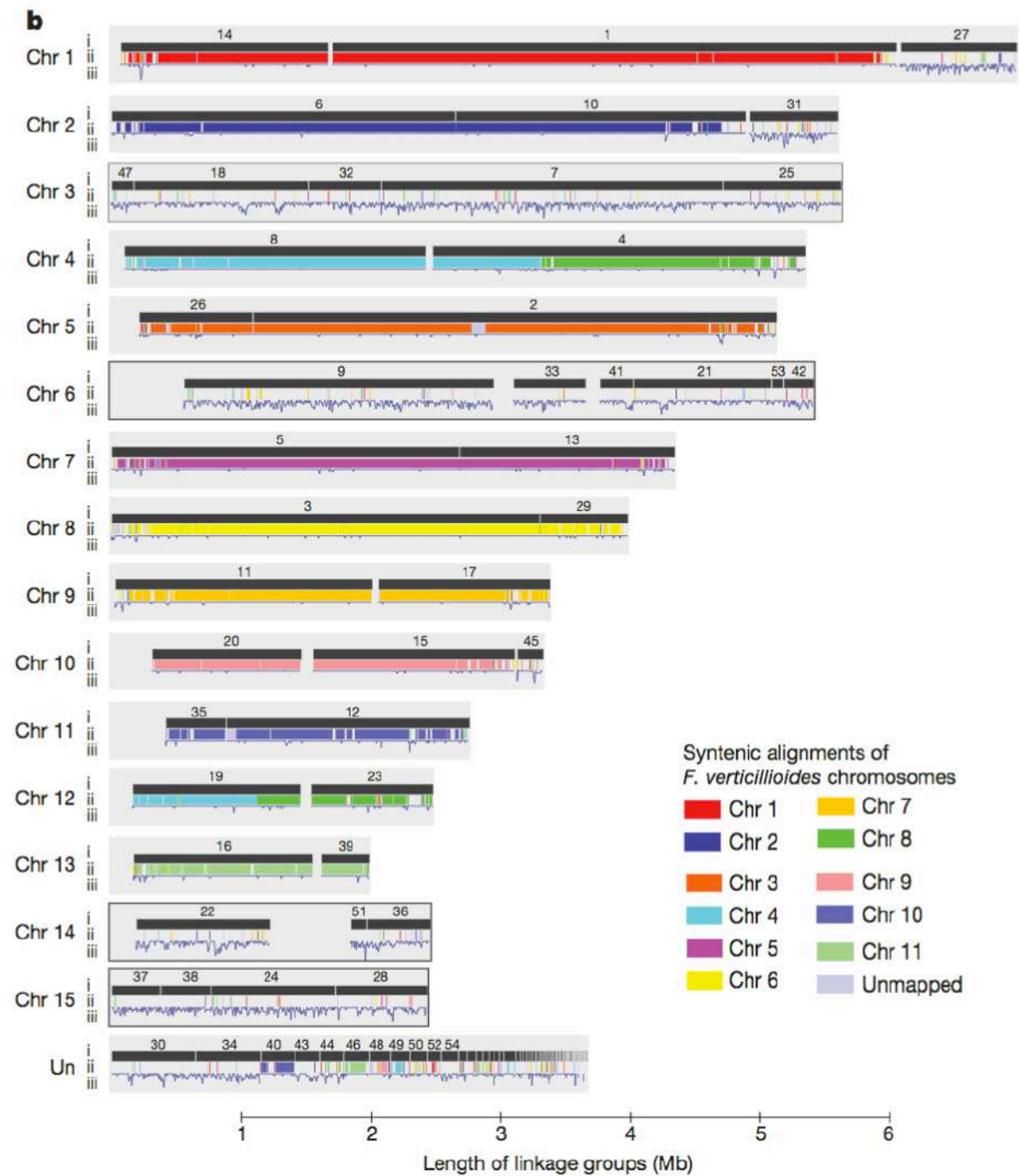
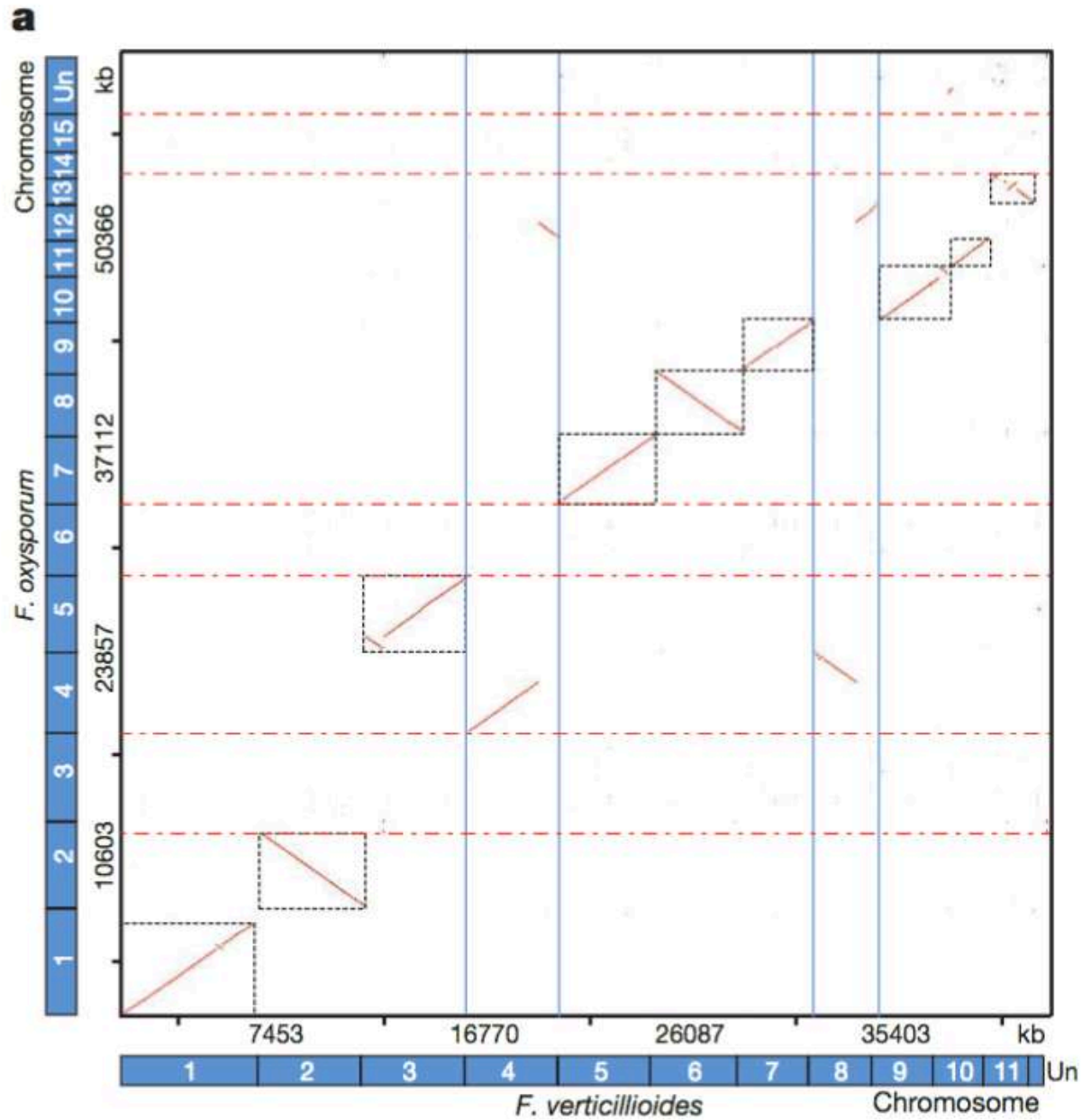
- 332 genomes, including 220 newly sequenced, covering ~1/3 of known budding yeasts
- Genome-scale inference of robust phylogeny and time tree of budding yeast subphylum
- Reconstruction of 45 metabolic traits infers complex budding yeast common ancestor
- Reductive evolution of traits and genes is a major mode of evolutionary diversification



Importance of genomics in fungi – plant pathogens

- The genomes of fungal plant pathogens can vary in size and composition, even between closely related species. Differences in the content of transposable elements cause variation in genome architecture.
- Variation in genome architecture results from differences in population genetic factors, including effective population size and the strength of genetic drift.
- During periods of low effective population size, non-adaptive mutations, such as transposable elements, can invade genomes and shape their architecture.
- Transposable elements contribute to the establishment and maintenance of rapidly evolving genome compartments that can comprise virulence genes. High mutation rates in these compartments support the evolution of new virulence phenotypes.

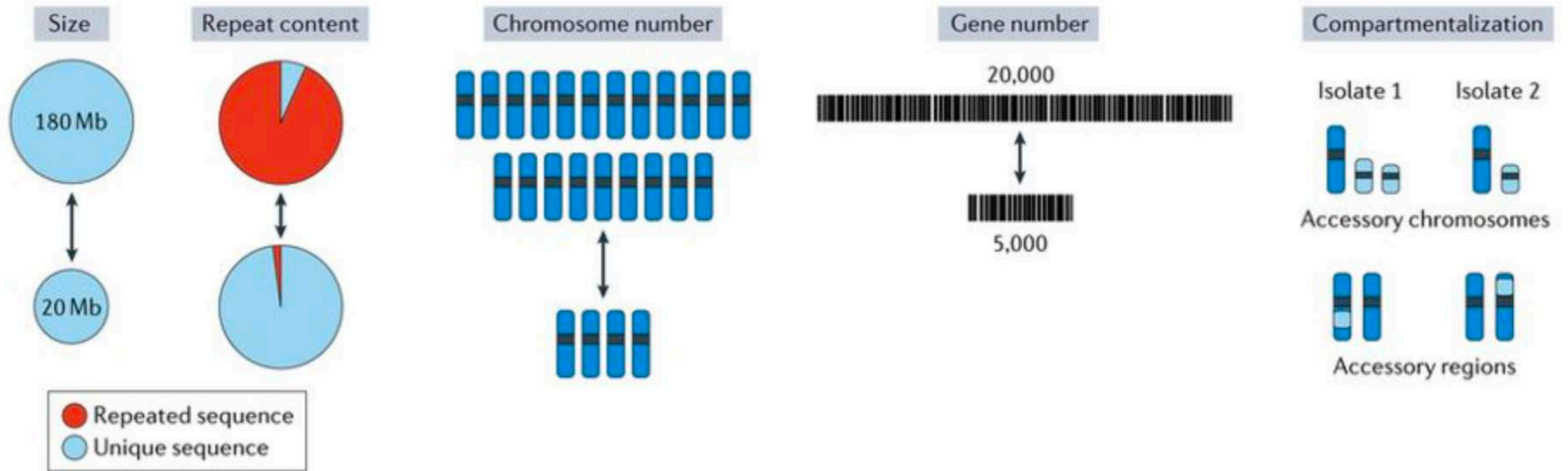




Importance of genomics in fungi – plant pathogens

- The genomes of fungal plant pathogens can vary in size and composition, even between closely related species. Differences in the content of transposable elements cause variation in genome architecture..

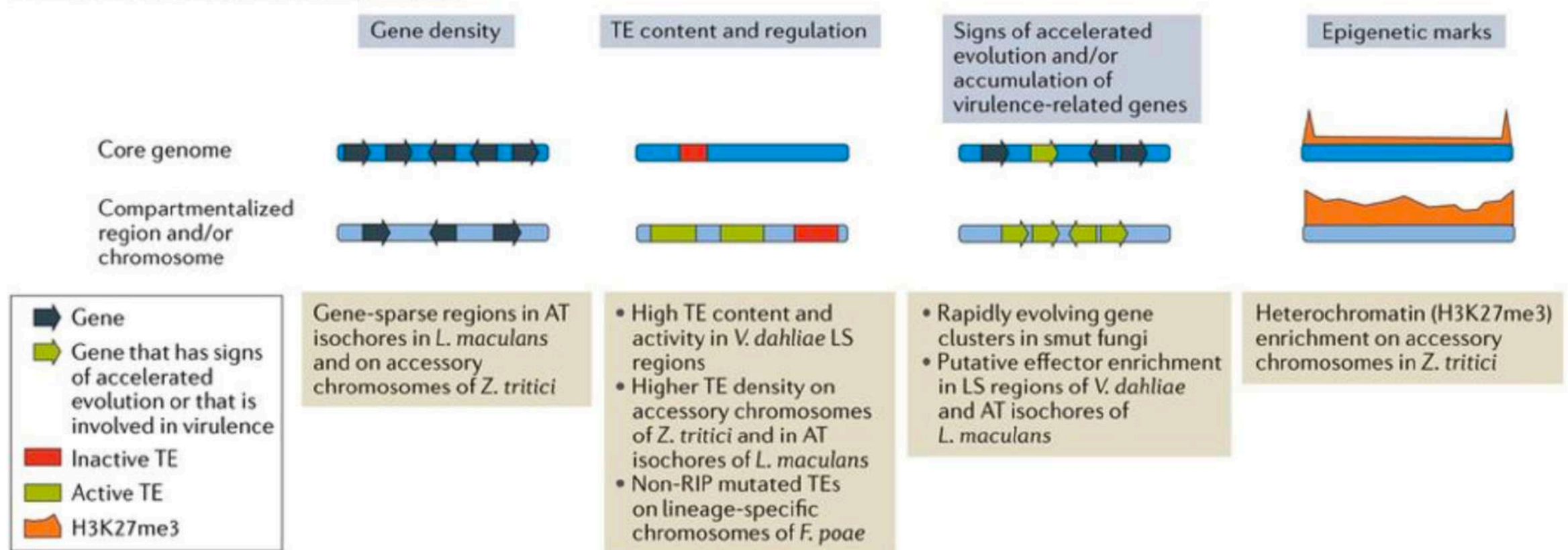
a Fungal plant pathogen genomes are highly diverse



Importance of genomics in fungi – plant pathogens

- The genomes of fungal plant pathogens can vary in size and composition, even between closely related species. Differences in the content of transposable elements cause variation in genome architecture..

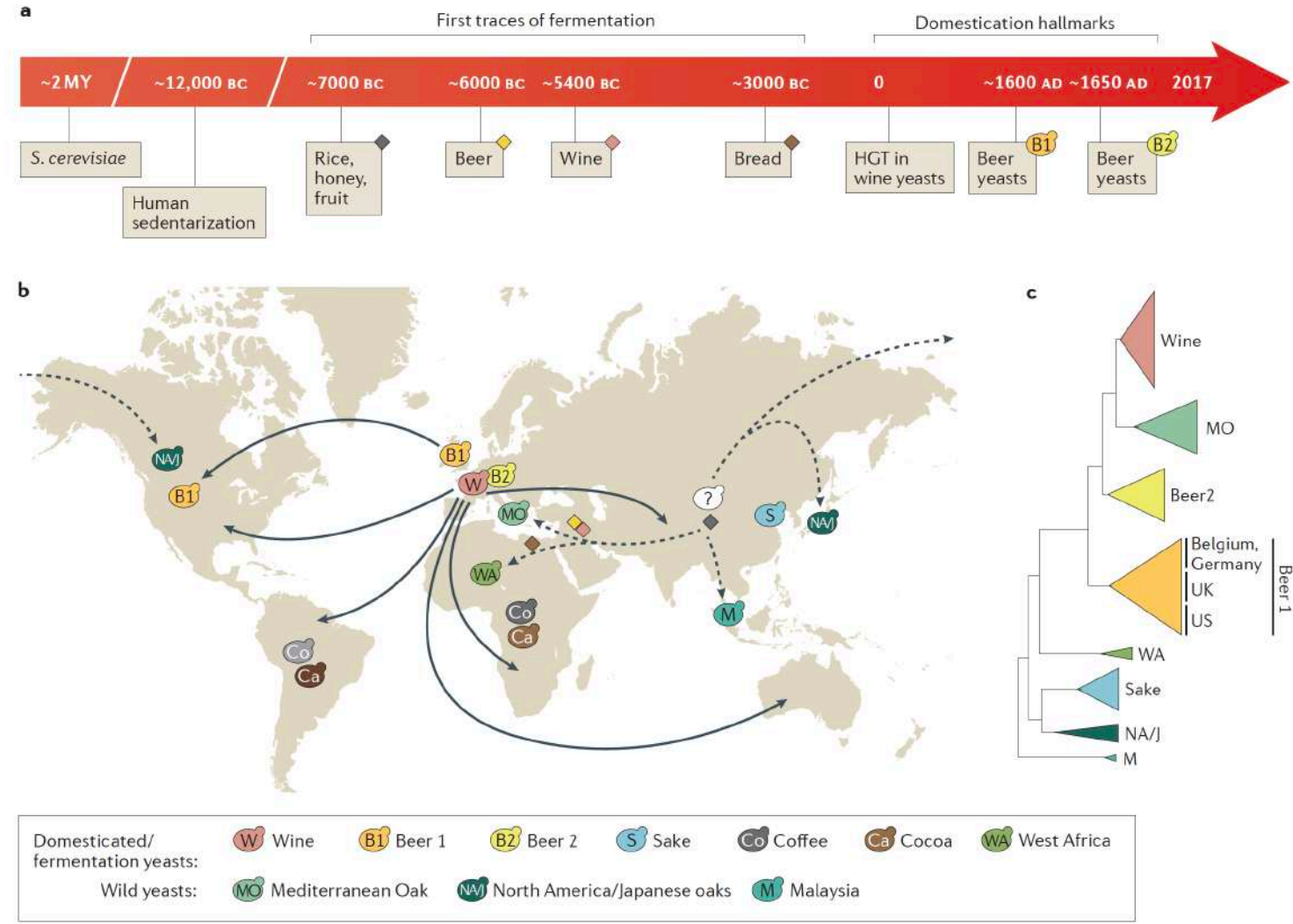
b Characteristics of genome compartments



Moving onto intraspecies variation - population genomics

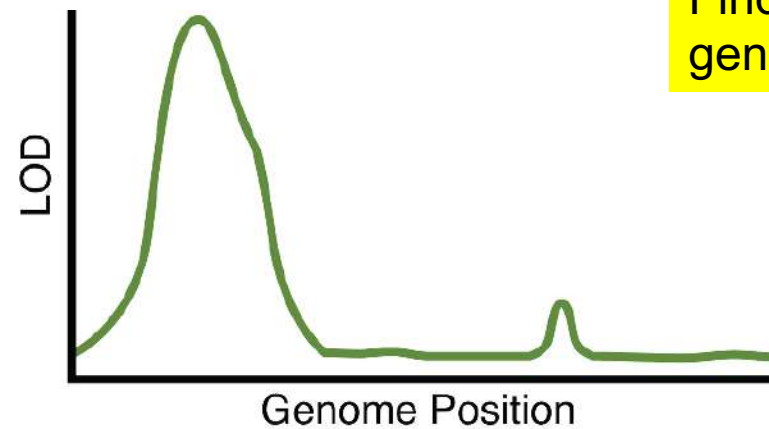
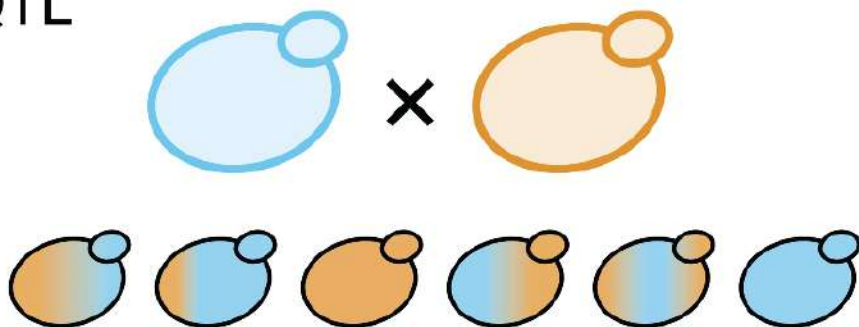
Importance of genomics in fungi - Population

- Population genomics and comparative genomics of *Saccharomyces* yeasts have revealed that hybridization occurred frequently throughout, and has had substantial effects on, yeast evolution. Hybridization could itself be a mechanism of adaptation and speciation.
- Genomic analysis of yeast strains associated with humans has revealed the history of yeast domestication and the mechanisms that have contributed to its adaptation to anthropogenic environments.

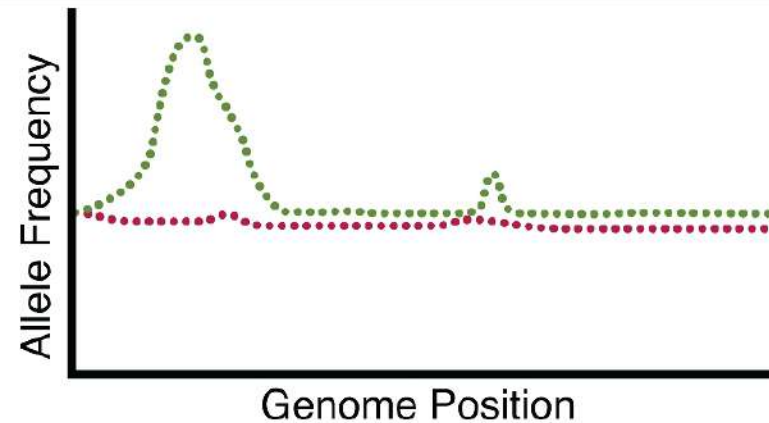
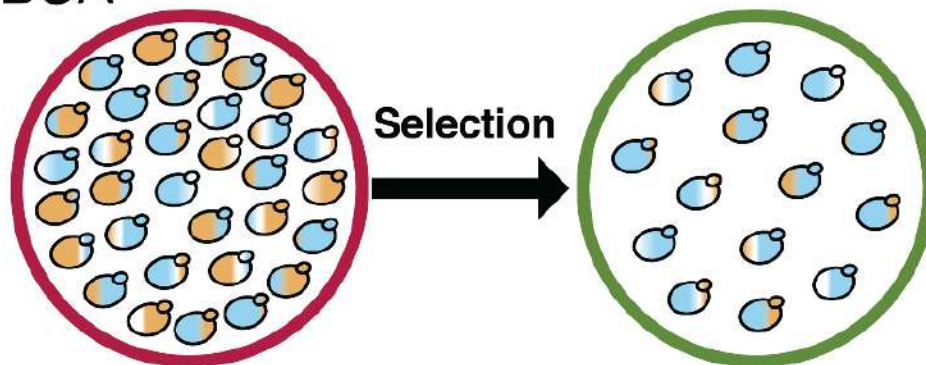


Once collection is available: genetic mapping

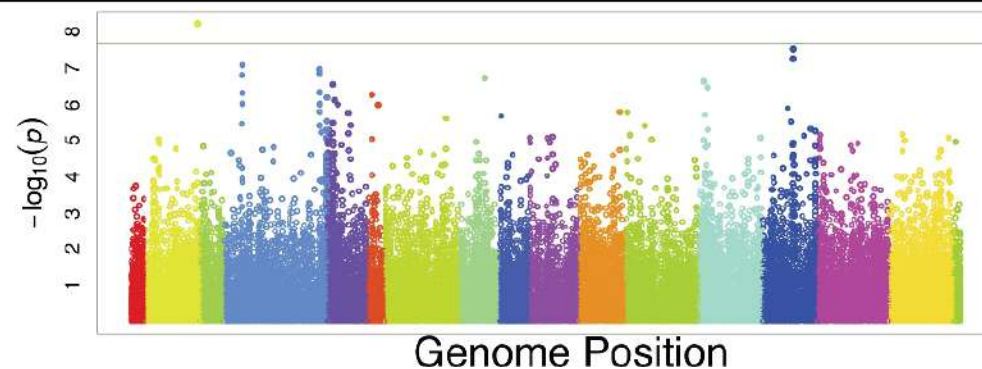
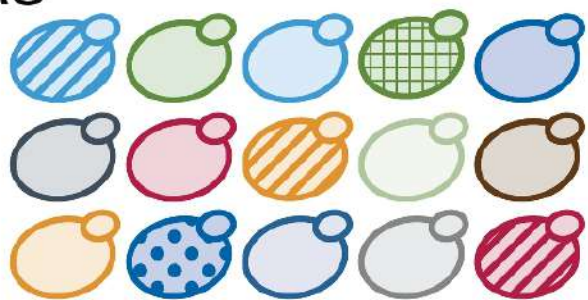
(A) QTL



(B) BSA



(C) GWAS

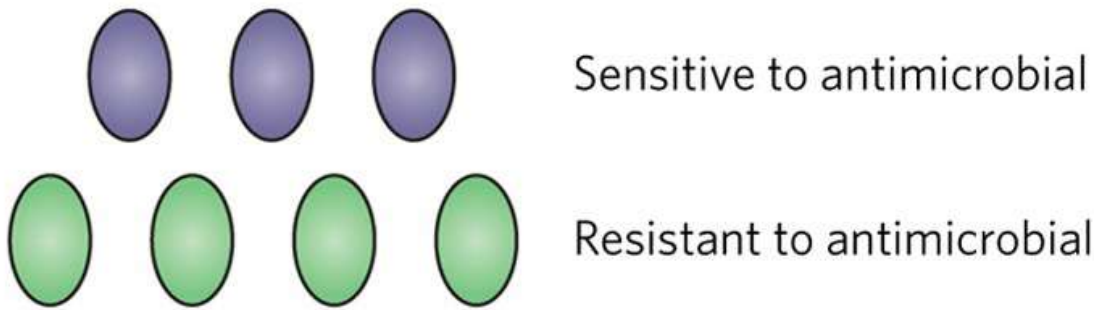


Genome-wide association study (GWAS) in microbes

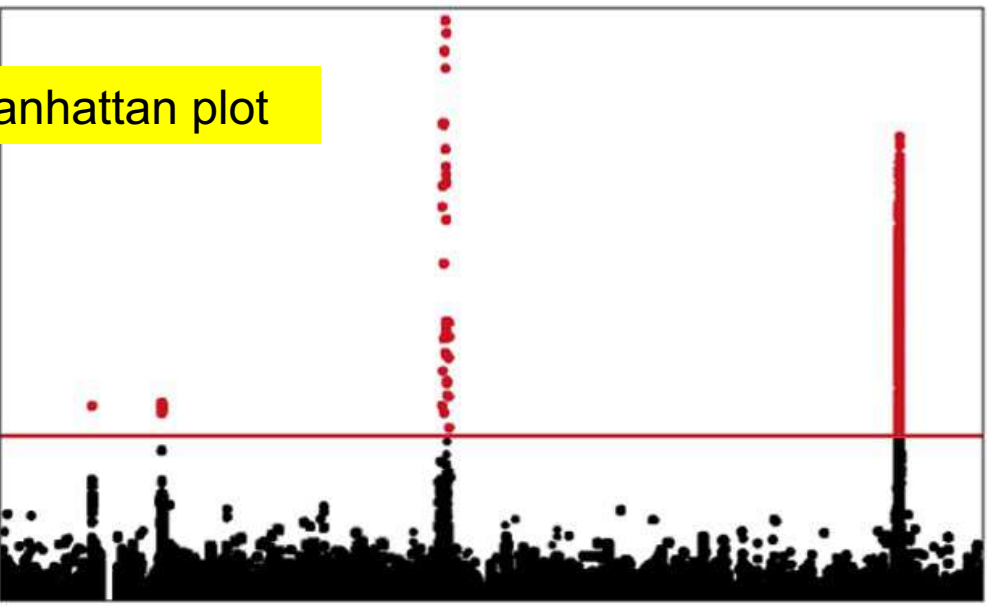
GWAS in humans



GWAS in bacteria

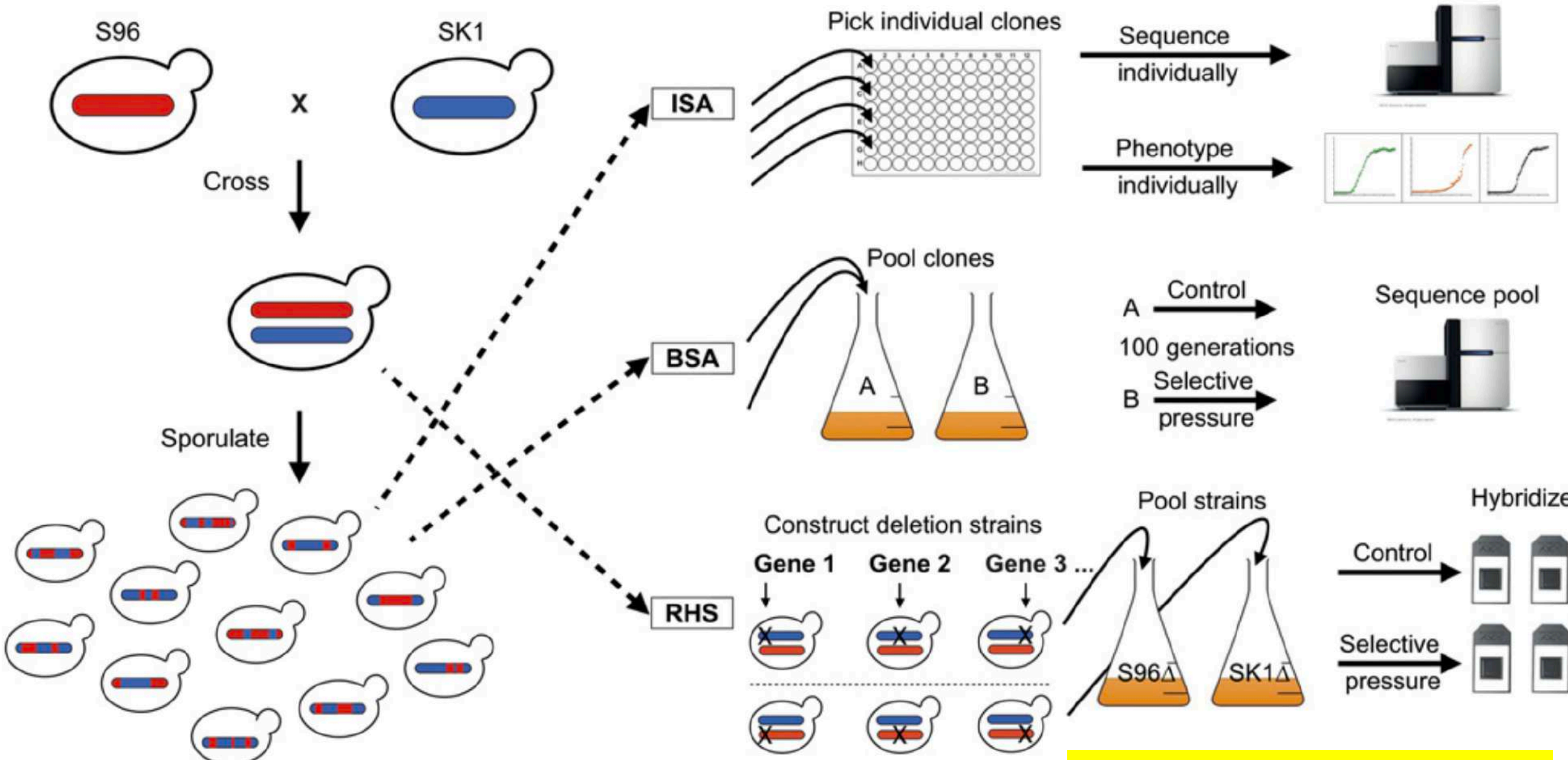


Manhattan plot



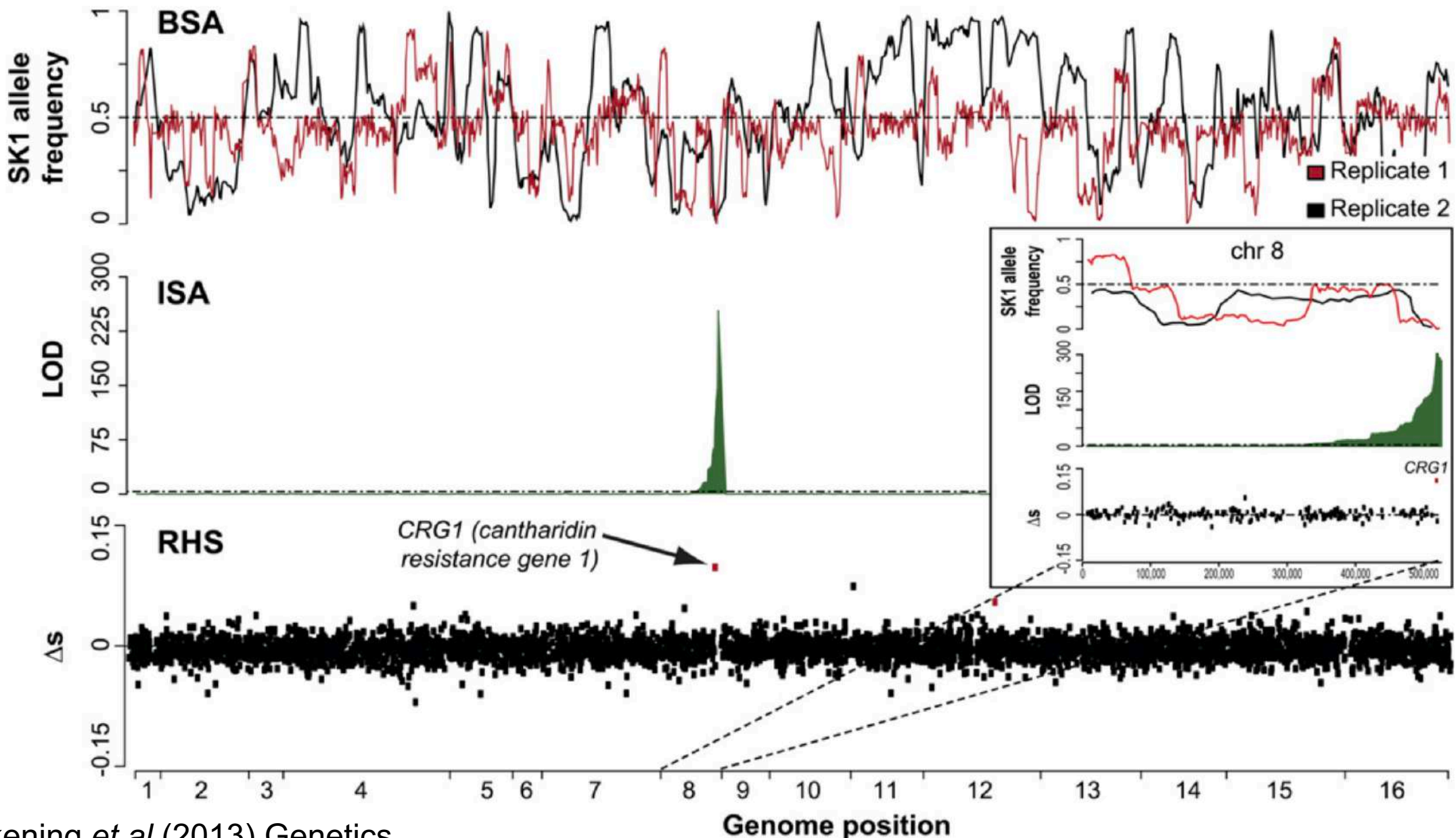
Need to control for possible confounding effects of genomic relatedness

Quantitative trait locus (QTL) mapping in yeast



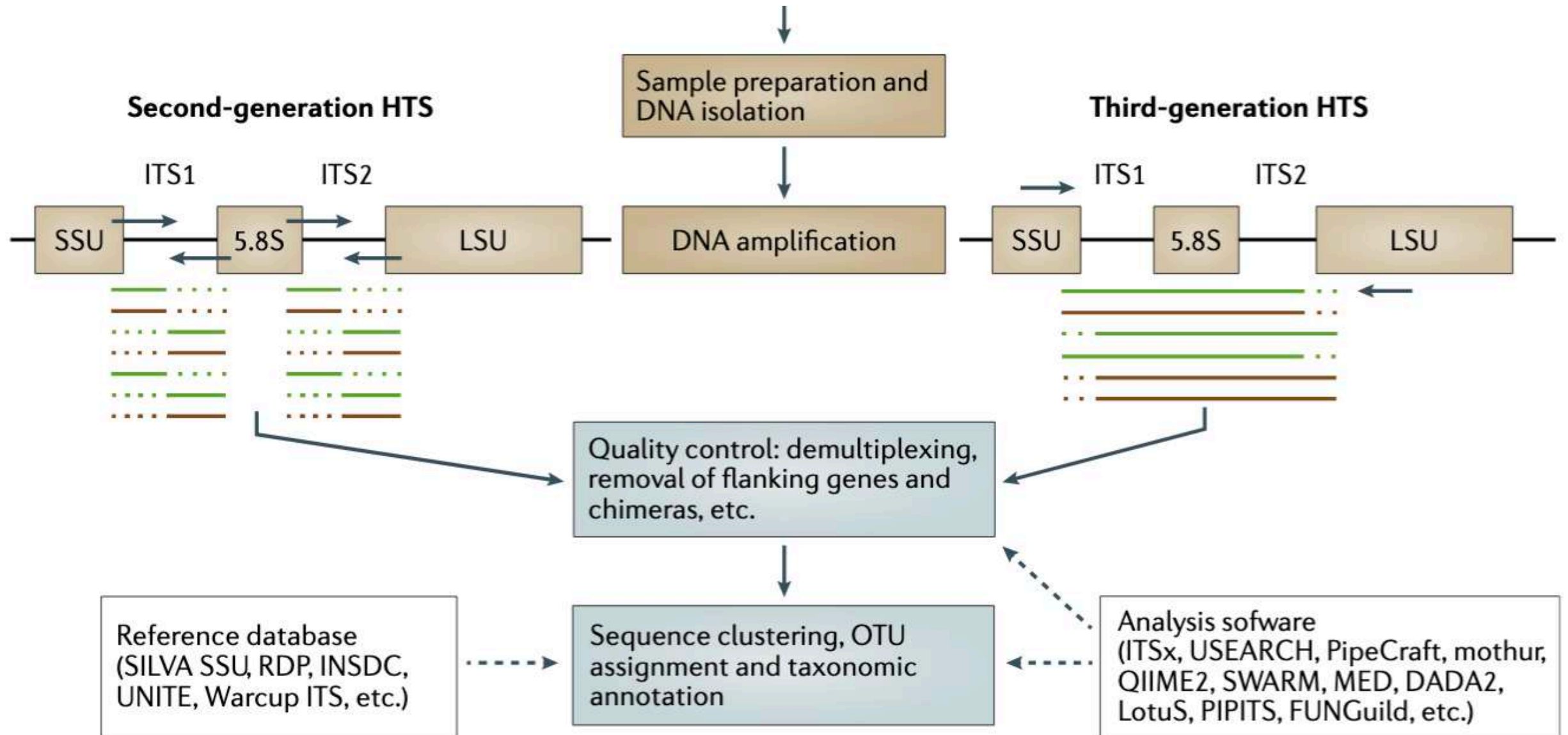
Reciprocal hemizyosity scanning

Quantitative trait locus (QTL) mapping in yeast

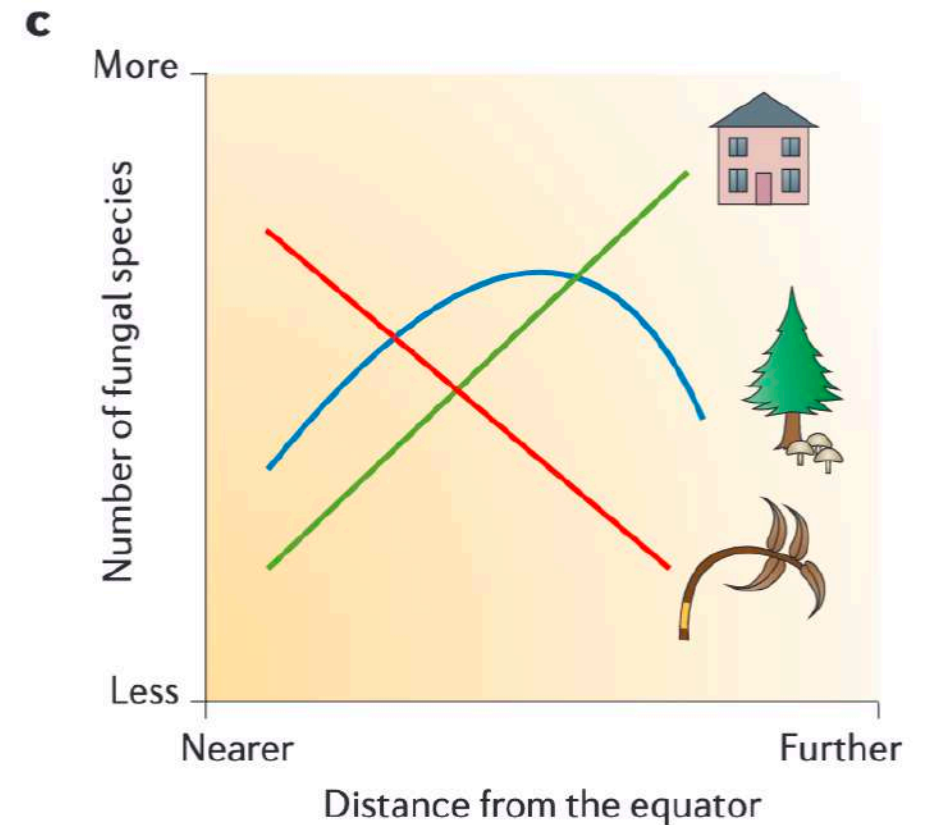
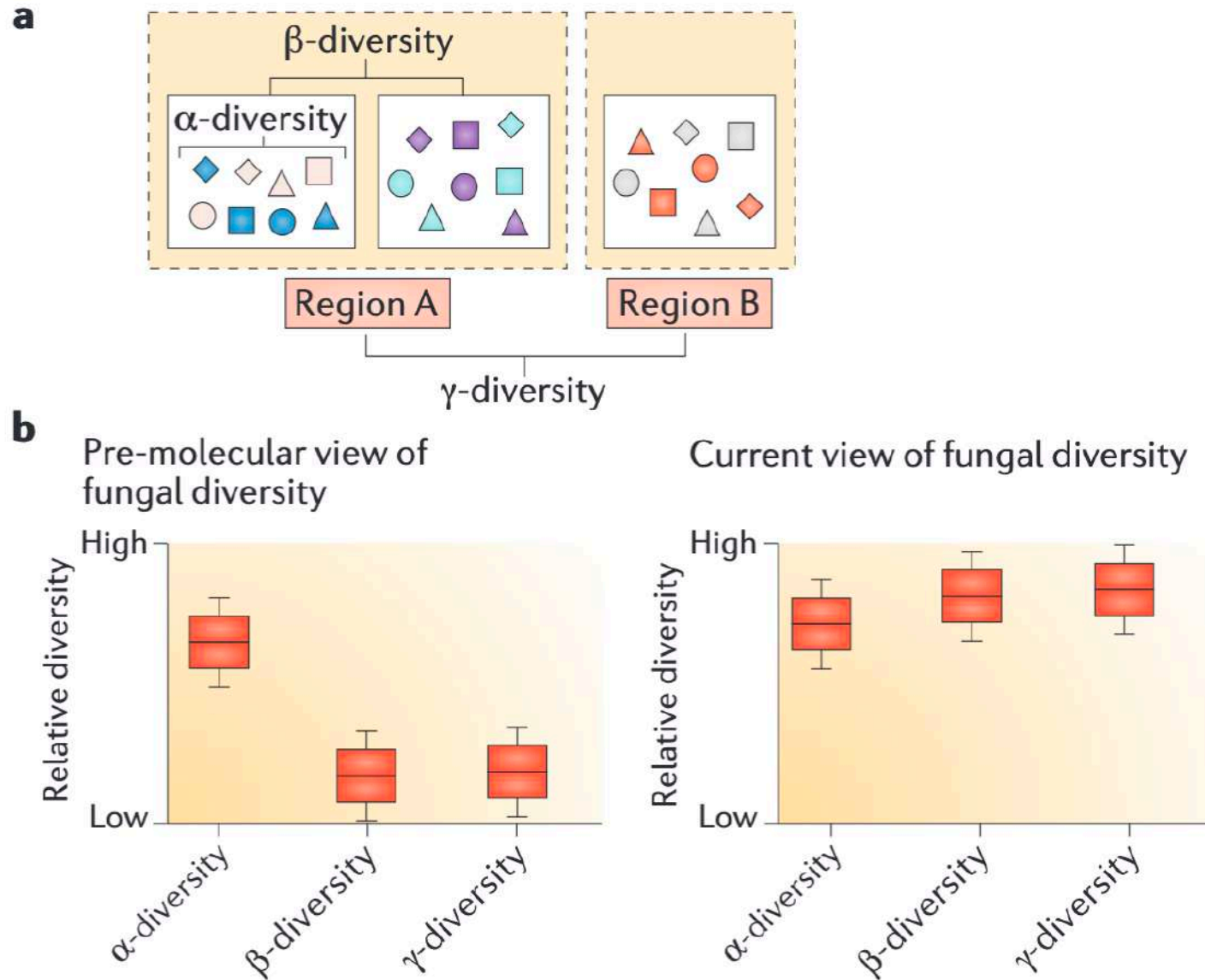


Community and microbial ecology
(already taught by previous lectures)

Main steps in a fungal metabarcoding project



Biogeography and emerging views of fungal diversity

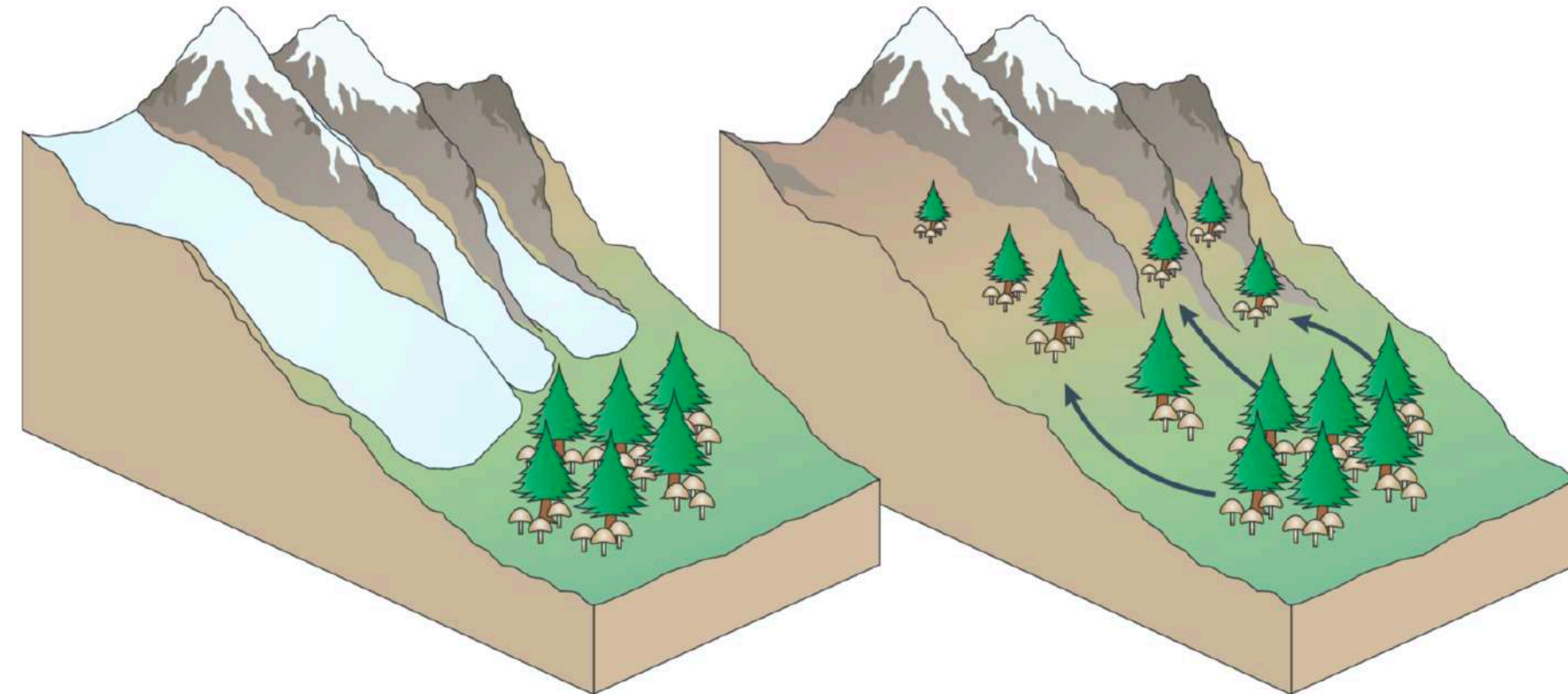


The evolution of new species or genetic diversity in fungi is often associated with dispersal or migration into new habitats

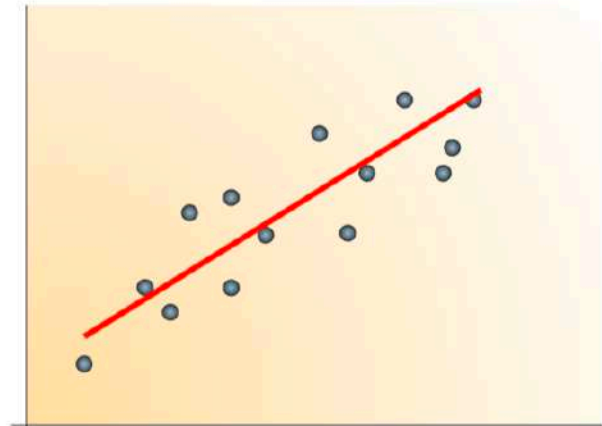
d

Glacial refugium

Post-glacial co-migration



Genetic differentiation

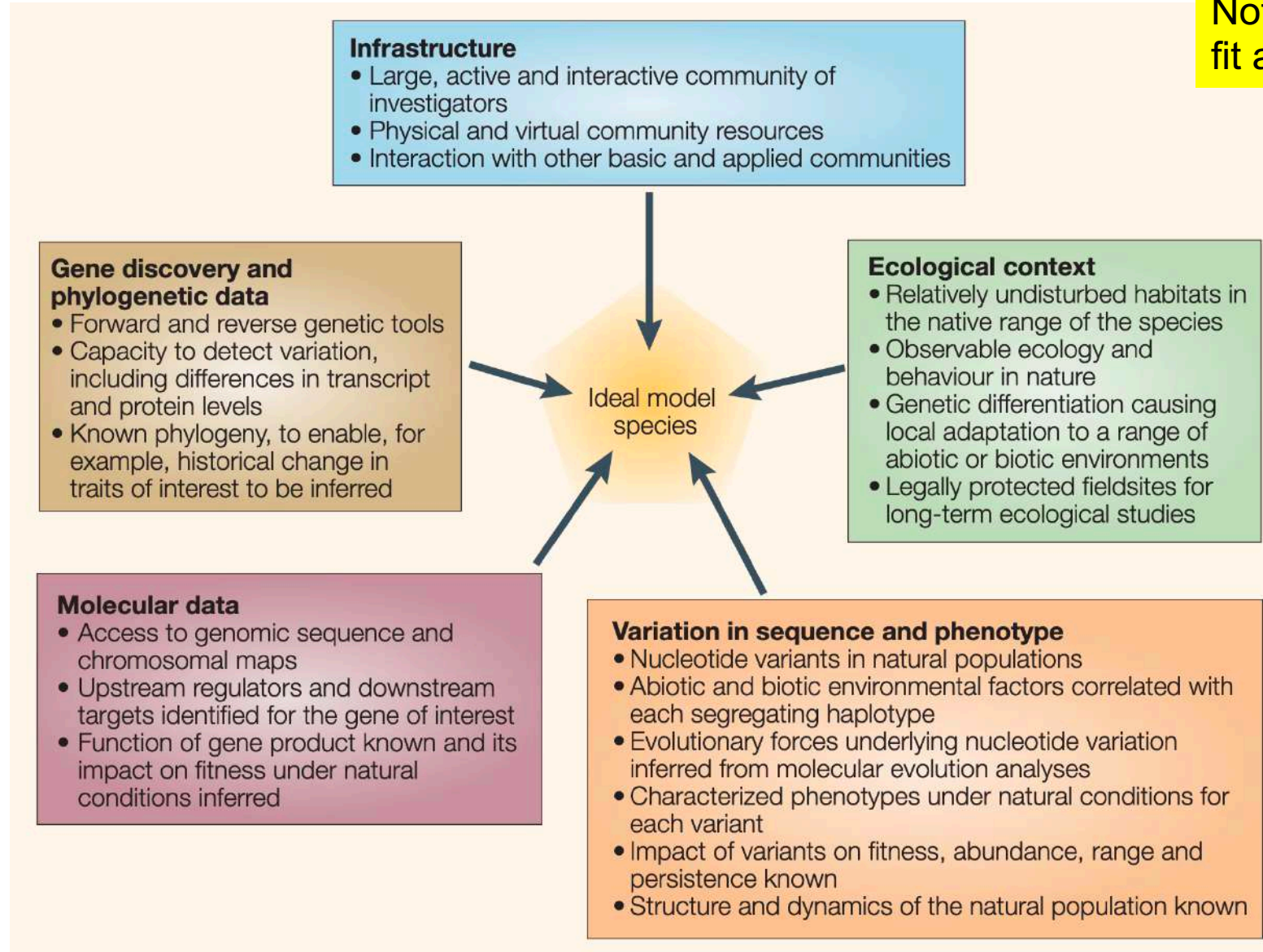


Distance between populations
(through a migration route)

Moving towards functions

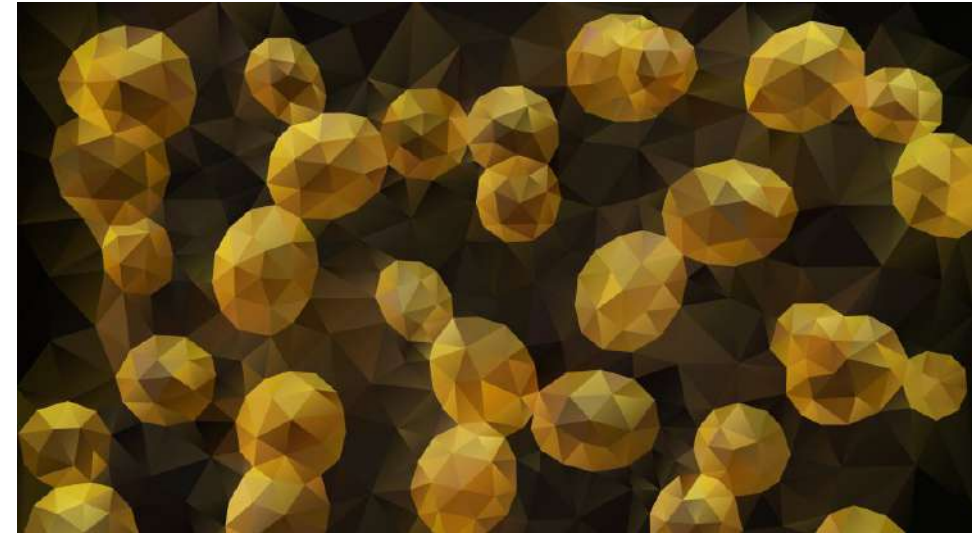
An ideal model organism for EG

Not many organisms fit all these criteria



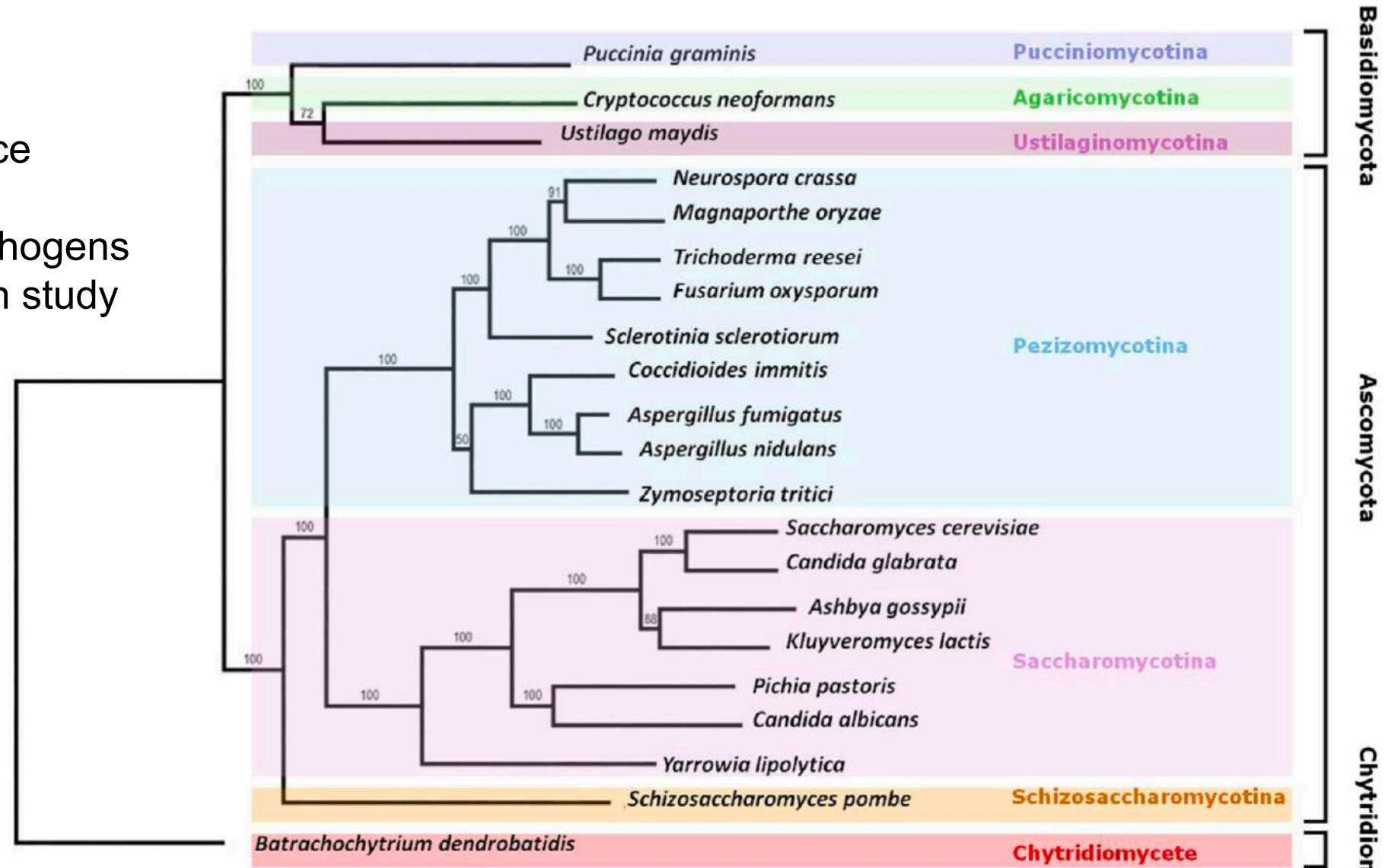
Model organisms

- Easy to maintain and breed in a laboratory setting.
- Many model organisms can breed in large numbers.
- Some have a very short generation time, which is the time between being born and being able to reproduce, so several generations can be followed at once
- Mutants allow scientists to study certain characteristics or diseases.
- Easy and cheap genetic manipulation
- Some model organisms have orthologs to humans.
- Model organisms can be used to create highly detailed genetic maps.
- Or they may occupy a pivotal position in the evolutionary tree



Research in other model fungi

- All have genome sequence available
- Most are animal/plant pathogens
- So much emphasis put on study of virulence



0,3

Research in model yeast *Saccharomyces cerevisiae*

Yeast: An Experimental Organism for 21st Century Biology

David Botstein*¹ and Gerald R. Fink[†]

*Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, and [†]Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Functional Genomics: Gene–Protein–Function Association via Mutants

Databases and Gene Ontology

Gene Expression and Regulatory Networks

Protein Interaction Networks

Gene Interaction Networks

Integrating Co-expression and Protein and Gene Interaction Networks

Leveraging Diversity to Understand Complex Inheritance

Strengths and Weaknesses of Genome-Scale Experimentation and Inference: Experimental Validation Is Essential

Evolution

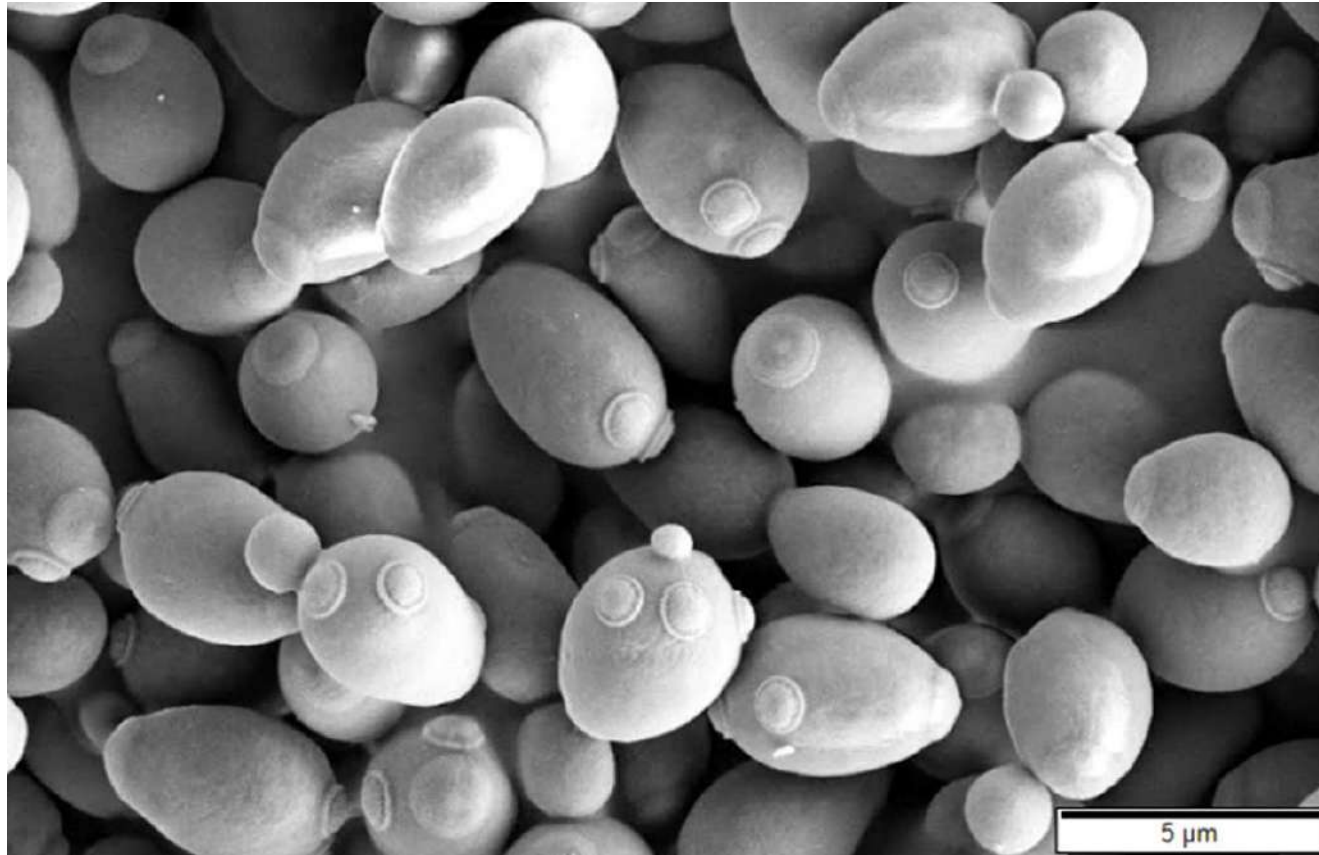
Evidence for the theory of duplication and divergence

Experimental evolution studies with yeast

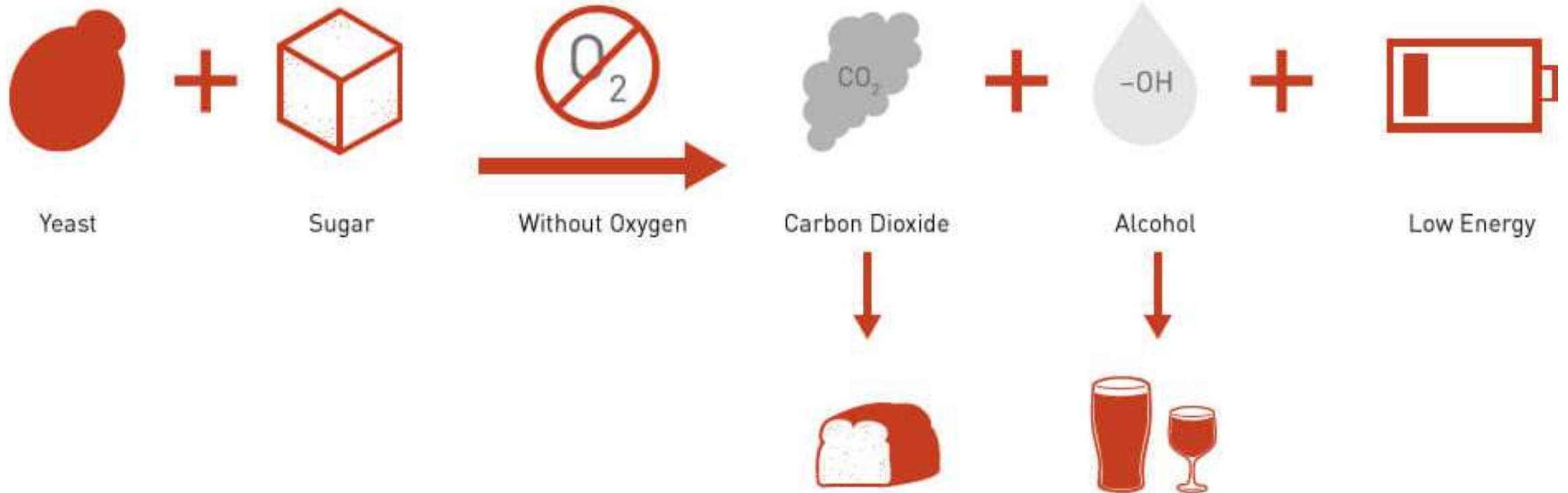
Human Disease

- Biotechnology
- Fermentation
- Synthetic biology
- High-throughput / Systematic
- Light sensing

Case studies: Origin of *Saccharomyces cerevisiae*



S. cerevisiae arguably one of the most domesticated species for its fermentation product



Revealing a 5,000-y-old beer recipe in China

Jiajing Wang^{a,b,1}, Li Liu^{a,b}, Terry Ball^c, Linjie Yu^d, Yuanqing Li^e, and Fulai Xing^f

^aStanford Archaeology Center, Stanford University, Stanford, CA 94305; ^bDepartment of East Asian Languages and Cultures, Stanford University, Stanford, CA 94305; ^cDepartment of Ancient Scripture, Brigham Young University, Provo, UT 84602; ^dZhejiang Research Institute of Chemical Industry, 310006 Hangzhou, China; ^eDepartment of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305; and ^fShaanxi Provincial Institute of Archaeology, 710054 Xi'an, China

Edited by Dolores R. Piperno, Smithsonian Institution, Fairfax, VA, and approved April 26, 2016 (received for review January 27, 2016)



'World's oldest brewery' found in cave in Israel, say researchers



True ecology of *Saccharomyces cerevisiae*?



FEMS Yeast Research, 15, 2015, fov009

doi: [10.1093/femsyr/fov009](https://doi.org/10.1093/femsyr/fov009)

Advance Access Publication Date: 27 February 2015

Commentary



COMMENTARY

Saccharomyces cerevisiae: a nomadic yeast with no niche?

Matthew R. Goddard^{1,2,*} and Duncan Greig^{3,4}

Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution

Irene Stefanini^{a,1}, Leonardo Dapporto^{b,c,1}, Jean-Luc Legras^{d,e,f}, Antonio Calabretta^{a,b}, Monica Di Paola^g, Carlotta De Filippo^h, Roberto Viola^h, Paolo Capretti^c, Mario Polsinelli^b, Stefano Turillazzi^{b,i}, and Duccio Cavalieri^{a,h,2}

^aDipartimento di Farmacologia, University of Florence, 50139, Florence, Italy; ^bDipartimento di Biologia Evoluzionistica, University of Florence, 50125, Florence, Italy; ^cDipartimento di Biotecnologie Agrarie, University of Florence, 50144, Florence, Italy; ^dINRA (Institut National de la Recherche Agronomique), UMR1083 (Unité Mixte de Recherche Sciences pour l'Oenologie), F-34060 Montpellier, France; ^eMontpellier SupAgro, UMR1083 (Unité Mixte de Recherche Sciences pour l'Oenologie), F-34060 Montpellier, France; ^fUniversité Montpellier I, UMR1083 (Unité Mixte de Recherche Sciences pour l'Oenologie), F-34060 Montpellier, France; ^gDipartimento di Scienze per la Salute della Donna e del Bambino, Ospedale Pediatrico Meyer, University of Florence, 50139, Florence, Italy; ^hCentre for Research and Innovation, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all'Adige, Trento, Italy; and ⁱCentro di Servizi di Spettrometria di Massa, University of Florence, Florence, Italy

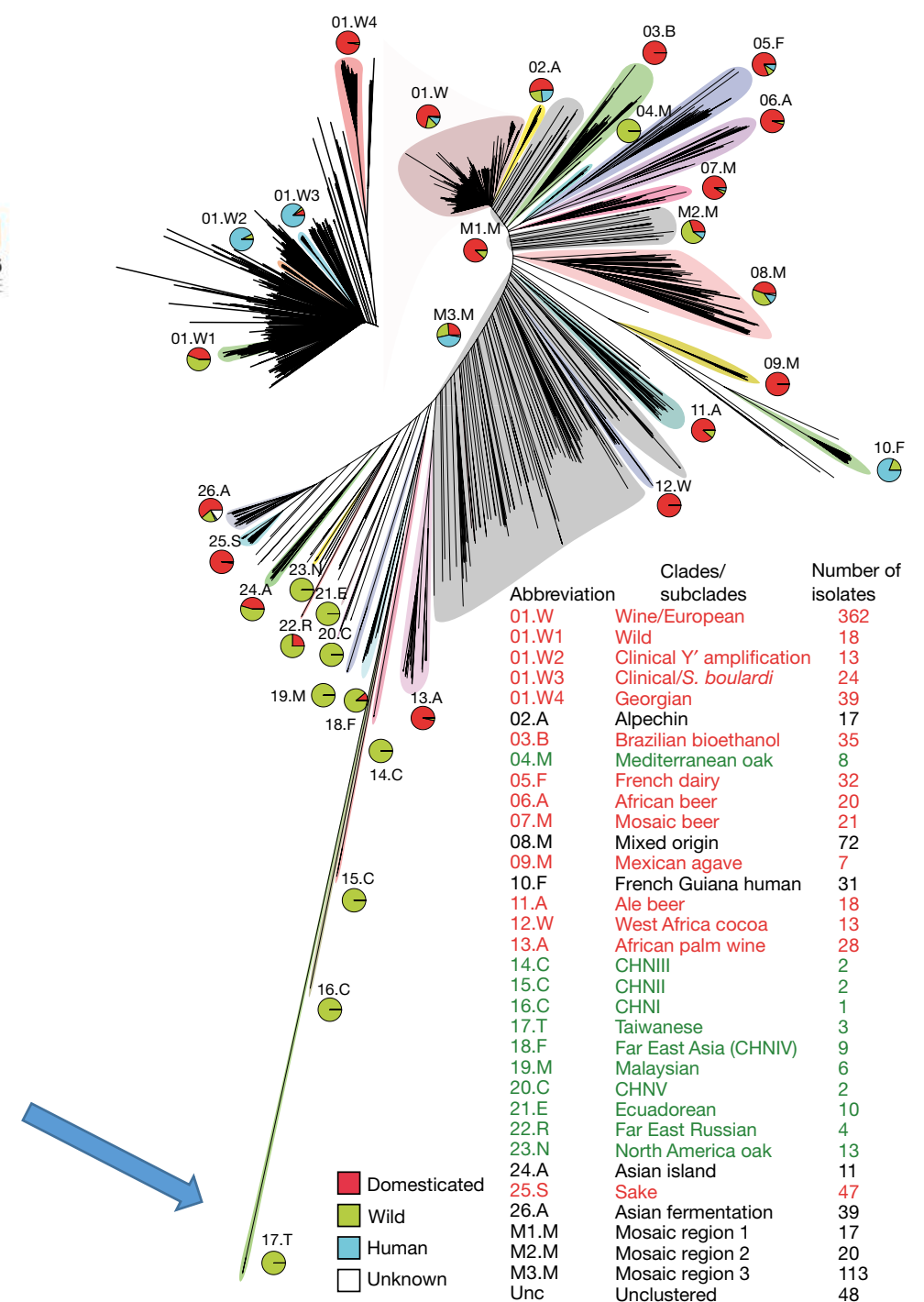
Edited by Nancy A. Moran, Yale University, West Haven, CT, and approved July 5, 2012 (received for review May 18, 2012)




Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates

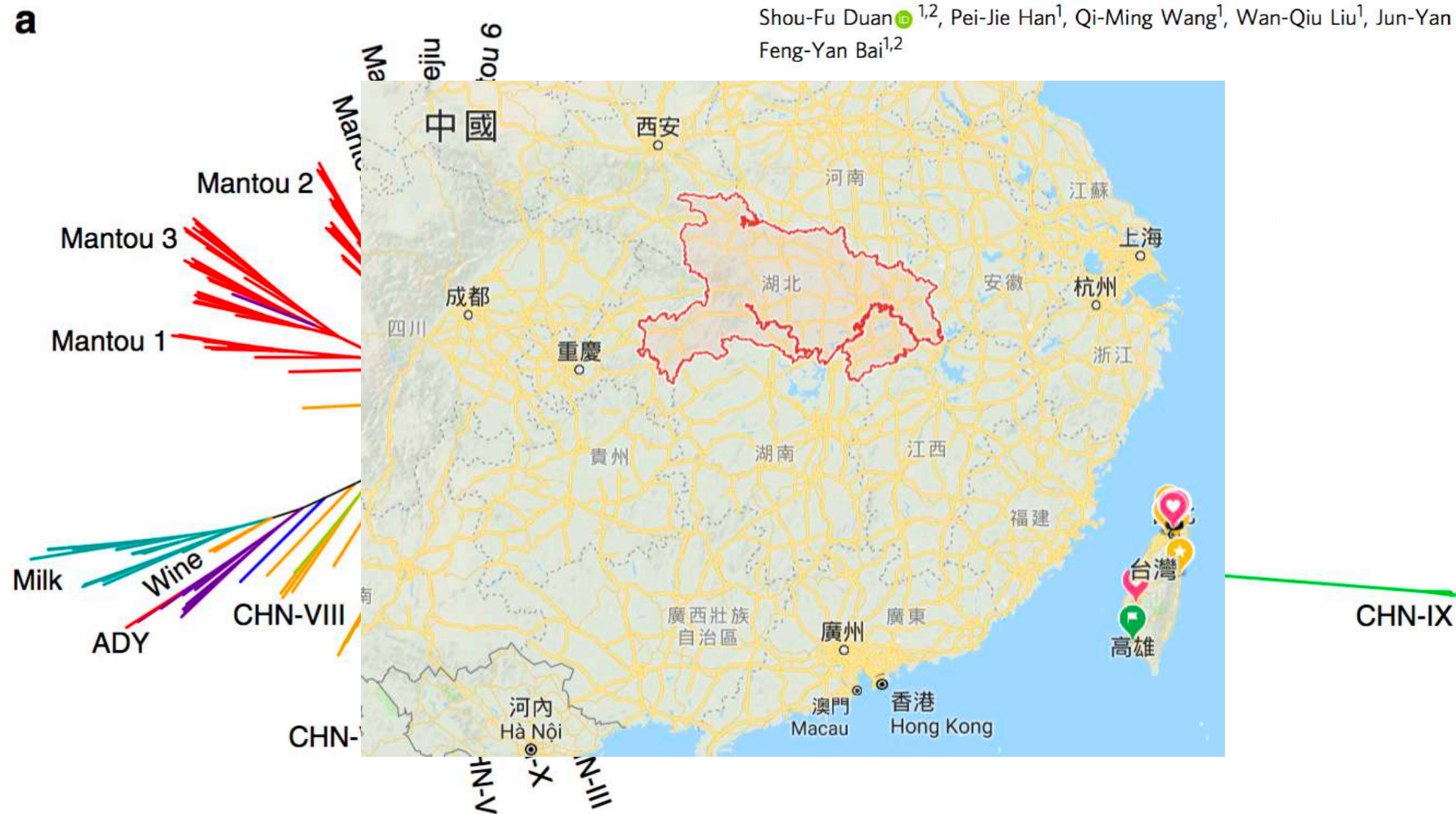
Jackson Peter^{1,6}, Matteo De Chiara^{2,6}, Anne Friedrich¹, Jia-Xing Yue², David Pflieger¹, Anders Bergström², Anastasie Sigwalt¹, Benjamin Barre², Kelle Freel¹, Agnès Llored², Corinne Cruaud³, Karine Labadie³, Jean-Marc Aury³, Benjamin Istace³, Kevin Lebrigand⁴, Pascal Barbry⁴, Stefan Engelen³, Arnaud Lemainque³, Patrick Wincker^{3,5,7}, Gianni Liti^{2,7*} & Joseph Schacherer^{1,7*}

“The Taiwanese wild lineage represents the most divergent population that has yet been described (average of 1.1% sequence divergence to non-Taiwanese strains).”



The origin and adaptive evolution of domesticated populations of yeast from Far East Asia

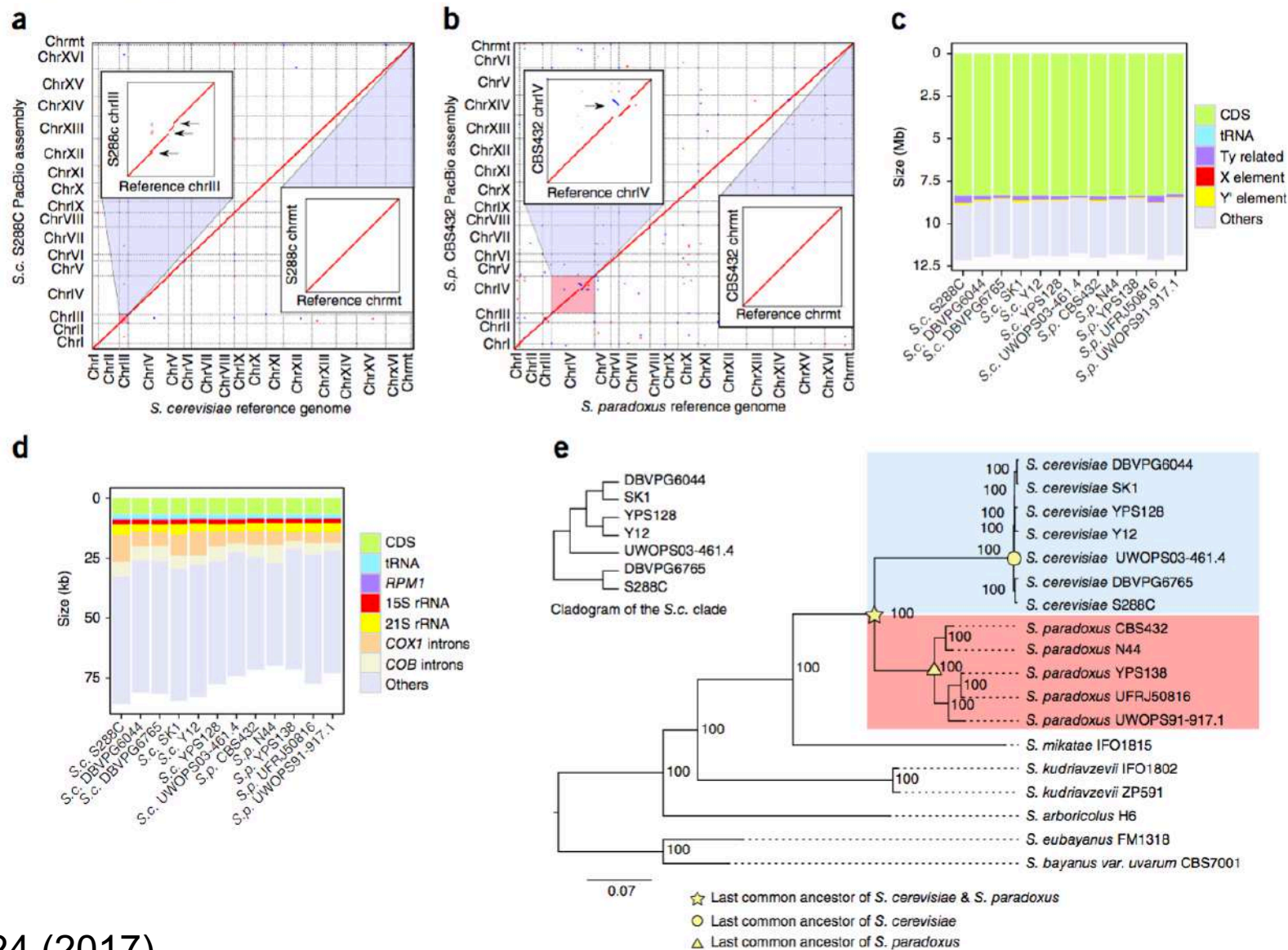
Shou-Fu Duan ^{1,2}, Pei-Jie Han¹, Qi-Ming Wang¹, Wan-Qiu Liu¹, Jun-Yan Shi^{1,2}, Kuan Li¹, Xiao-Ling Zhang¹ & Feng-Yan Bai^{1,2}



Contrasting evolutionary genome dynamics between domesticated and wild yeasts

Jia-Xing Yue¹, Jing Li¹, Louise Aigrain², Johan Hallin¹, Karl Persson³, Karen Oliver², Anders Bergström², Paul Coupland^{2,5}, Jonas Warringer³, Marco Cosentino Lagomarsino⁴, Gilles Fischer⁴, Richard Durbin² & Gianni Liti¹

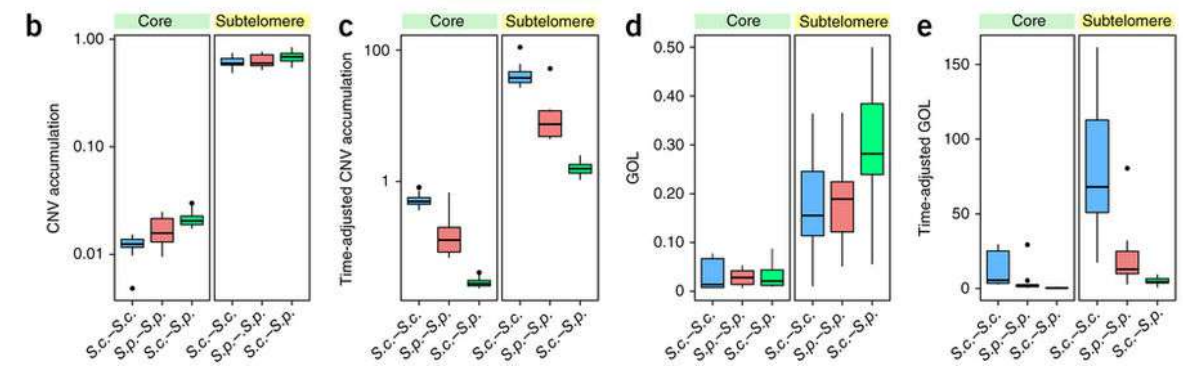
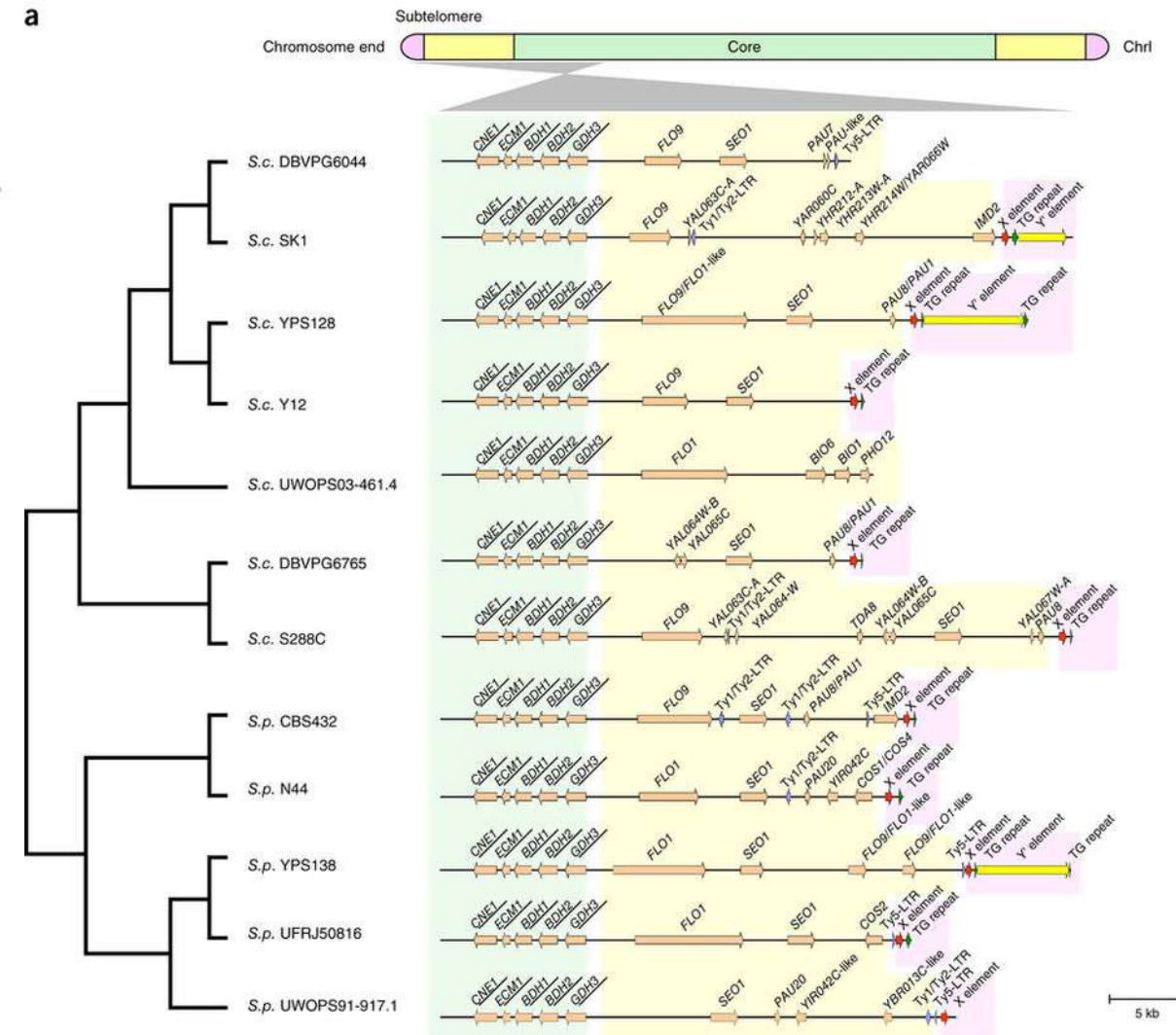
- long-read sequencing to generate **end-to-end genome assemblies** for **12 strains** representing major subpopulations of the partially domesticated yeast *Saccharomyces cerevisiae* and its wild relative *S. paradoxus*.



Contrasting evolutionary genome dynamics between domesticated and wild yeasts

Jia-Xing Yue¹, Jing Li¹, Louise Aigrain², Johan Hallin¹, Karl Persson³, Karen Oliver², Anders Bergström², Paul Coupland^{2,5}, Jonas Warringer³, Marco Cosentino Lagomarsino⁴, Gilles Fischer⁴, Richard Durbin² & Gianni Liti¹

- enable precise definition of chromosomal boundaries between cores and subtelomeres
- *S. paradoxus* shows faster accumulation of balanced rearrangements (inversions, reciprocal translocations and transpositions), *S. cerevisiae* accumulates unbalanced rearrangements (novel insertions, deletions and duplications) more rapidly.
- Such striking contrasts between wild and domesticated yeasts are likely to reflect the influence of human activities on structural genome evolution.



Case study I

The ISME Journal (2017) 11, 1189–1204

© 2017 International Society for Microbial Ecology All rights reserved 1751-7362/17

www.nature.com/ismej

ORIGINAL ARTICLE

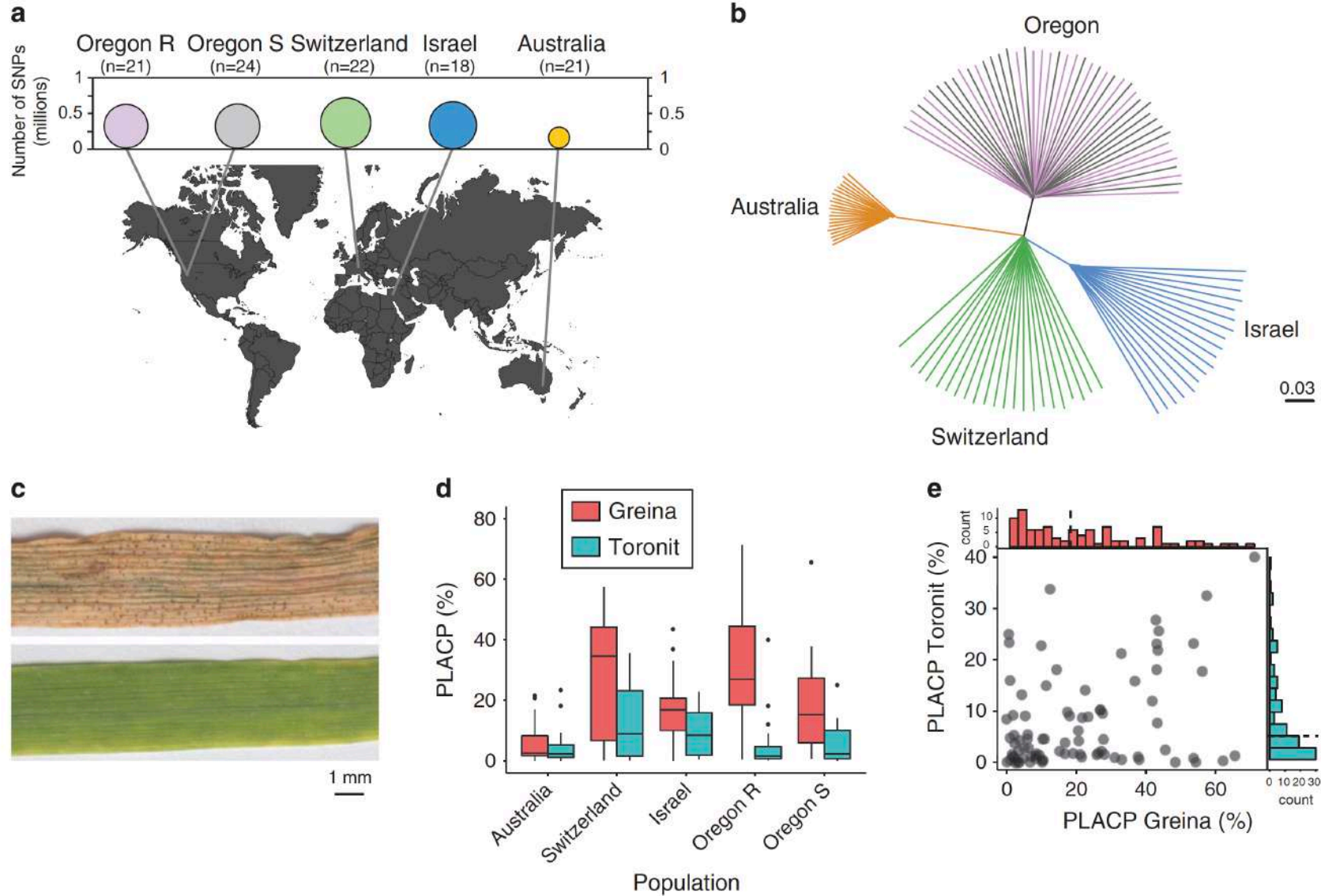
A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements

Fanny E Hartmann, Andrea Sánchez-Vallet, Bruce A McDonald and Daniel Croll¹
Plant Pathology, Institute of Integrative Biology, Zurich, Switzerland

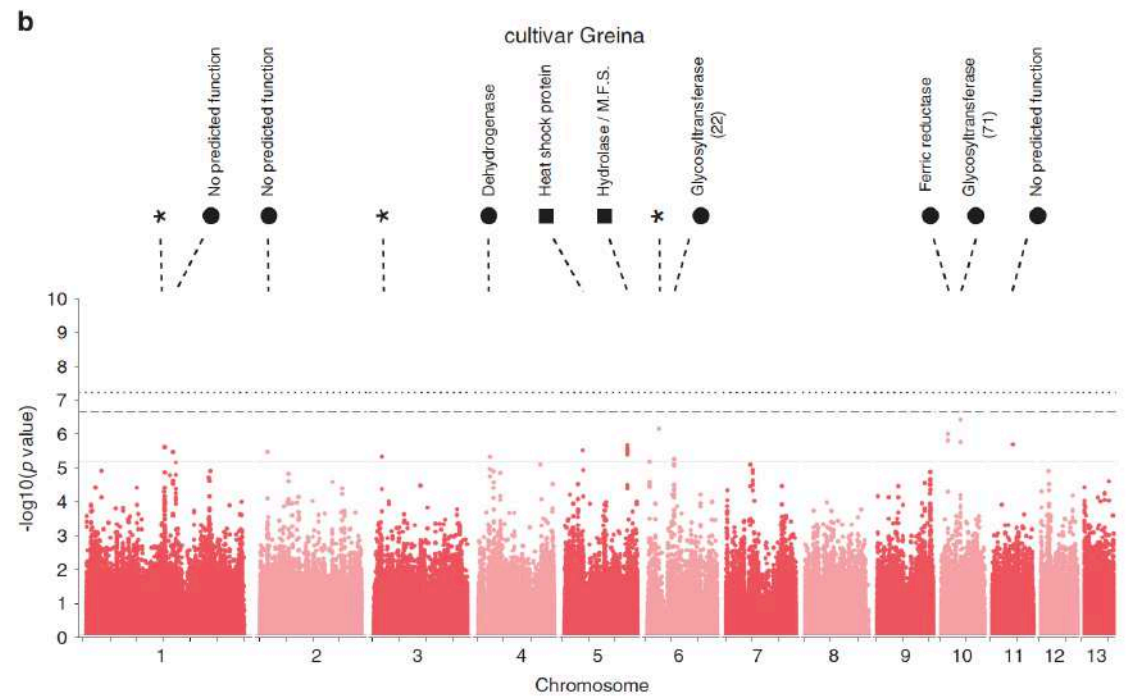
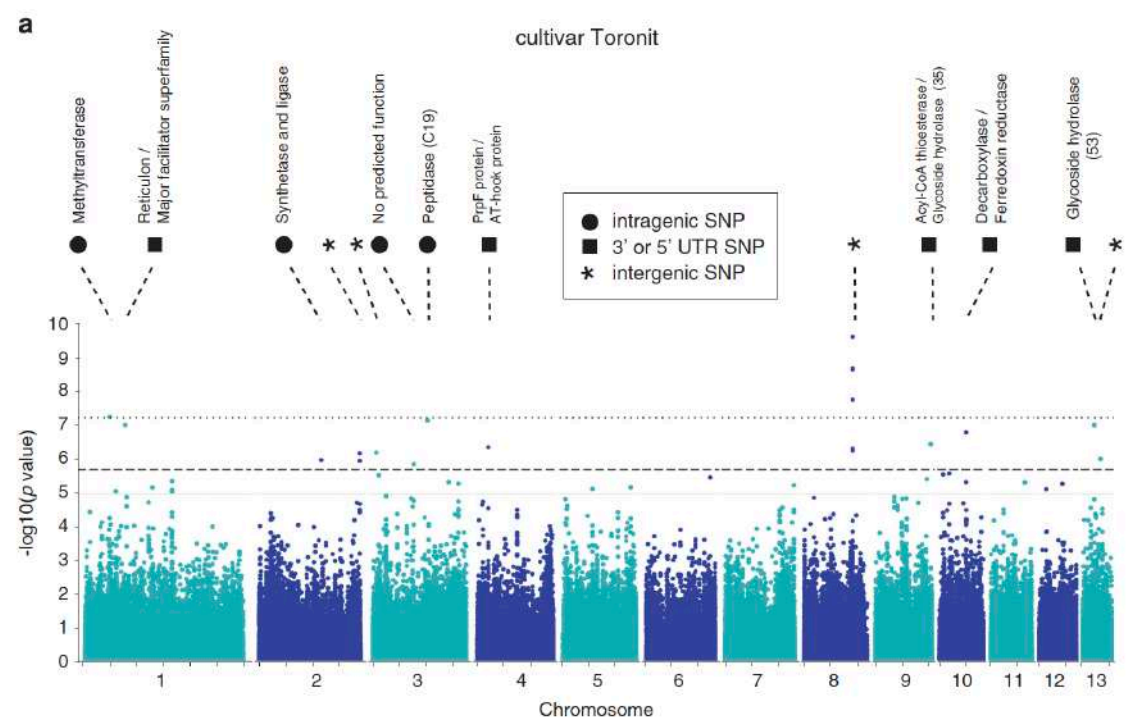
Case study

Fungal pathogens can rapidly evolve virulence towards resistant crops in agricultural ecosystems. Gains in virulence are often mediated by the mutation or deletion of a gene encoding a protein recognized by the plant immune system. However, the loci and the mechanisms of genome evolution enabling rapid virulence evolution are poorly understood. We performed genome-wide association mapping on a global collection of 106 strains of *Zymoseptoria tritici*, the most damaging pathogen of wheat in Europe, to identify polymorphisms linked to virulence on two wheat varieties. We found 25 distinct genomic loci associated with reproductive success of the pathogen. However, no locus was shared between the host genotypes, suggesting host specialization. **The main locus associated with virulence encoded a highly expressed, small secreted protein.** Population genomic analyses showed that the gain in virulence was explained by a segregating gene deletion polymorphism. The deletion was likely adaptive by preventing detection of the encoded protein. **Comparative genomics** of closely related species showed that the locus emerged *de novo* since speciation. A large cluster of transposable elements in direct proximity to the locus generated extensive rearrangements leading to multiple independent gene losses. Our study demonstrates that rapid turnover in the chromosomal structure of a pathogen can drive host specialization.

Genetic polymorphism, population structure, and virulence phenotypes of 106 *Zymoseptoria tritici* isolates

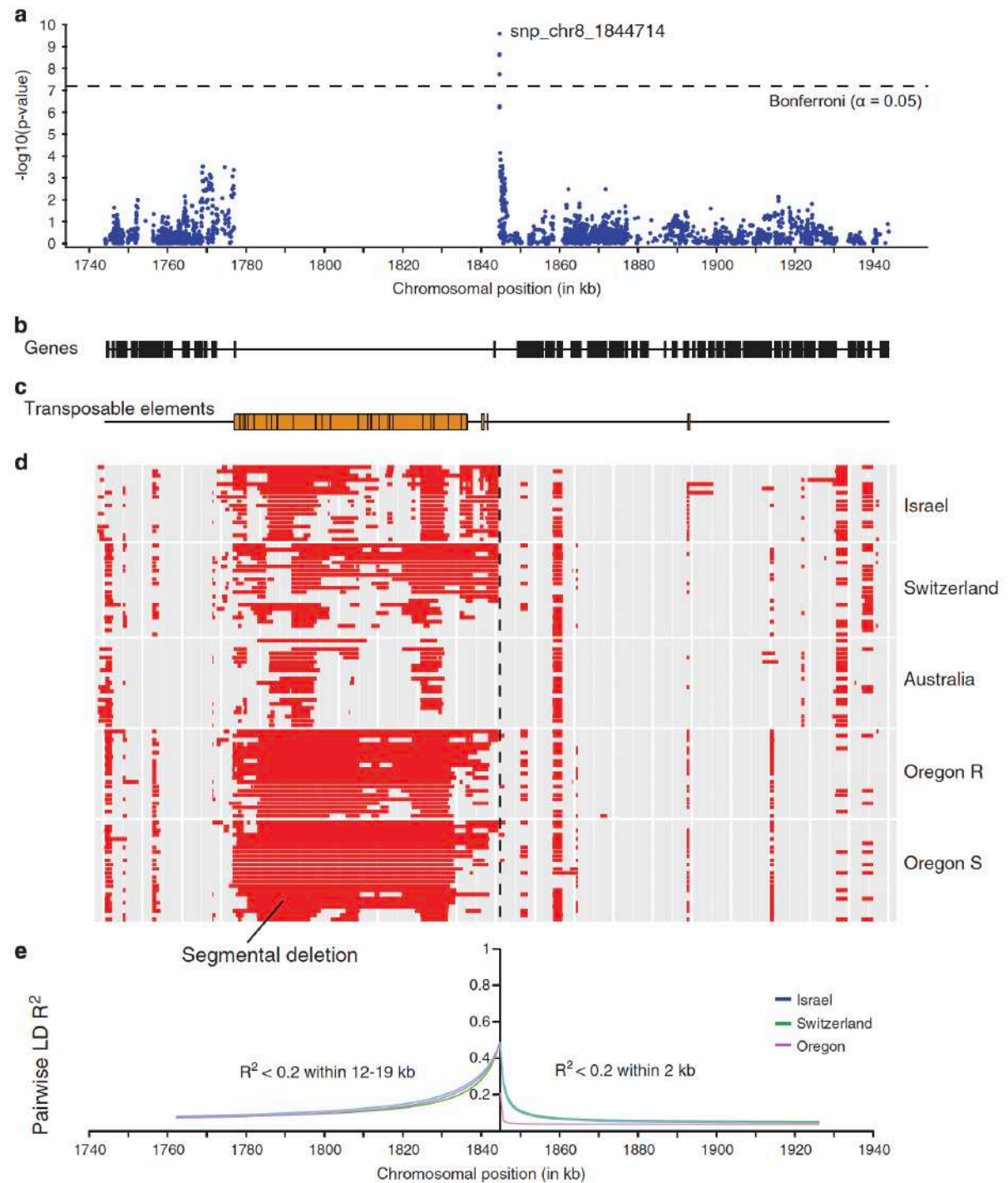


GWAS

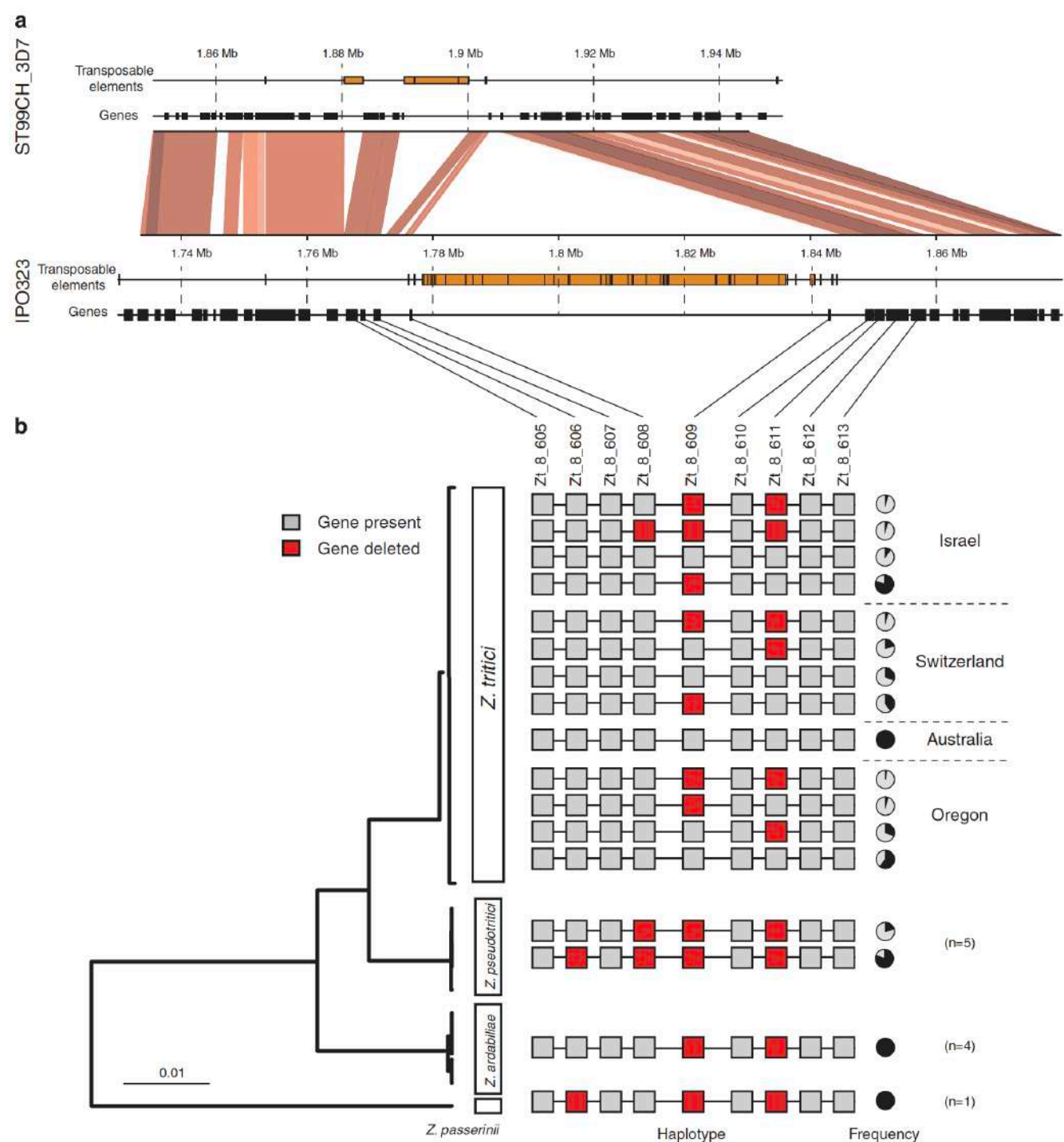


No loci were shared between two hosts

Population genomic analyses for the most significant GWAS locus associated with virulence of *Z. tritici*



The evolutionary history and structural variation at the major virulence locus detected by GWAS.



Case study II

RESEARCH ARTICLE

Genome-wide association across *Saccharomyces cerevisiae* strains reveals substantial variation in underlying gene requirements for toxin tolerance

Maria Sardi^{1,2^{⌘a}}, Vaishnavi Paithane¹, Michael Place¹, De Elegant Robinson², James Hose³, Dana J. Wohlbach^{1^{⌘b}}, Audrey P. Gasch^{1,3*}

1 Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, **2** Microbiology Training Program, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, **3** Laboratory of Genetics, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

^{⌘a} Current address: Cargill, Minneapolis, Minnesota, United States of America

^{⌘b} Current address: Department of Biology, Dickinson College, Carlisle, Pennsylvania, United States of America

* agasch@wisc.edu

Complex trait - hydrolysate-toxin (HT) tolerance

“One significant hurdle with regards to microbial fermentation is the presence of toxic compounds in the processed plant material, or hydrolysate, including weak acids, furans and phenolics released or generated by the pretreatment process.”

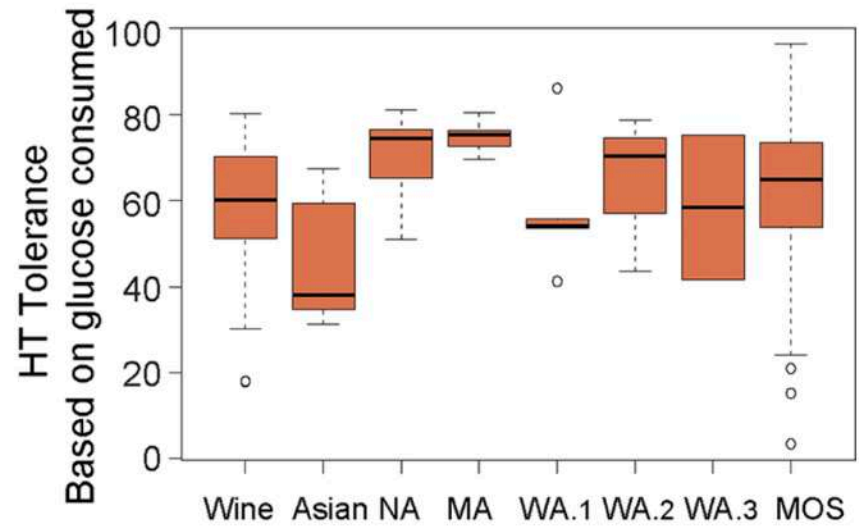
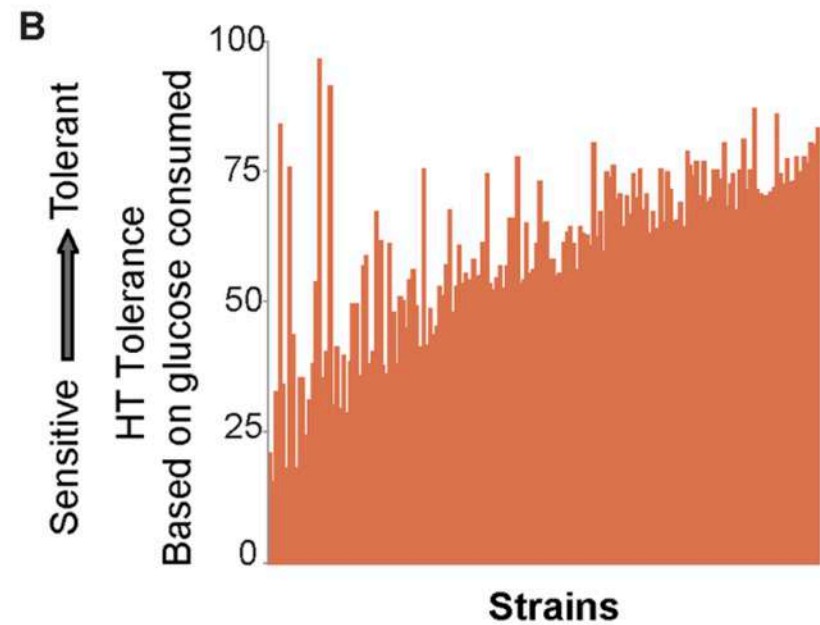
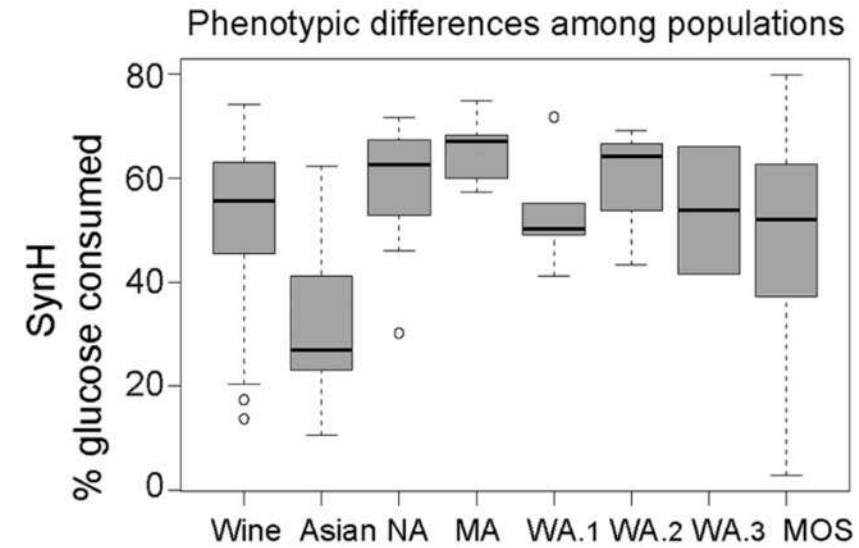
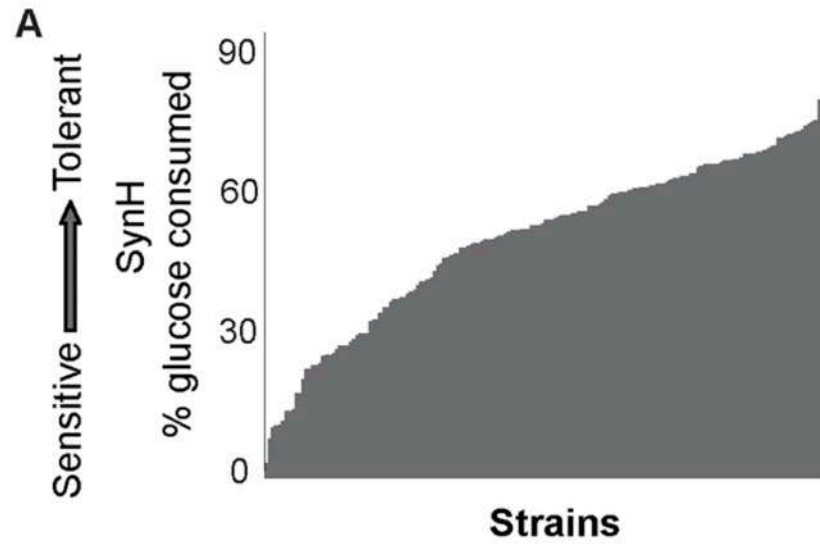
“We used genome-wide association (GWA) in *S. cerevisiae* strains responding to synthetic hydrolysate (SynH), both to identify new genes and processes important for HT tolerance and to explore the extent to which genetic background influences mechanism.

We tested 20 genes associated with HT tolerance and swapped alleles across strains to validate several allele-specific effects. However, in the process of allele exchange we discovered striking differences in gene contributions to the phenotype: out of 14 gene knockouts tested in two strains with opposing phenotypes, 8 (57%) had a statistically significant effect on HT tolerance in one of the backgrounds but little to no significant effect in the other background.

In most of these cases, the specific allele had little observable contribution to the phenotype. Thus, although GWA successfully implicated new genes and processes involved in HT tolerance, the causal variation in the tested strains is not at the level of the allele but rather whether or not the gene’s function is important for the phenotype in that background.

This raises important implications for considering natural variation in functional networks to explain phenotypic variation”

Strain-specific difference for SynH and HT tolerance



Distribution of SNP alleles

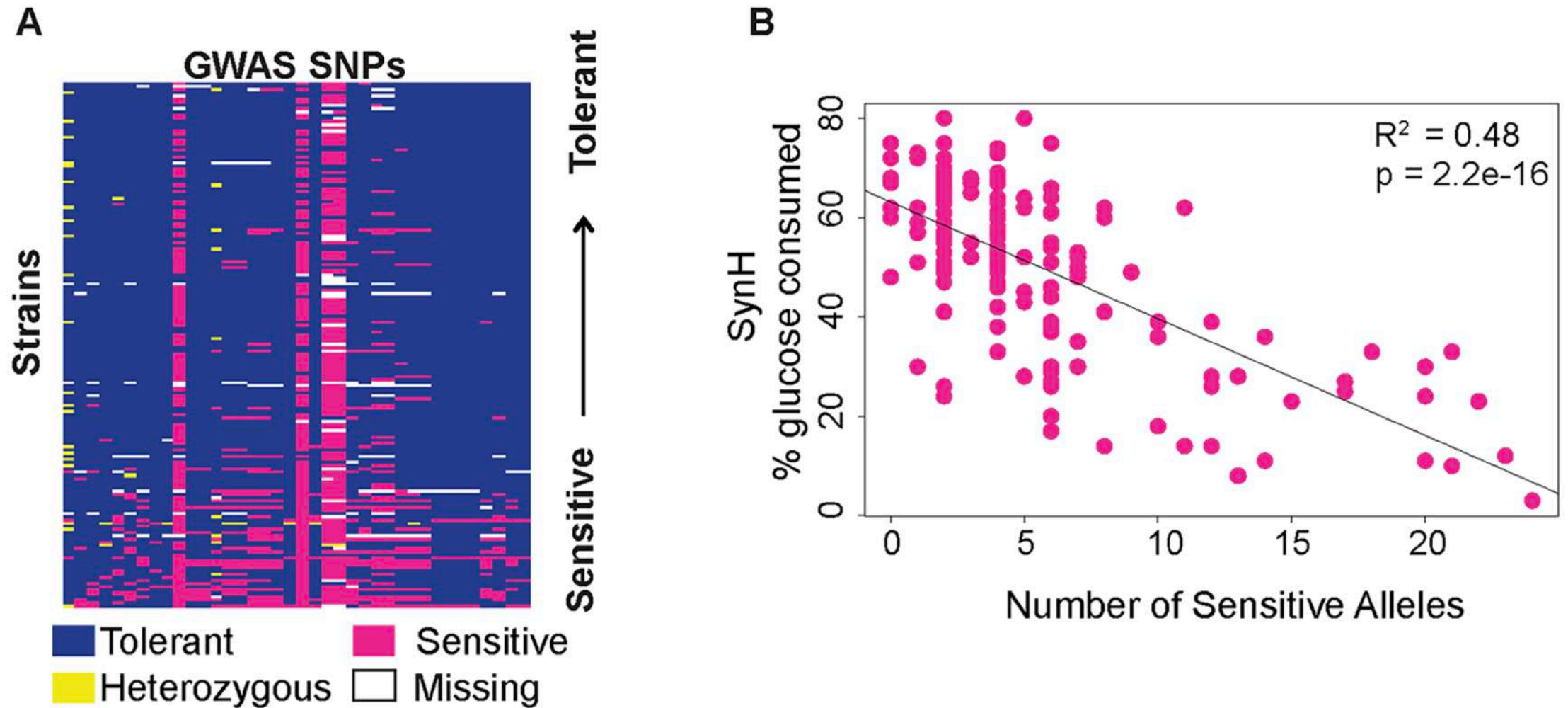
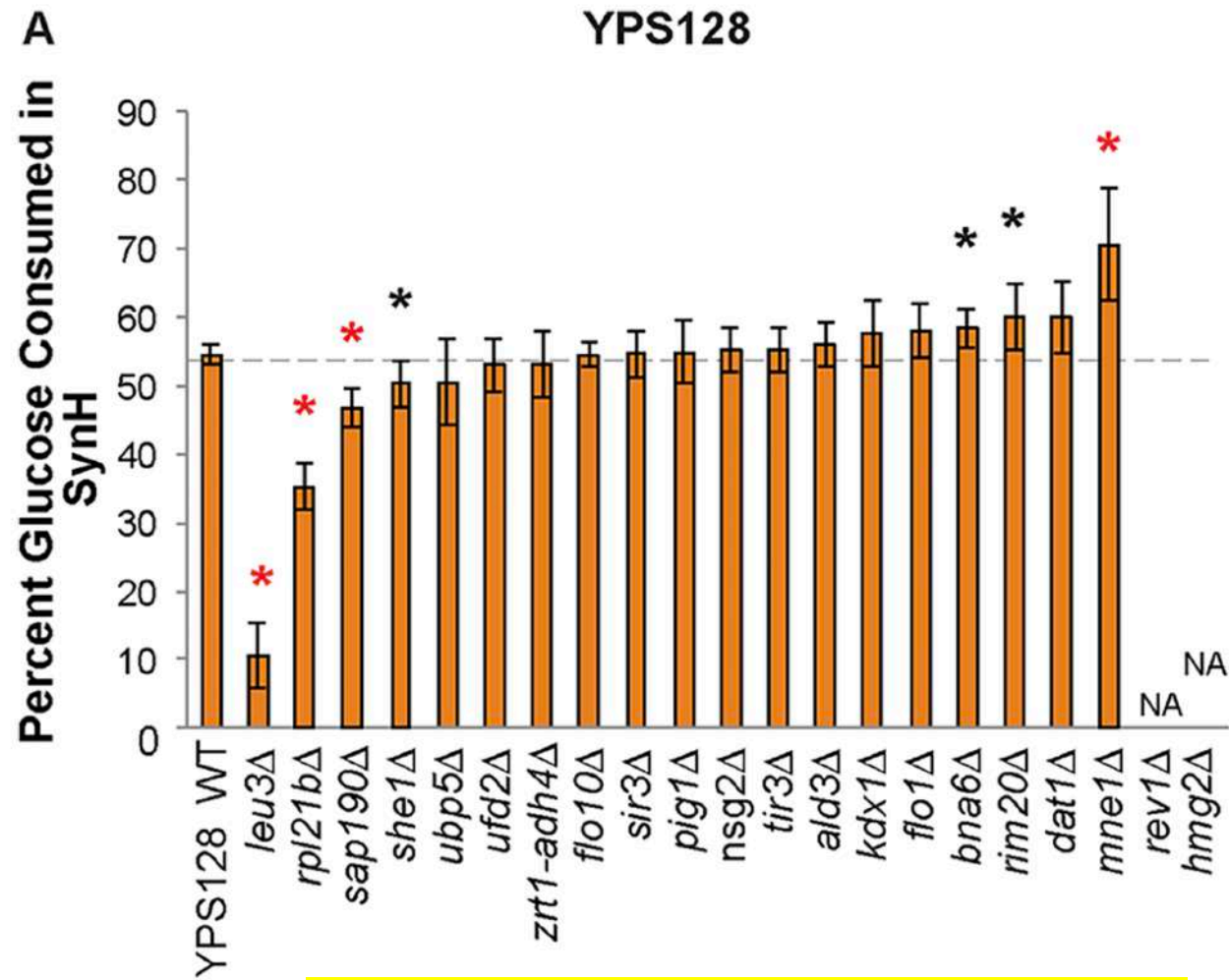
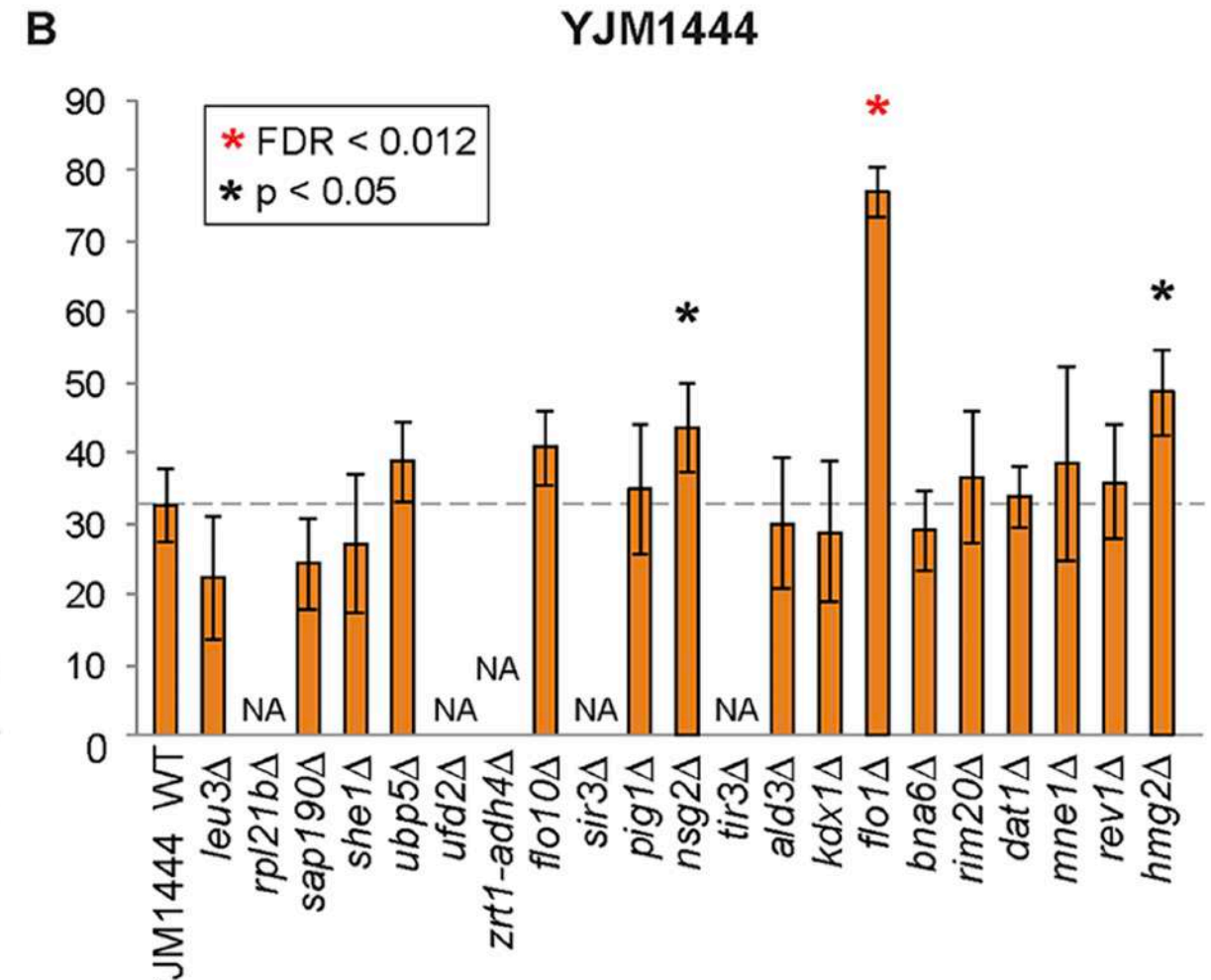


Fig 3. Distribution of SNP alleles. (A) A heat map of the 38 SNPs found in the GWA analysis (columns) in each strain (rows), where the alleles associated with the sensitive or resistance phenotypes are color-coded according to the key. Strains were organized from tolerant (top) to sensitive (bottom). (B) Percent glucose consumed in SynH + HTs was plotted against the number of sensitive alleles identified in each strain. Correlation of the two is indicated by the R^2 and linear fit line.

Knockout effects of genes containing SNPs found in GWA



Extensive background effects influence gene involvement in SynH tolerance



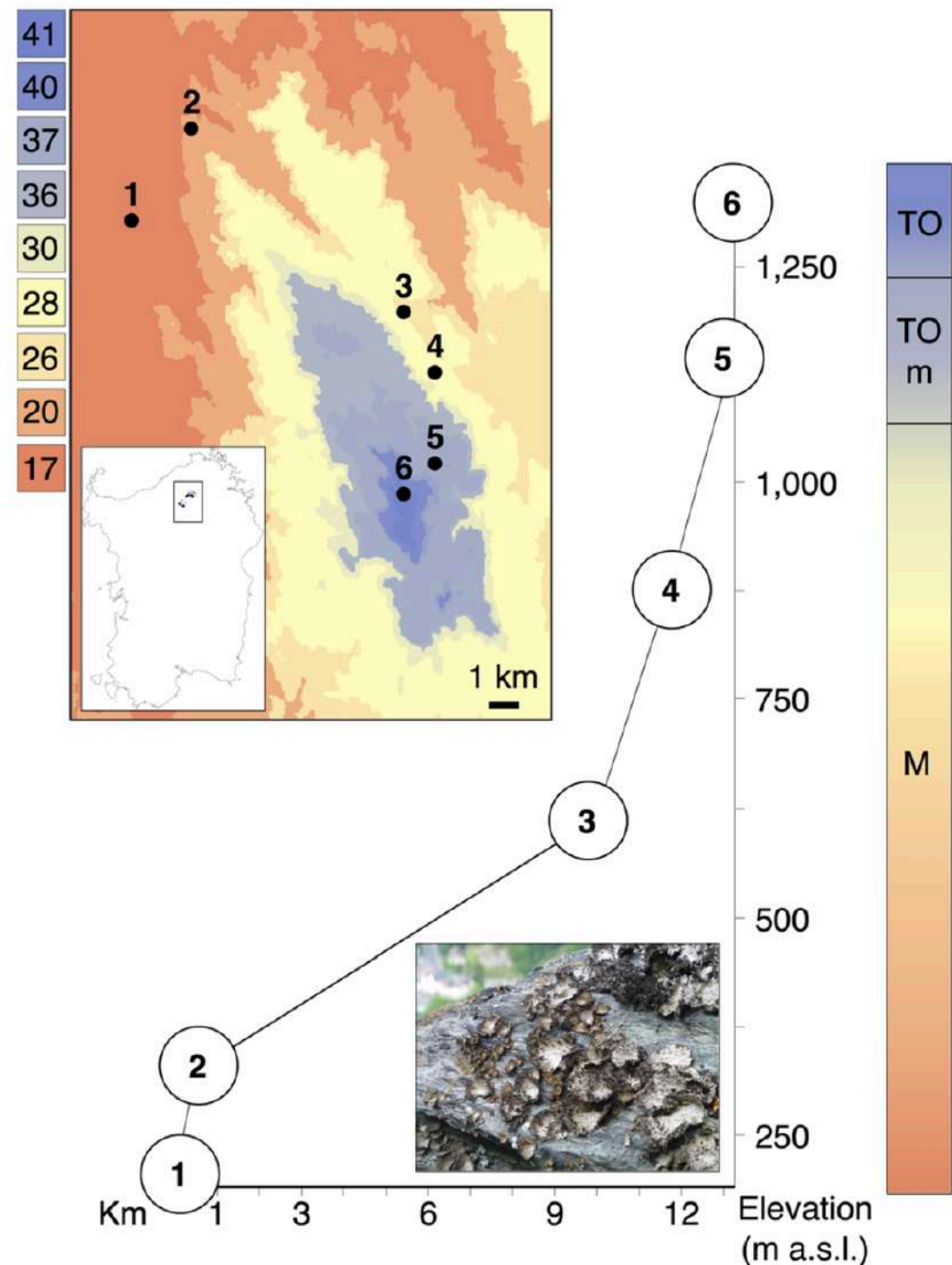
This indicates substantial epistatic interactions with the genetic background, such that the gene is important in one strain and but dispensable in another



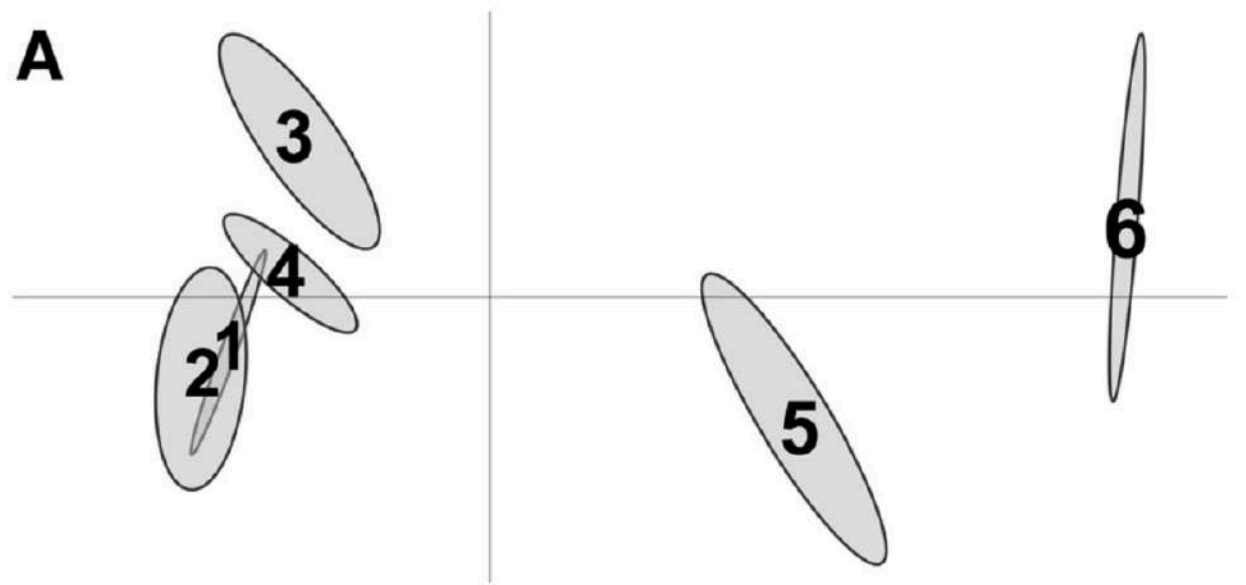
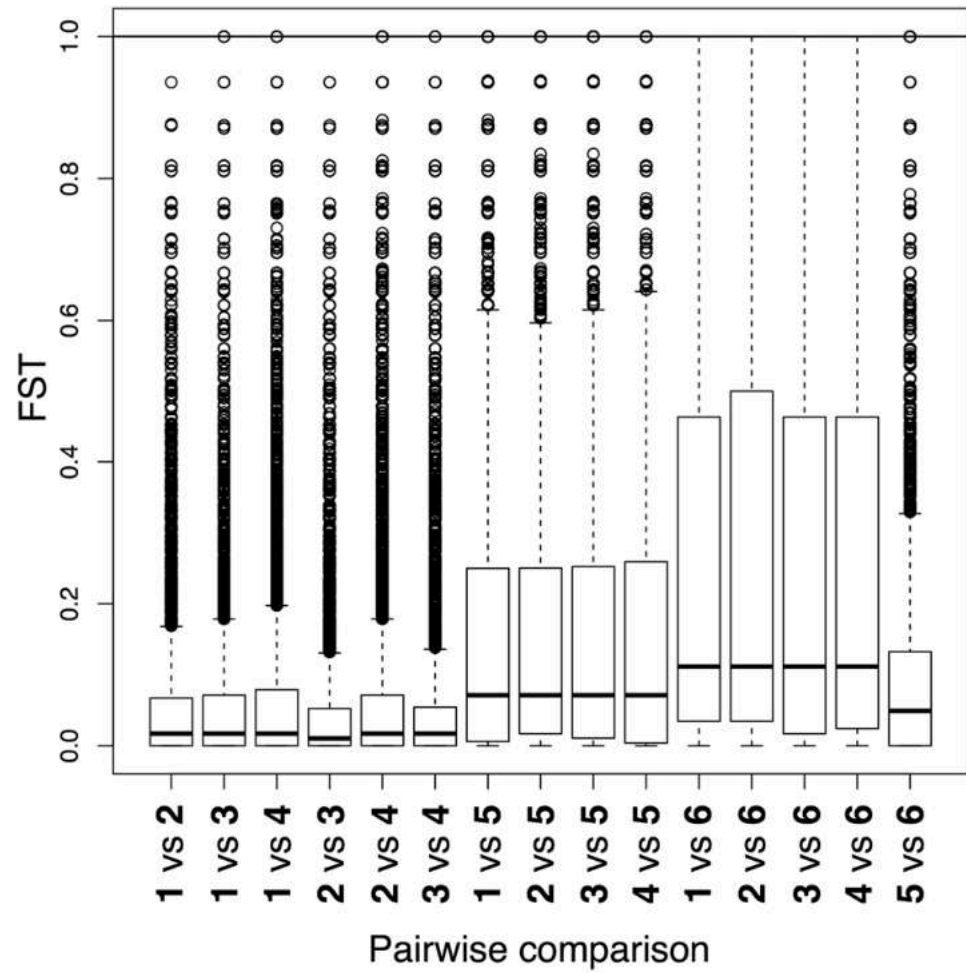
Adaptive differentiation coincides with local bioclimatic conditions along an elevational cline in populations of a lichen-forming fungus

Francesco Dal Grande^{1*}, Rahul Sharma², Anjuli Meiser^{1,2}, Gregor Rolshausen¹, Burkhard Büdel³, Bagdevi Mishra¹, Marco Thines¹, Jürgen Otte¹, Markus Pfenninger^{1,2} and Imke Schmitt^{1,2*}

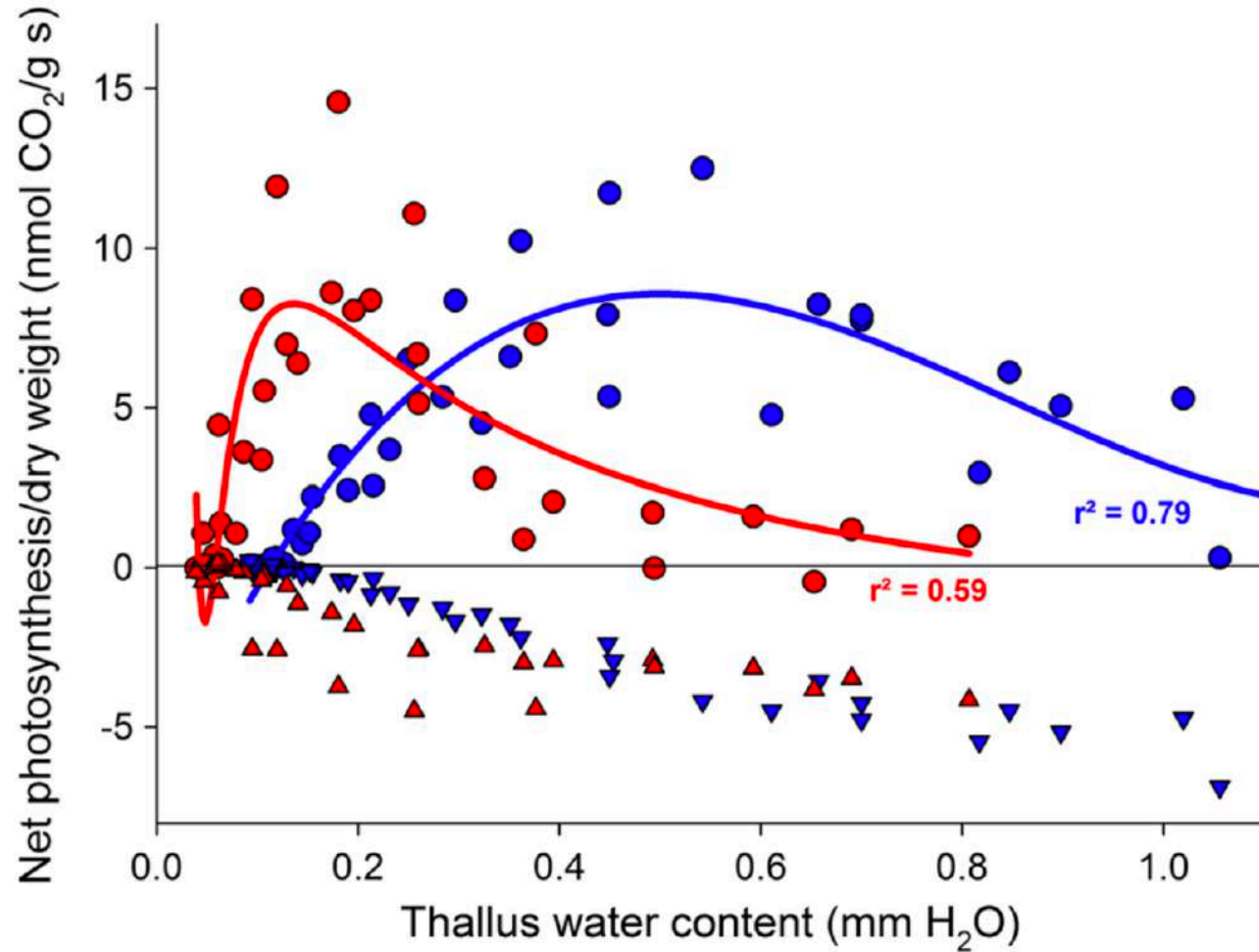
Here we report on the population genomics of a lichen-forming ascomycete along an altitudinal gradient in the Mediterranean region. As model, we chose *Lasallia pustulata* (Umbilicariaceae), a species with a distribution from southern Europe to northern Scandinavia, which forms dense populations on exposed, siliceous rocks [30]. Using genomic data from geographically close populations along a steep altitudinal gradient in northern Sardinia (Italy), we analyzed whether genetic clusters were present, and whether relatedness between clusters was correlated with signatures of local adaptation. Heat, drought, and radiation stress constitute determining factors in the composition of biological communities inhabiting rocky outcrops and boulders in Mediterranean mountains [31]. Therefore we tested the hypothesis that environmental factors shape genome-wide population differentiation in lichenized fungi which occur across different bioclimatic regions. Specifically, we addressed the following questions: i) what is the genome-wide population structure and connectivity between geographically close populations along an elevation gradient?, ii) what are putative functions of highly differentiated genes between the genetic clusters?, iii) what are putative functions of the genes showing strong correlation with local climatic factors?, and iv) do individuals belonging to different genetic clusters (and environments) display fitness differences?



Strong genetic structure separating lower altitude populations and rest



Photosynthetic CO₂ gas exchange of *L. pustulata* highland (population 6; blue) and lowland population (populations 1 to 5; red) related to thallus water content (TWC).



The genetic separation coincided with differences in physiological responses to thallus water content (WC)

“Future studies ..., and quantitative trait locus mapping experiments of the candidate genes in controlled and field settings will help to elucidate the drivers of local adaptation in this and other fungal species.”

Moving beyond EG

What's next?

Genotypes

- Genotype frequencies
- Genomic variations
- Population genomics
- Comparative genomics

Phenotypes

- Phenotype frequencies
- Phenotype plasticity
- Development

Ecology

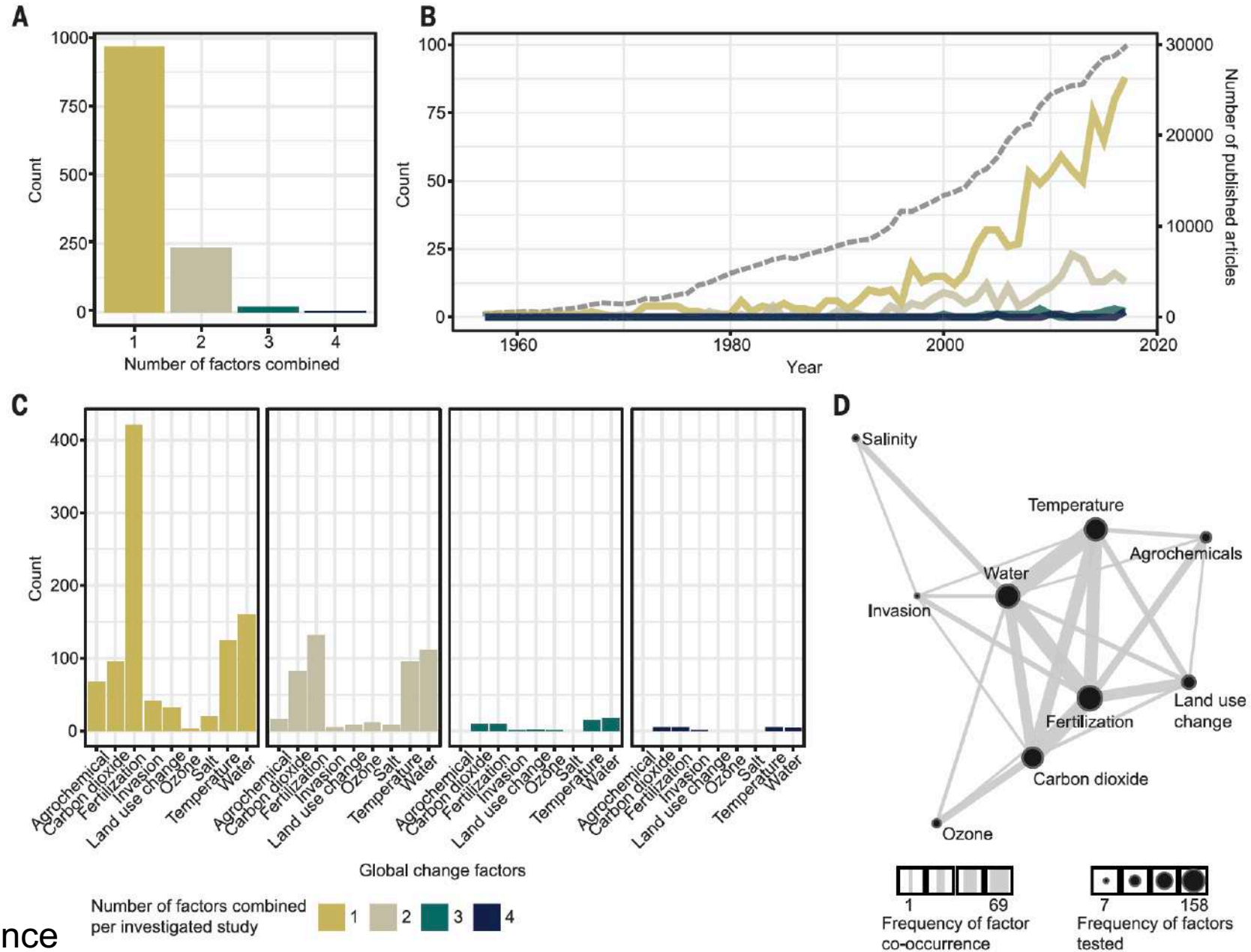
- Abiotic
- Biotic
- Short term / long term

SOIL ECOLOGY

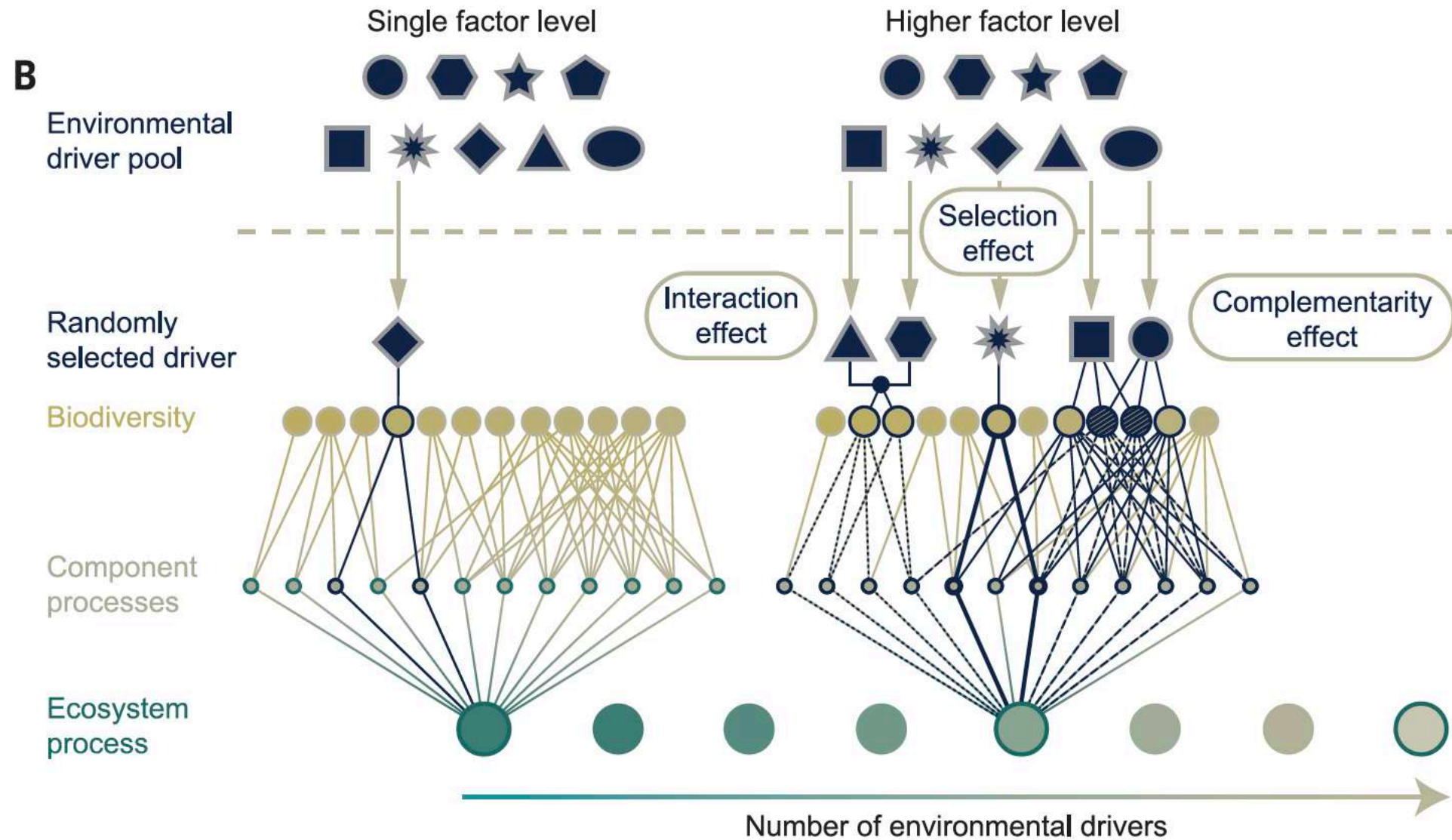
The role of multiple global change factors in driving soil functions and microbial biodiversity

Matthias C. Rillig^{1,2*†}, Masahiro Ryo^{1,2*}, Anika Lehmann^{1,2}, Carlos A. Aguilar-Trigueros^{1,2}, Sabine Buchert^{1,2}, Anja Wulf^{1,2}, Aiko Iwasaki^{1,2}, Julien Roy^{1,2}, Gaowen Yang^{1,2}

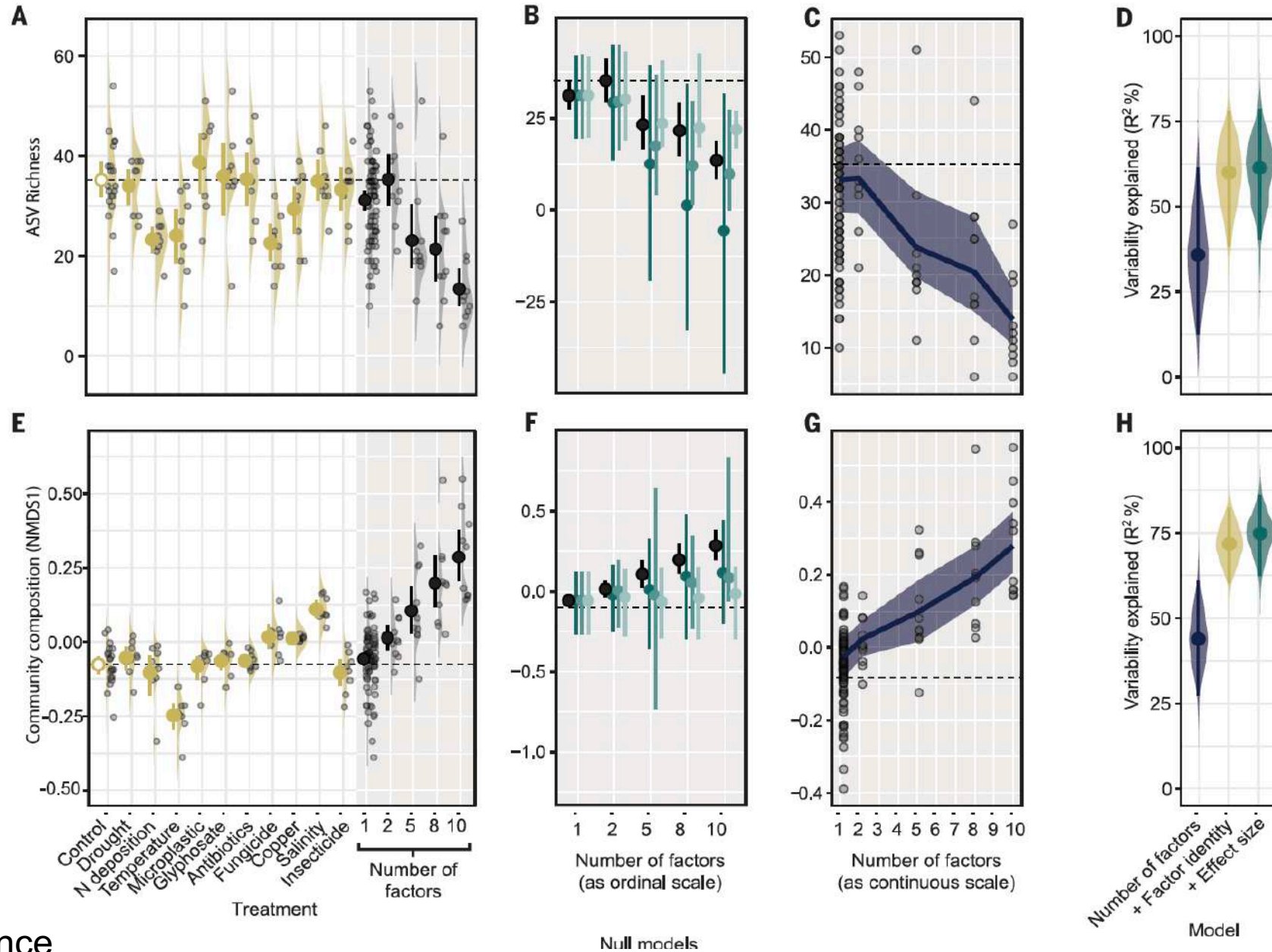
Factor in soil ecology experiments



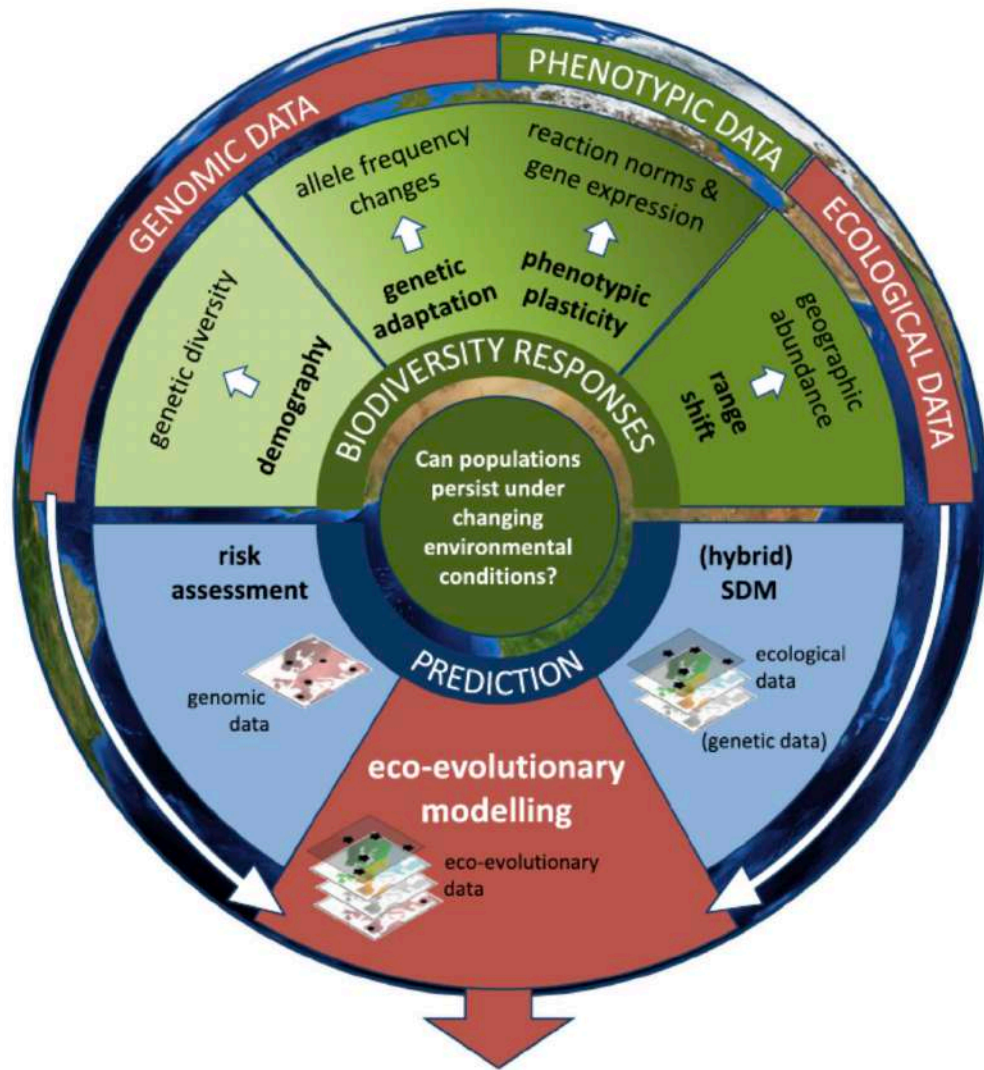
the number of global change factors alone might predict general trends in changes of biodiversity and ecosystem processes



Effects on the soil fungal community of different global change factors applied singly and using different numbers of factors

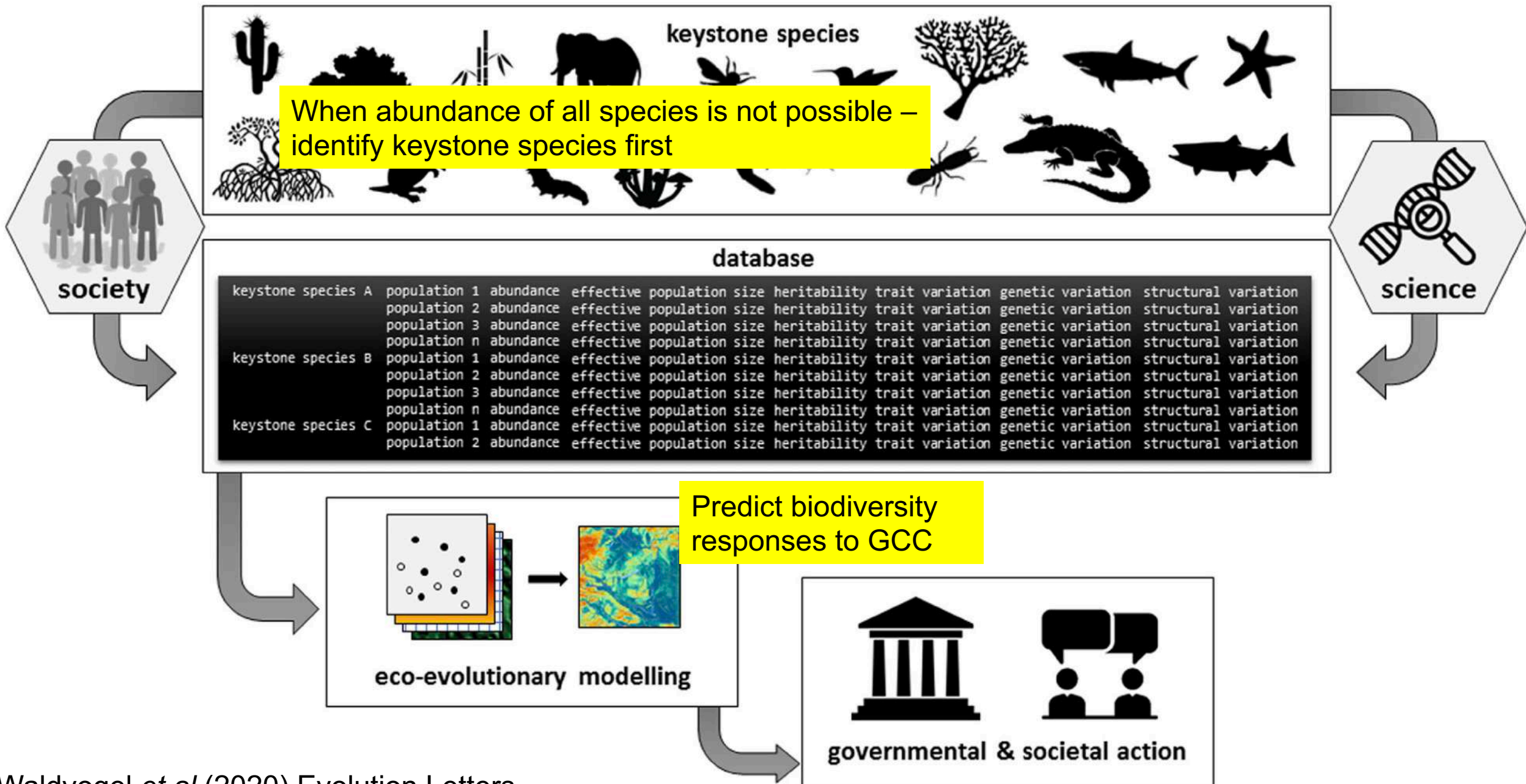


Predicting species' responses to climate change



promising strategy to use genomics for predictions:
fast enough and for a broad taxon spectrum

**ecosystem management &
conservation strategies**



Questions and knowledge gaps

KEY UNRESOLVED QUESTIONS:



What is the relative importance of fungal adaptation, migration and acclimatisation?

How does climate change affect the yield of fungal spore-bearing structures?

How does climate change affect fungal growth and activity?

How do fungi mediate ecosystem responses to climate change?

How do changes in the phenology of spore-bearing structure production reflect changes in activity, abundance, biomass and distribution?

Can fungi track climate space shifts?

KEY INFORMATION GAPS:



Long-term data

Large-scale data

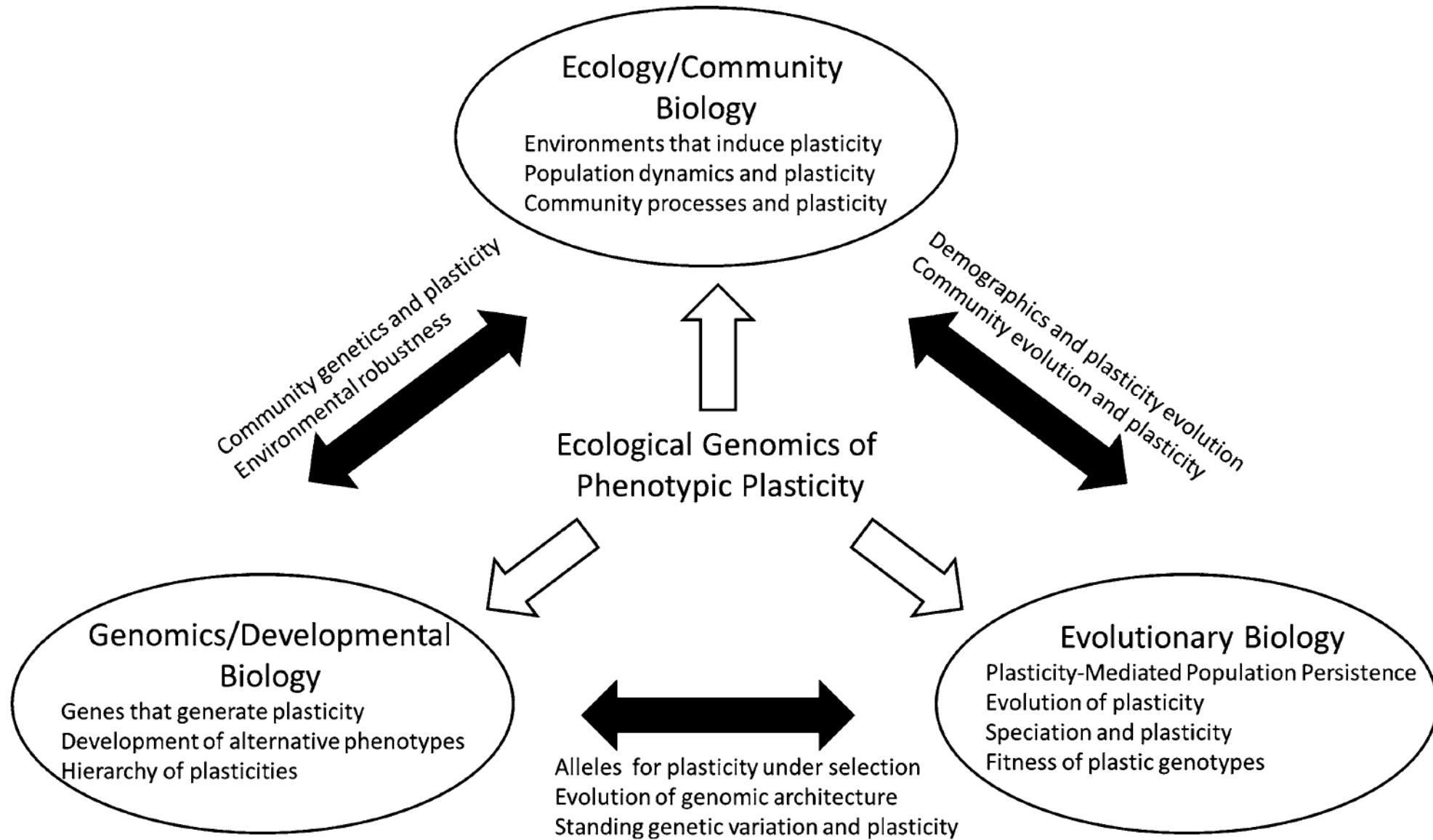
Data from tropical, subtropical and warm temperate ecosystems

Experimental data from fungi associated with trees, rather than seedlings

Fungal response and effect traits

Data from multiple simultaneous drivers of change (nitrogen deposition, carbon dioxide, ozone, UV, temperature, drought, fire)

Not covered but very important – phenotypic plasticity



More references

#Some nice papers

<https://www.nature.com/subjects/fungi>

Fungal Phylogenetics and Phylogenomics

<https://www.sciencedirect.com/bookseries/advances-in-genetics/vol/100/suppl/C>

The Impact of Molecular Data in Fungal Systematics

<https://www.sciencedirect.com/science/article/pii/S0065229605420029>

Genome Diversity and Evolution in the Budding Yeasts (Saccharomycotina)

<http://www.genetics.org/content/206/2/717>

Evolutionary biology through the lens of budding yeast comparative genomics

<https://www.nature.com/articles/nrg.2017.49>

Dimensions of biodiversity in the Earth mycobiome

<https://www.nature.com/articles/nrmicro.2016.59>

Evolution and genome architecture in fungal plant pathogens

<https://www.nature.com/articles/nrmicro.2017.76>

Ecological genomics: understanding gene and genome function in the natural environment

Ungerer *et al* (2007) Heredity

<https://www.nature.com/articles/6800992>