Divided bacterial genomes: structure, function, and evolution

George diCenzo
georgecolin.dicenzo@unifi.it

University of Florence, Florence, Italy

Outline

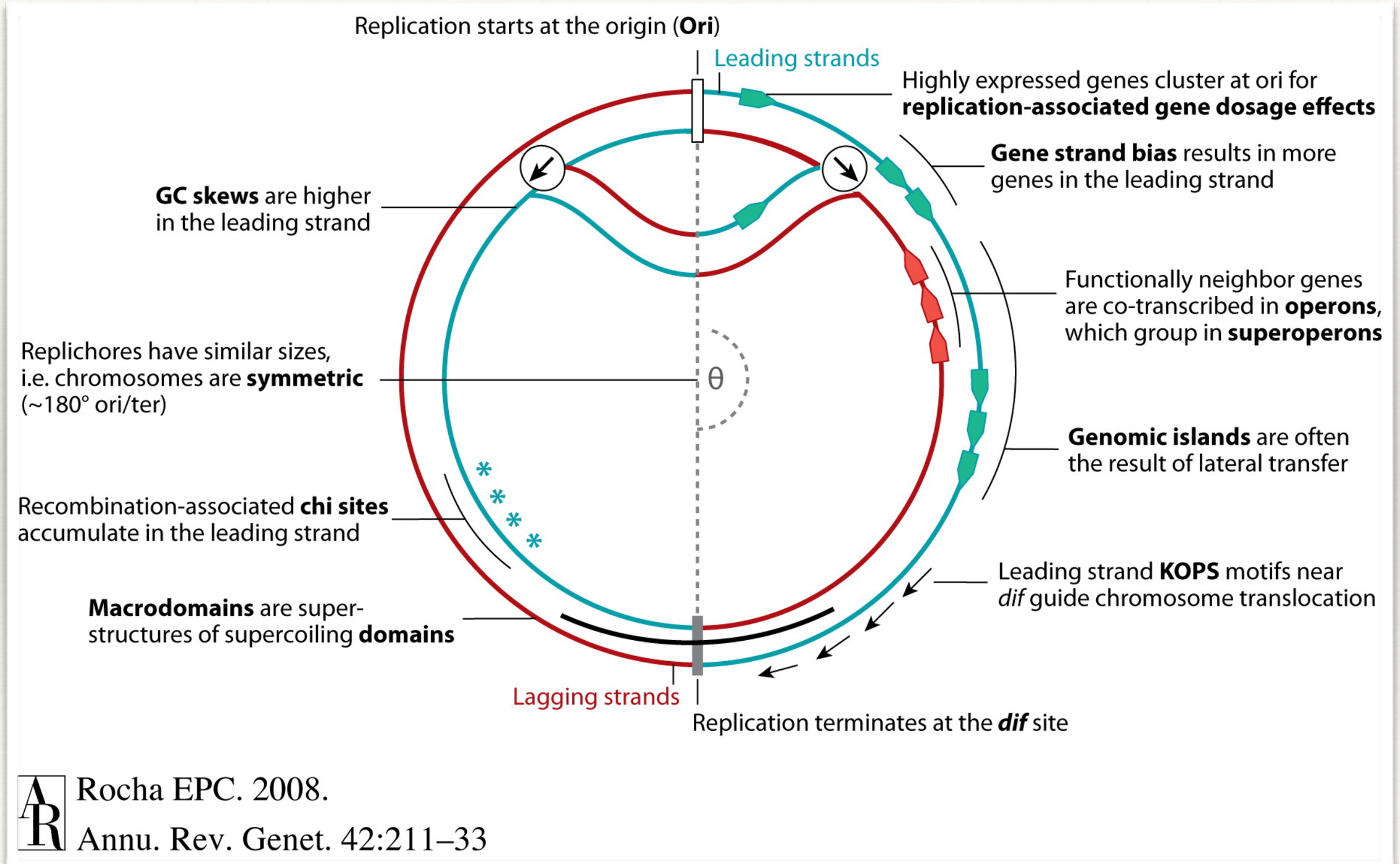
1. Genome organization and introduction to divided genomes
2. Characteristics of divided genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

Outline

1. Genome organization and introduction to divided genomes
2. Characteristics of multipartite genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

Bacterial genomes: organization = function

- Bacterial genomes are highly organized



Rocha EPC. 2008.
Annu. Rev. Genet. 42:211–33

Bacterial genomes: organization = function

- Bacterial genomes are highly organized
- Pathways often encoded by genes organized in an operon
- Related operons or regulators in often in close proximity

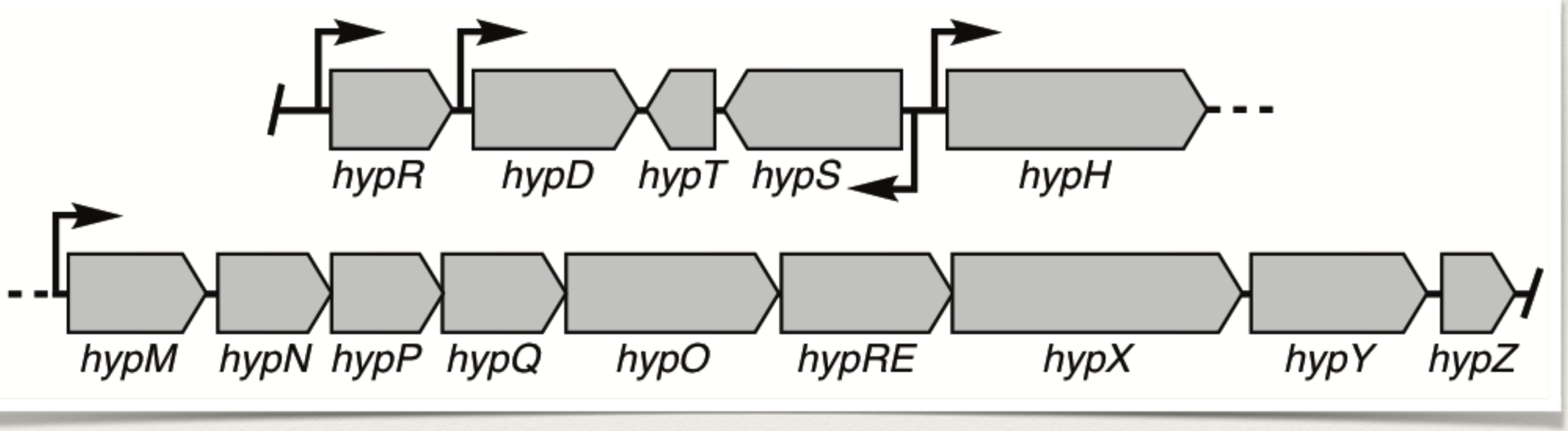


Image from Chen *et al*. 2016. J Bacteriol 198: 1171

Bacterial genomes: organization = function

- Bacterial genomes are highly organized
- Pathways often encoded by genes organized in an operon
- Related operons or regulators in often in close proximity
- Chromosomal location influences expression level

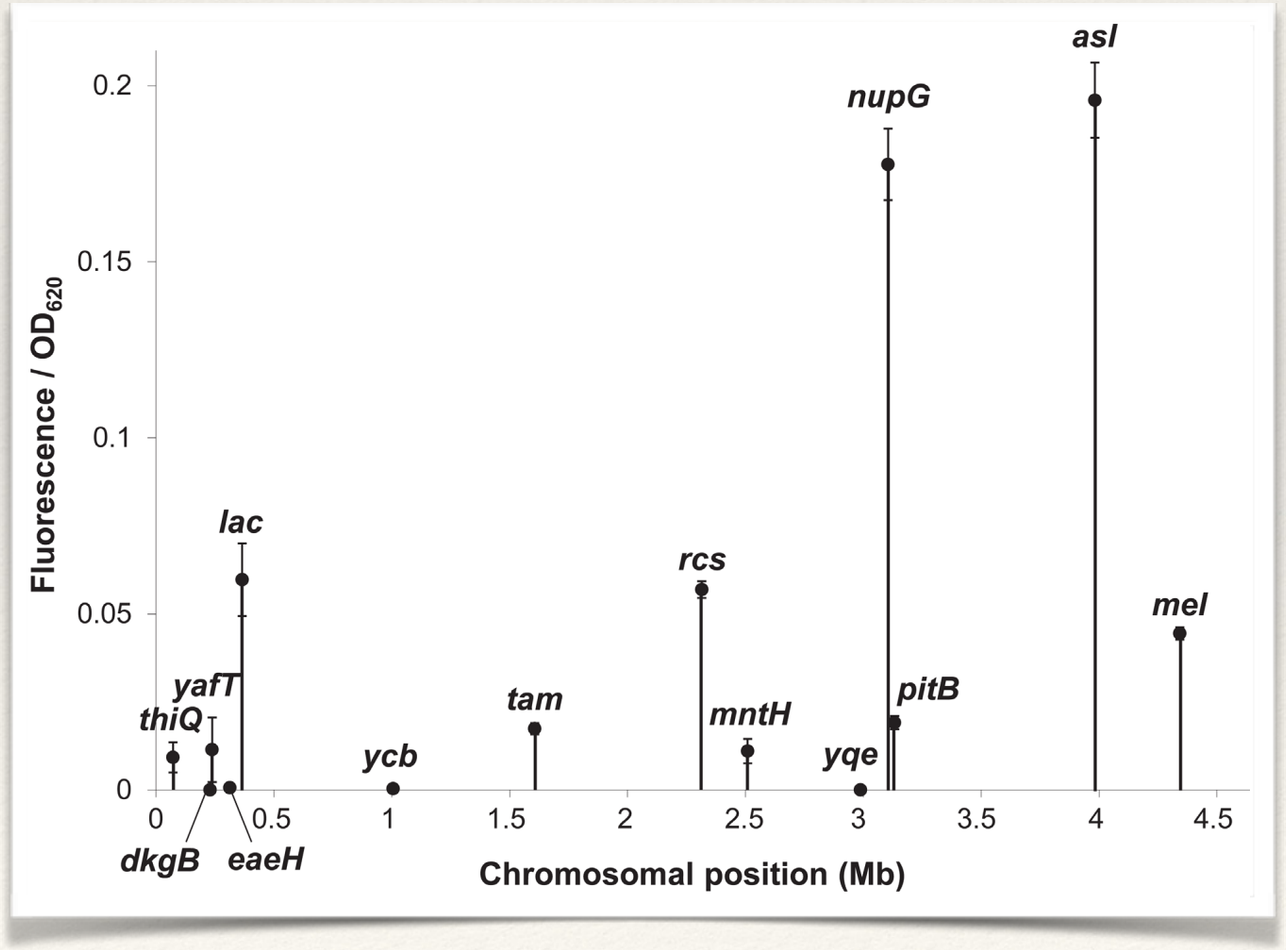


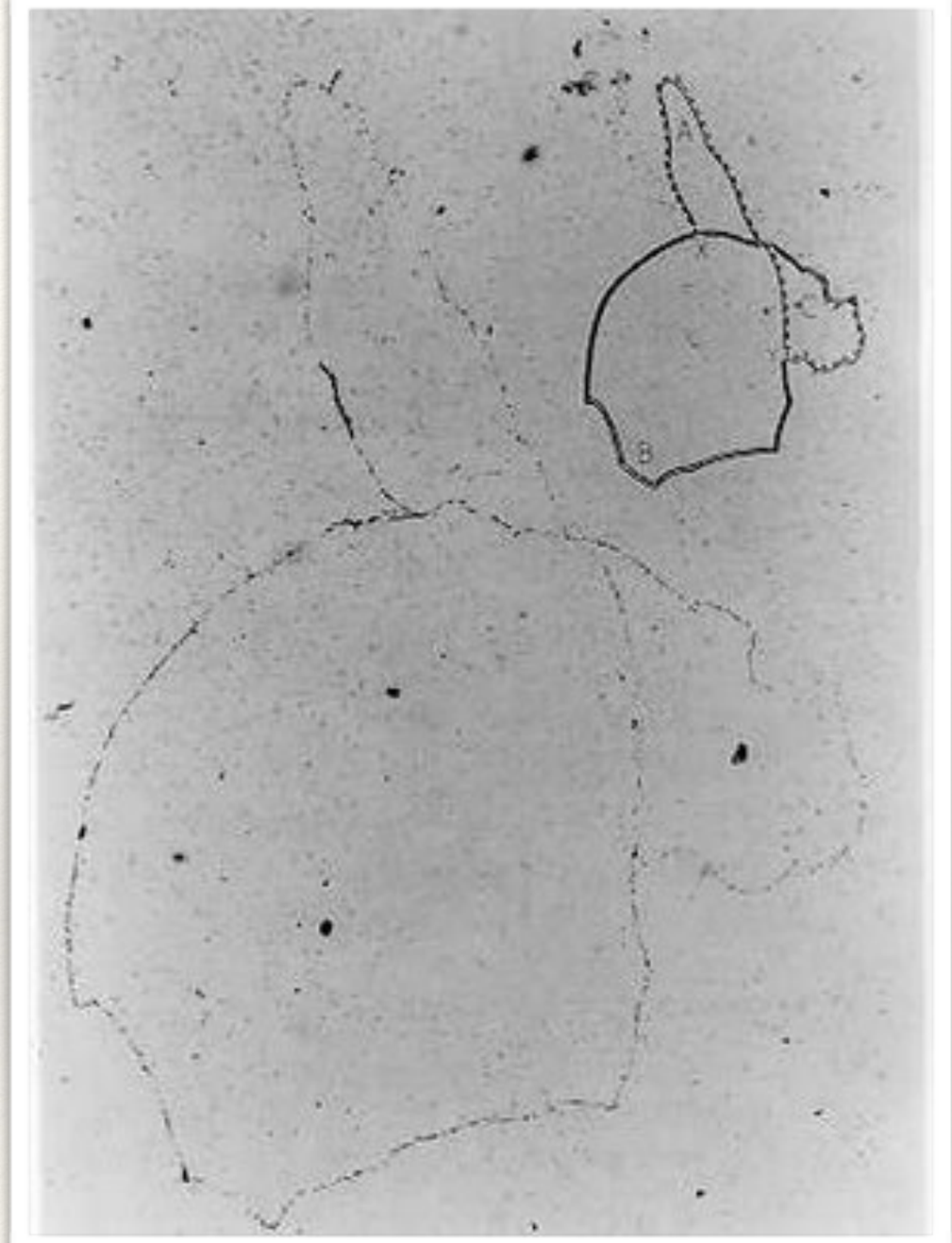
Image from Bryant *et al*. 2014. Nucleic Acids Res. 42: 11383

Bacterial genomes: organization = function

- Bacterial genomes are highly organized
- Pathways often encoded by genes organized in an operon
- Related operons or regulators in often in close proximity
- Chromosomal location influences expression level
- ‘Hot spots’ for horizontal gene transfer

Evolving view of the bacterial genome

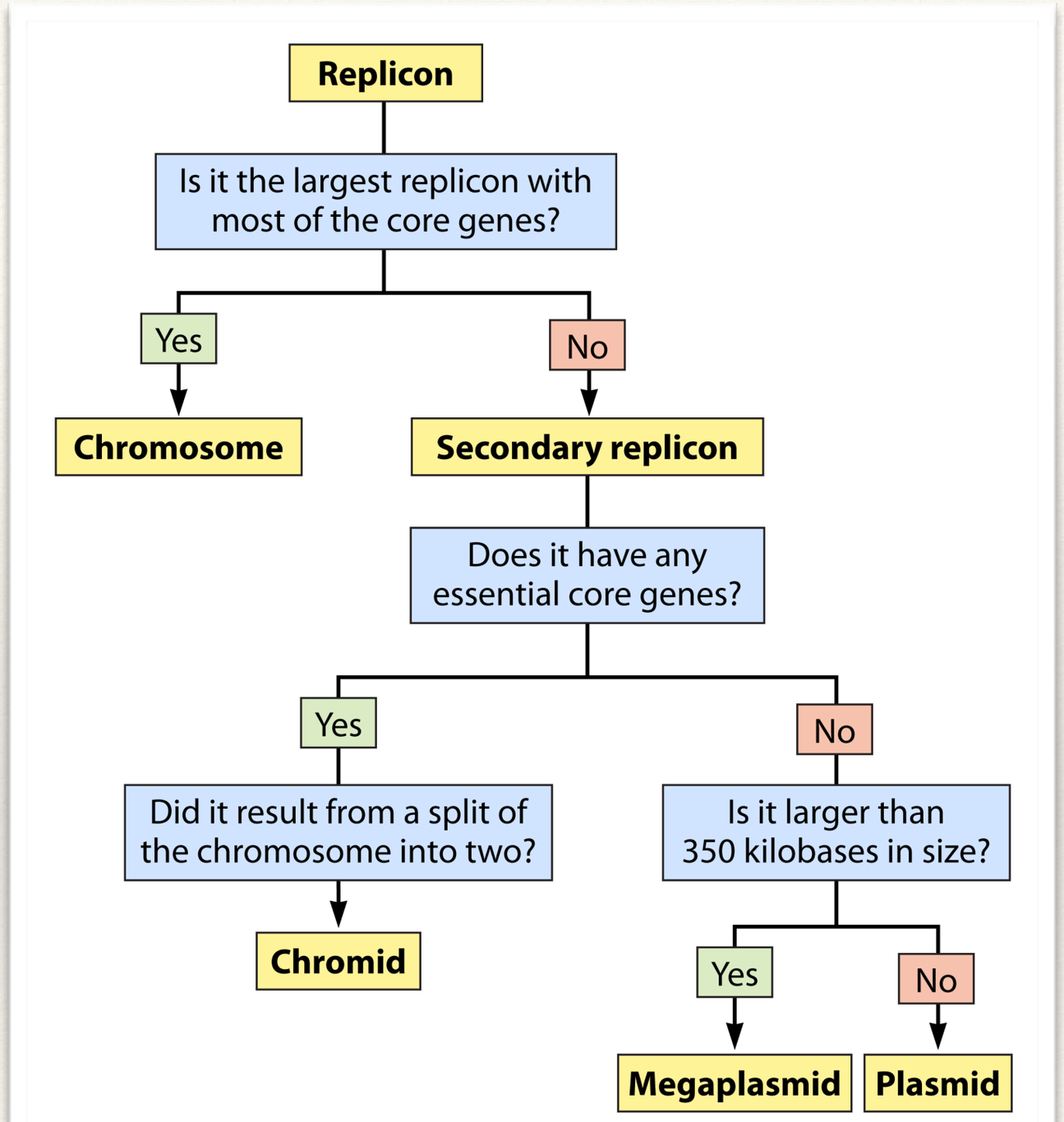
- 1963: *Escherichia coli* genome is a single, circular chromosome
- 1981: Identification of a *Sinorhizobium meliloti* megaplasmid
- 1989: *Borrelia burgdorferi* chromosome is linear
- 1989: *Rhodobacter sphaeroides* genome has a second chromosome (chromid)



Circular *E. coli* chromosome undergoing replication (Cairns 1963. Cold Spring Harbor Symposia on Quantitative Biology. 28: 44)

Terminology

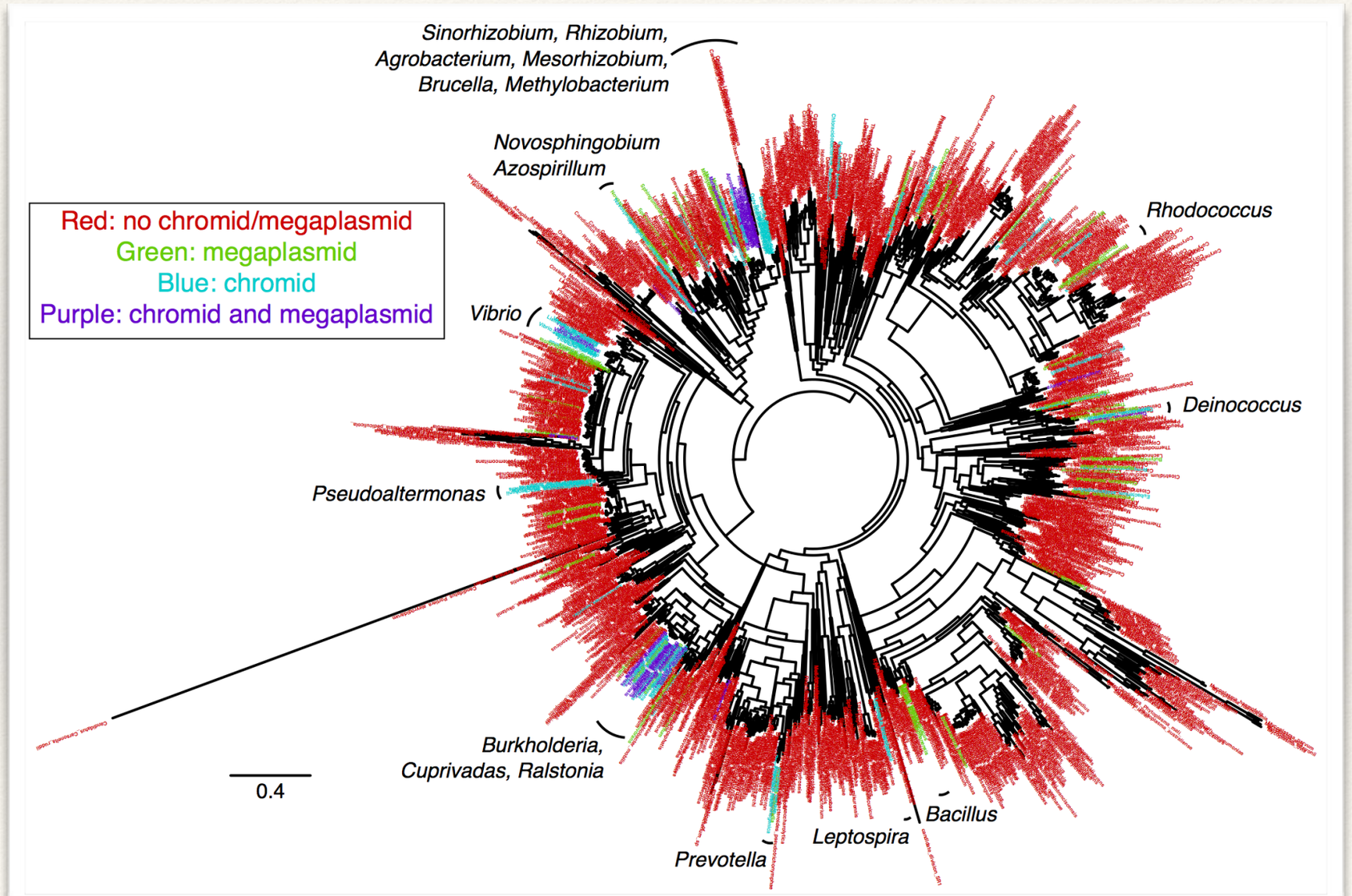
- Replicon: a general term to refer to any DNA molecule
  - Technically refers to DNA molecules with a single origin of replicons, excluding many archaeal chromosomes
- Chromosome: the primary replicon that encodes almost all essential cellular proteins
- Megaplasmid: a large (> 350 kb) replicon that lacks essential genes
- Chromid: intermediate between a chromosome and a megaplasmid, with at least one core gene
- Multipartite/divided genome: Contains a megaplasmid/chromid



Flow chart for classification of bacterial replicons (diCenzo and Finan. 2017. MMBR. 81: e00019)

Phylogenetic distribution of multipartite genomes

- Multipartite genomes (particularly chromids) are most common in the Proteobacteria, but can be found throughout the bacterial phylogeny
- Appear enriched in species relevant to human society



Phylogeny of 1,708 bacterial species, coloured based on genome structure

(diCenzo and Finan. 2017. MMBR. 81: e00019)

Outline

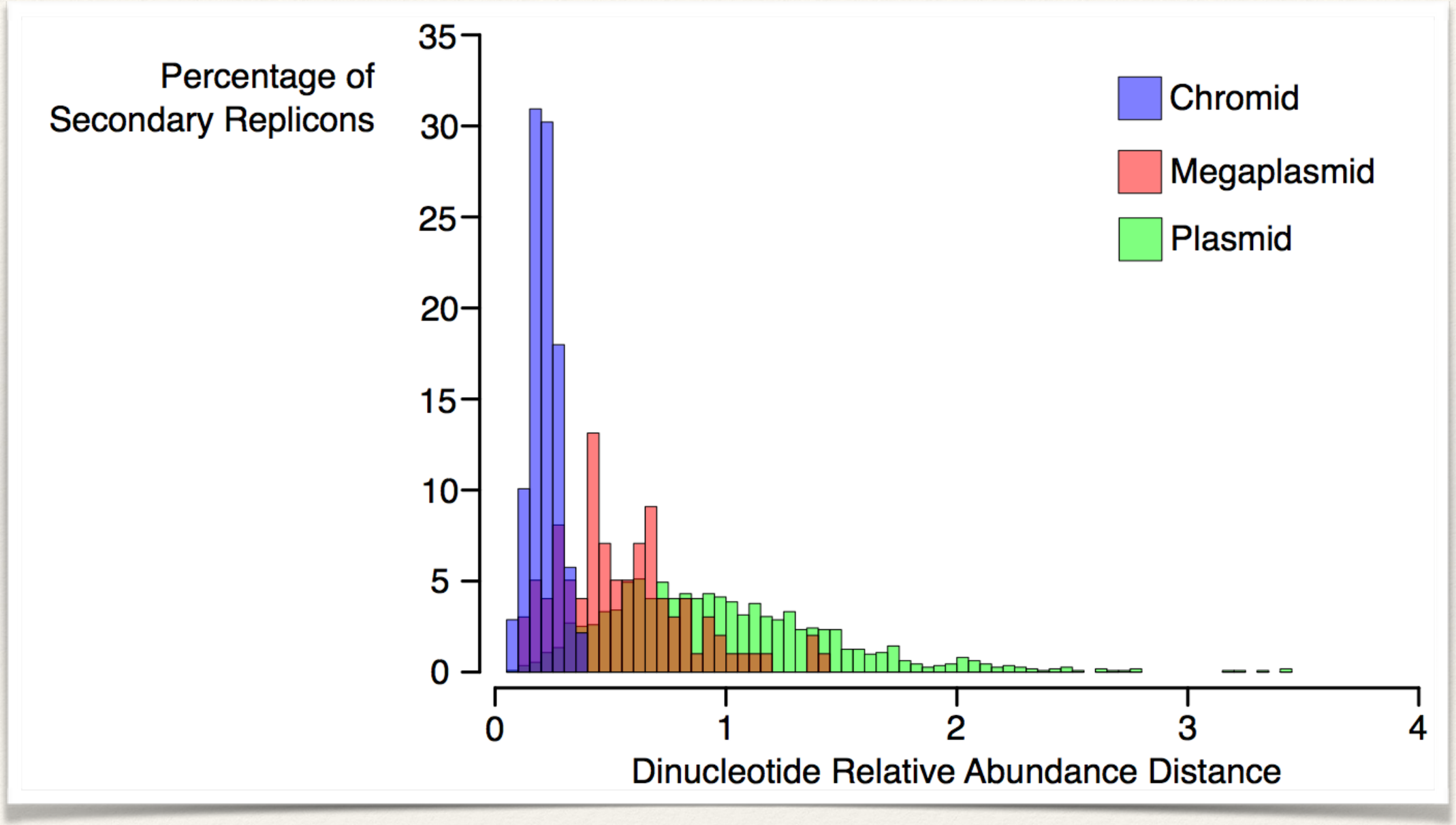
1. Genome organization and introduction to divided genomes
2. Characteristics of divided genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

Genomic signatures

- Genomic signature: "any specific quantitative characteristic of a sequence that is pervasive along the genome, while being dissimilar for sequences originating from organisms of different species" (Karamichalis *et al*. 2016. BMC Bioinformatics. 17: 313)
  - GC content (% guanine/cytosine)
  - Codon usage (ratio of synonymous codons)
  - Dinucleotide relative abundance (frequency of each pair of nucleotides)
- Can also (on average) distinguish chromosomes from chromids from megaplasmids from plasmids

Genomic signatures

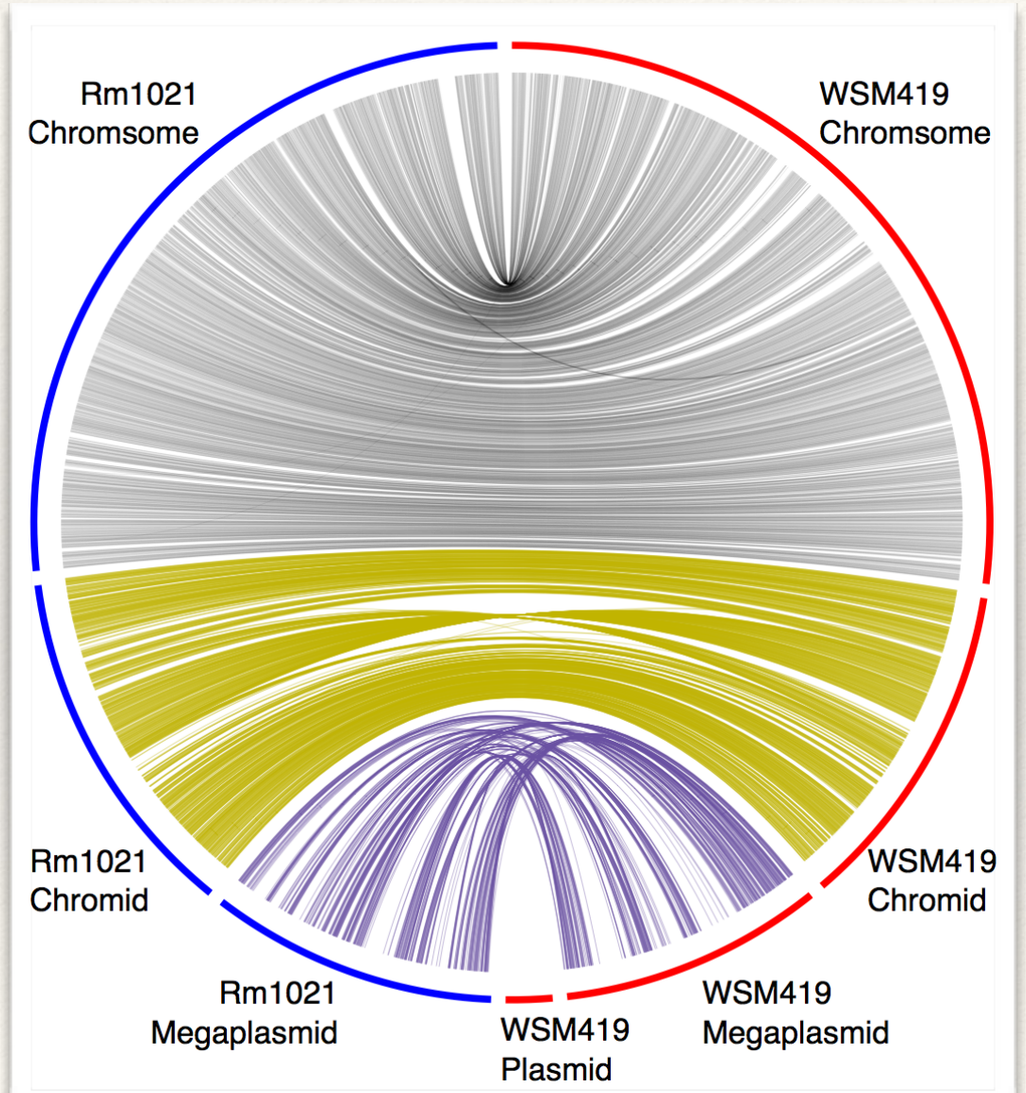
- Chromids most similar to chromosomes, plasmids least similar



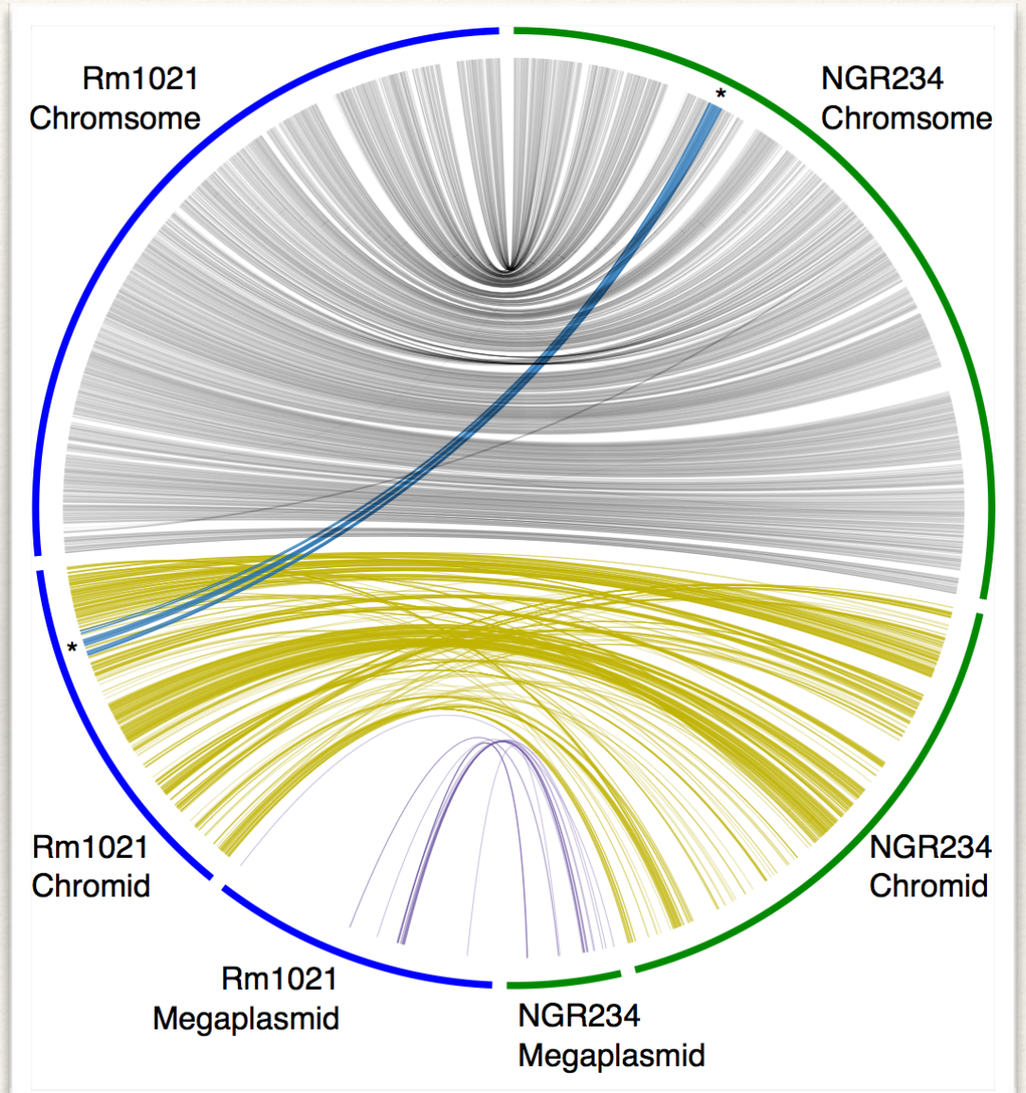
Distribution of the difference in dinucleotide profiles of secondary replicons compared to the corresponding chromosome (diCenzo and Finan. 2017. MMBR. 81: e00019)

Genetic variability

- Chromosomes are most genetically stable, followed by chromids, and last by megaplasmids



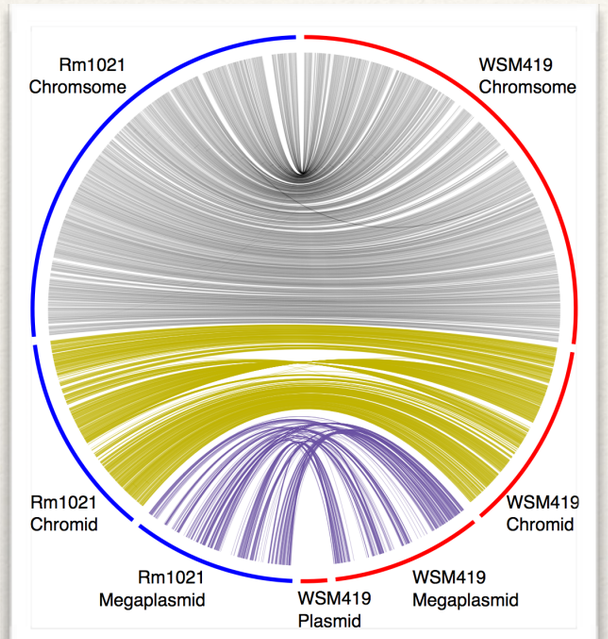
Comparing *S. meliloti* and *S. medicae* (closely related)

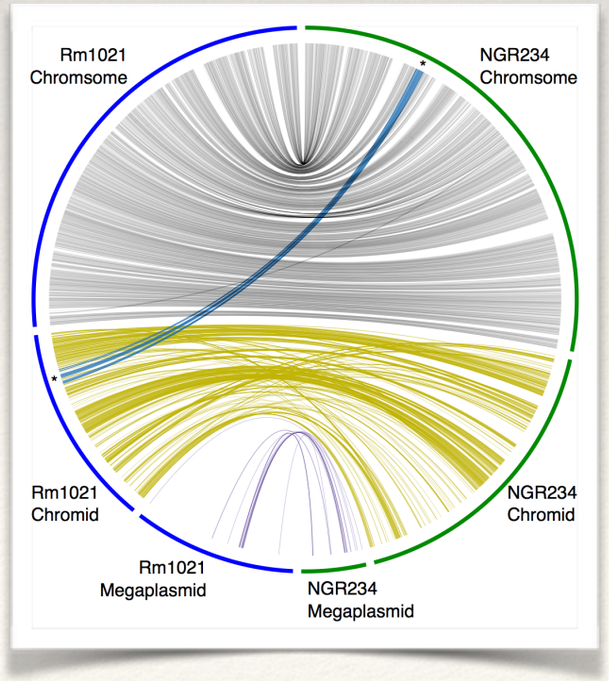


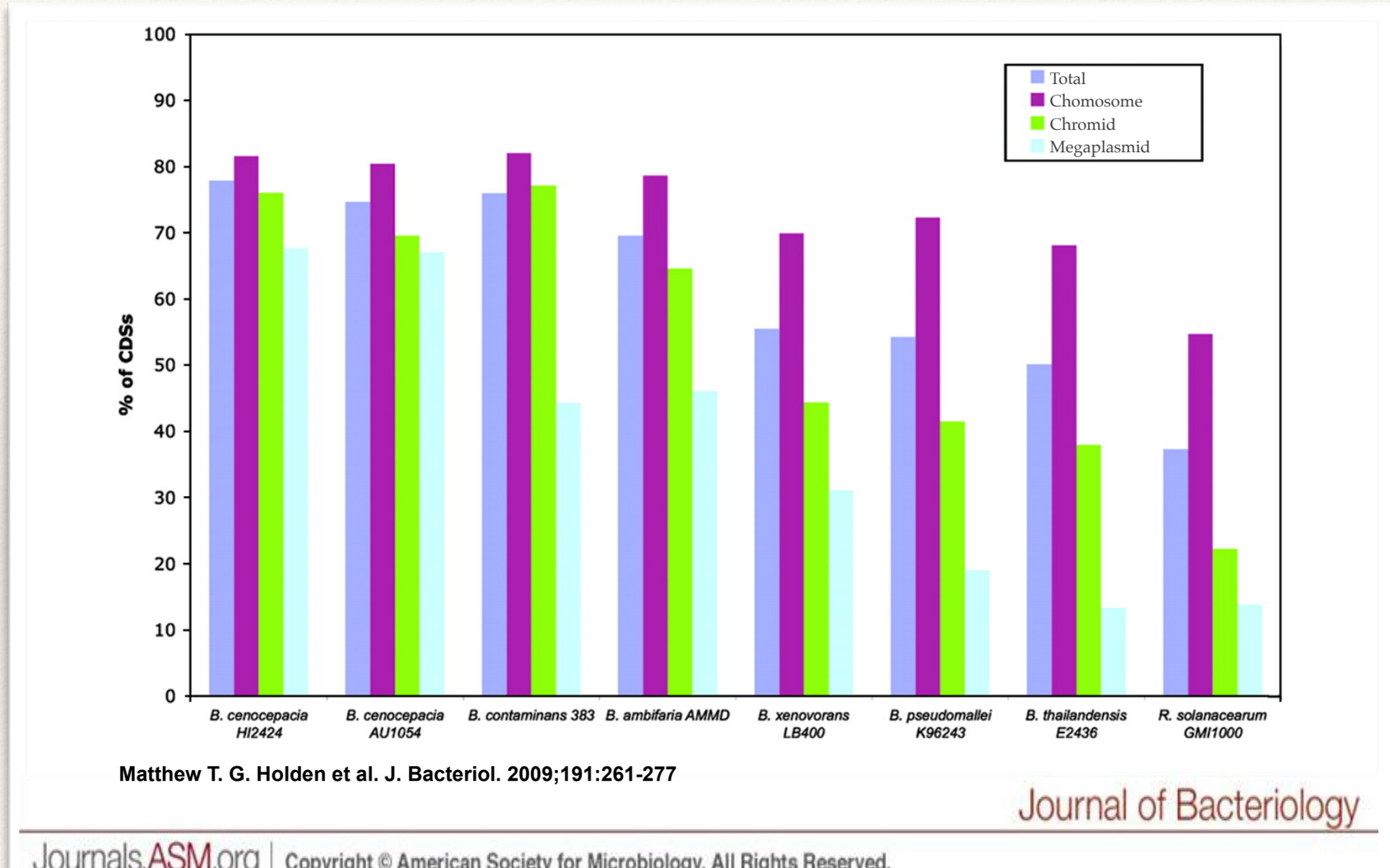
Comparing *S. meliloti* and *S. fredii* (less closely related)

Genetic variability

- Chromosomes are most genetically stable, followed by chromids, and last by megaplasmids







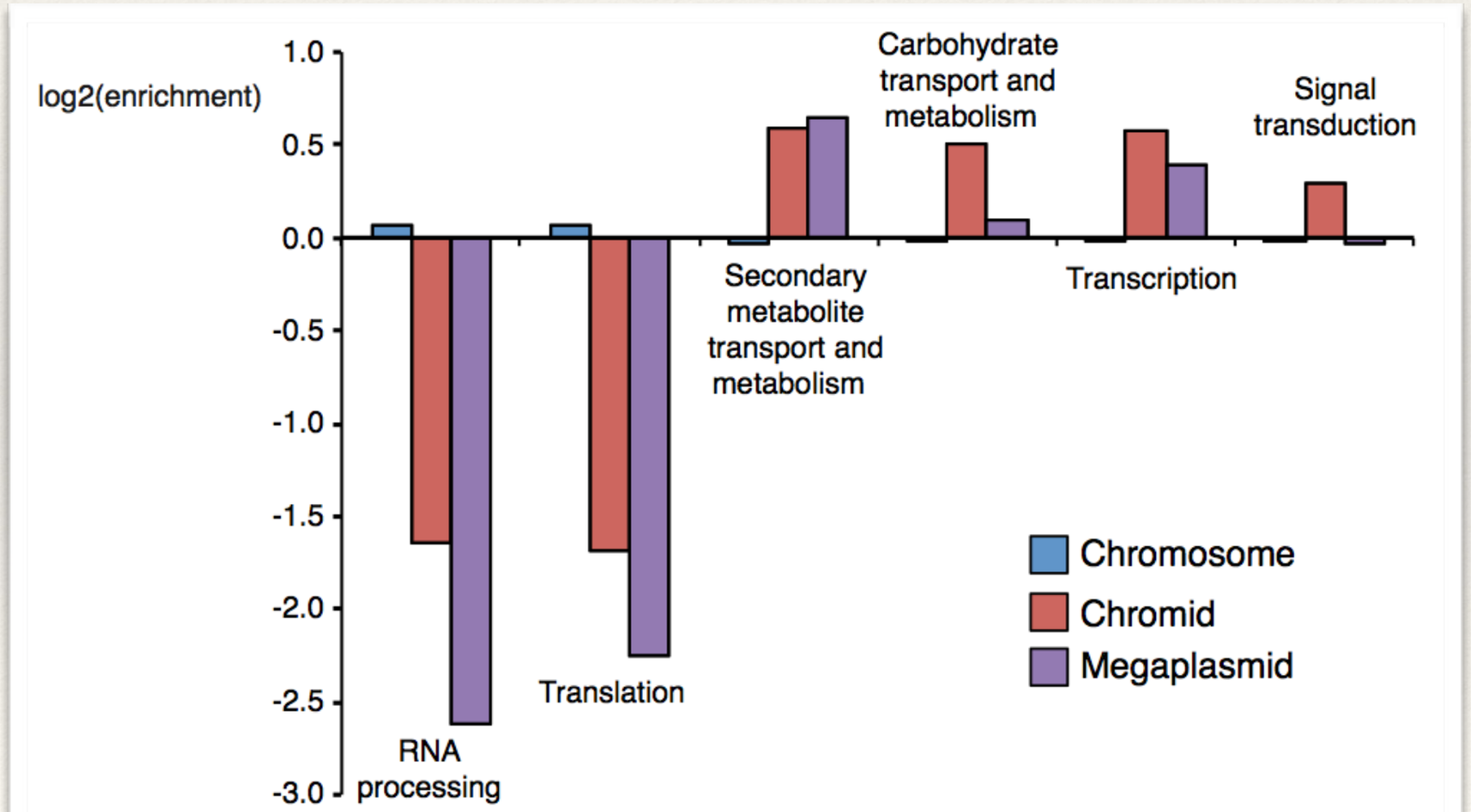
Matthew T. G. Holden et al. J. Bacteriol. 2009;191:261-277

Journal of Bacteriology

Number of *B. cenocepacia* genes conserved in other strains/species grouped by replicon

Functional biases

- Global biases can be detected in the function of proteins encoded by chromosomes, chromids, and megaplasmids
  - Chromosome enriched in core functions (e.g., RNA processing, translation, cytoskeleton, cell division)
  - Secondary replicons enriched in functions such as transport, metabolism, regulation



Enrichment of 5 functional categories on different replicon classes

Inter-replicon communication

- In many ways, each replicon functions independently. Yet, there is some co-operation among replicons
  - Transcription factors (e.g., CtrA in *S. meliloti*)
  - Biosynthetic pathways (e.g., pantothenate biosynthesis in *R. etli*)
  - Pervasive genetic interactions
- Dependent on length of co-evolution

Replicon	chr	p42a	p42b	p42c	p42d	p42e	p42f
chr	4.22	0.013	0.135	0.177	0.116	0.312	0.351
p42a	0.3	0.989	0.021	0.021	0.153	0.027	0.032
p42b	3.345	0.024	1.339	0.127	0.048	0.394	0.357
p42c	3.085	0.017	0.089	1.427	0.085	0.295	0.607
p42d	1.341	0.079	0.022	0.056	1.059	0.082	0.152
p42e	2.769	0.011	0.142	0.15	0.063	1.18	0.477
p42f	2.493	0.01	0.102	0.247	0.094	0.382	0.989

Measure of functional interactions between replicons in *R. etli*

(Gonzalez *et al*. 2006. PNAS. 103: 3834)

Outline

1. Genome organization and introduction to divided genomes
2. Characteristics of divided genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

How do megaplasmids form?

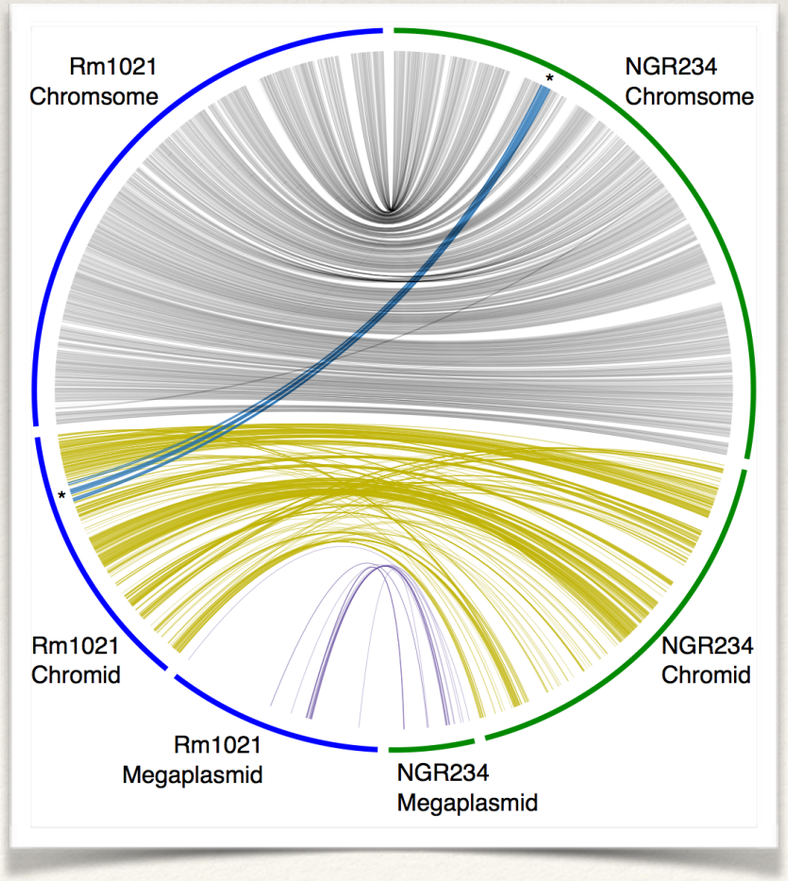
- Horizontal gene transfer

How do chromids form?

- Two hypotheses have been proposed:
- Schism hypothesis: split of an ancestral chromosome into two replicons
- Plasmid hypothesis: evolve from a megaplasmid
  - introgression of genomic signatures and gain of essential genes from the chromosome

How do chromids form?

- Two hypotheses have been proposed:
- Schism hypothesis: split of an ancestral chromosome into two replicons
- Plasmid hypothesis: evolve from a megaplasmid
  - introgression of genomic signatures and gain of essential genes from the chromosome
  - Accounts for genomic signature differences and unequal essential gene distribution

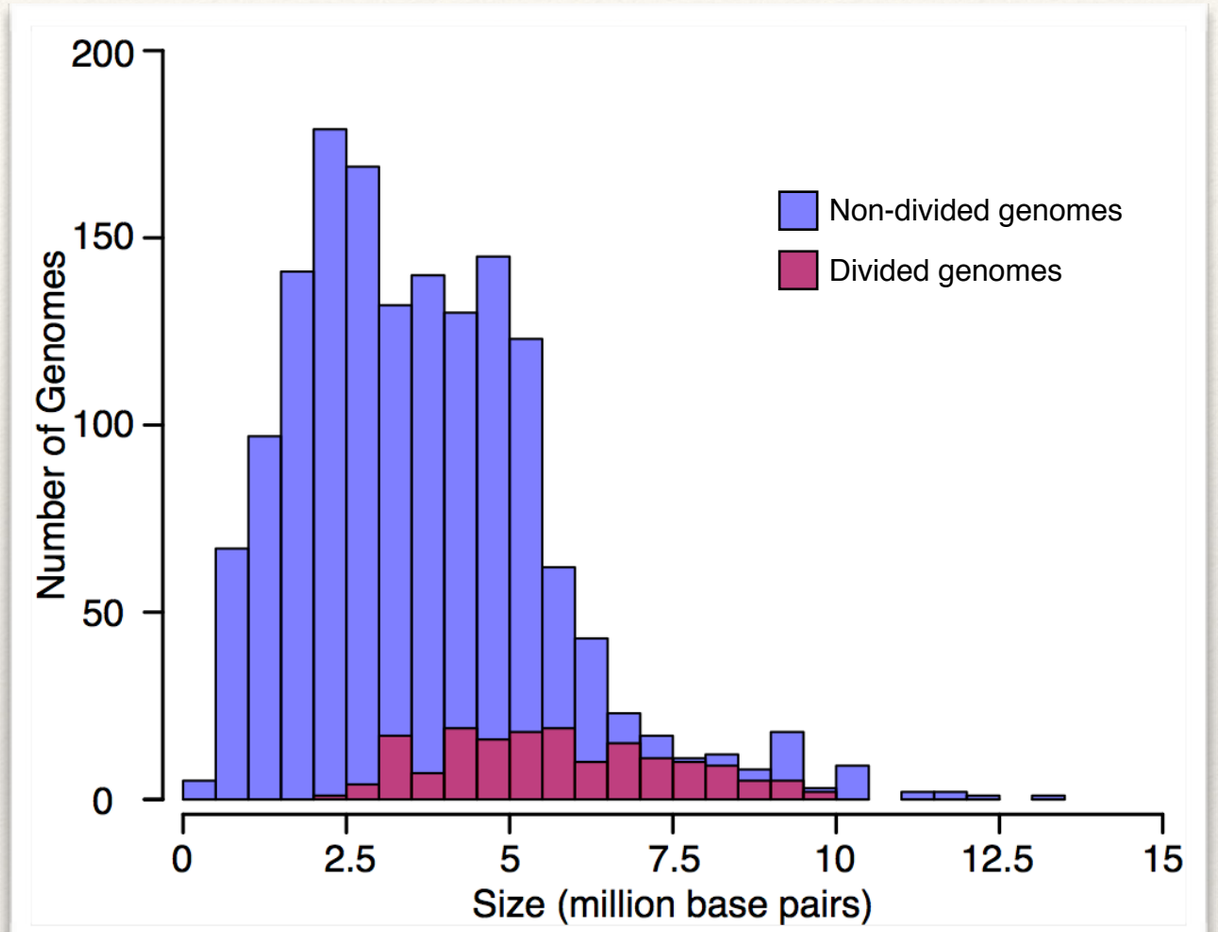


Outline

1. Genome organization and introduction to divided genomes
2. Characteristics of divided genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

Why maintain a multipartite genome?

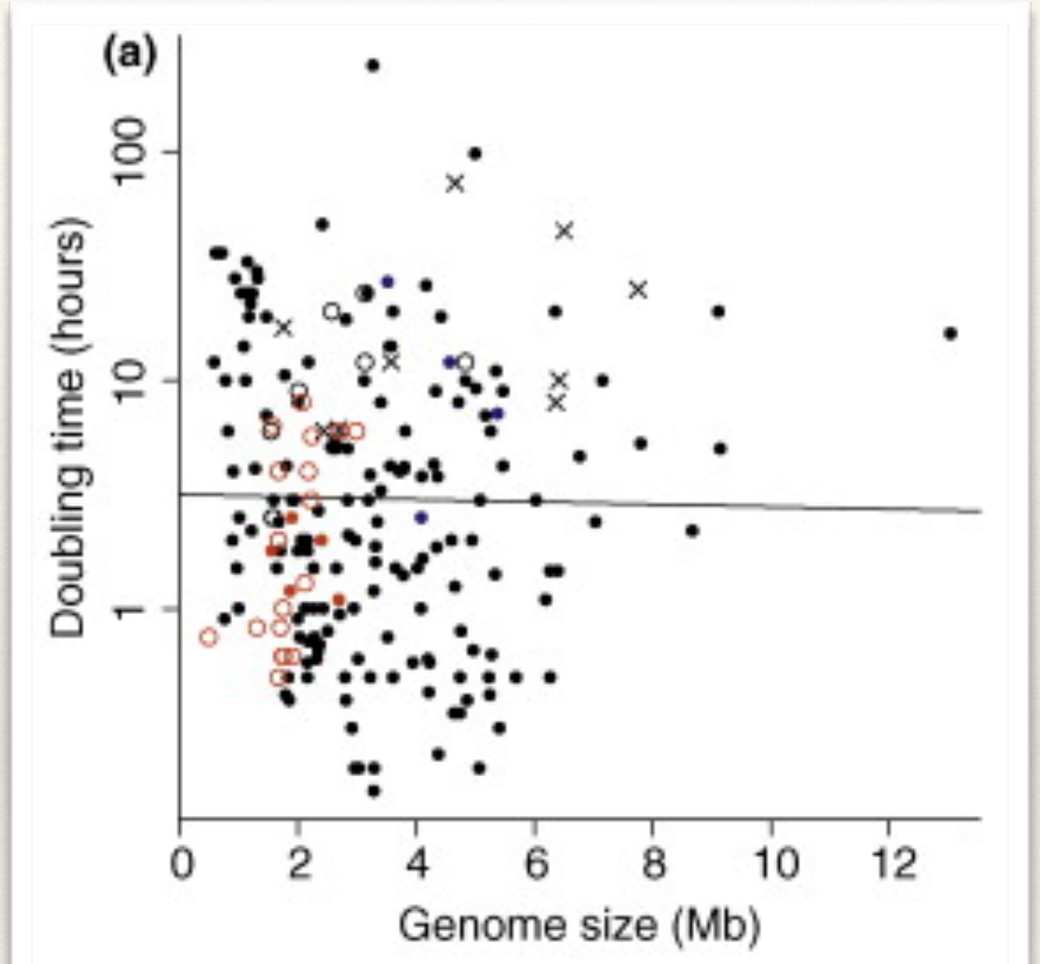
- Increased genome size
  - Pro: Multipartite genomes are larger on average
  - Con: Majority of largest genomes are not divided



Distribution of genome sizes for divided and non-divided genomes (diCenzo and Finan. 2017. MMBR. 81: e00019)

Why maintain a multipartite genome?

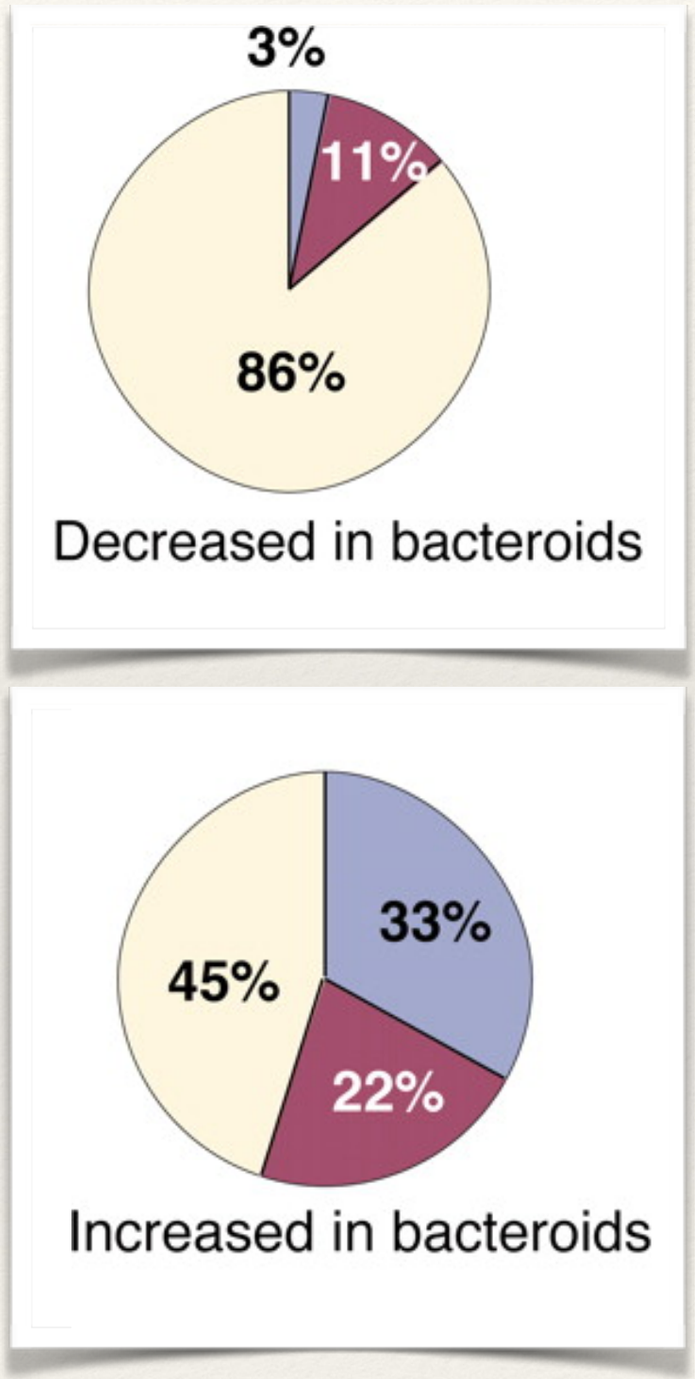
- Increased genome size
  - Pro: Multipartite genomes are larger on average
  - Con: Majority of largest genomes are not divided
- Increased growth rate
  - Pro: Some of the fastest growing species have a divided genome
  - Con: No correlation between genome size and growth rate



Relationship (or lack thereof) of genome size and growth rate (Vieira-Silva *et al*. 2010. Trends Ecol Evol. 25: 319)

Why maintain a multipartite genome?

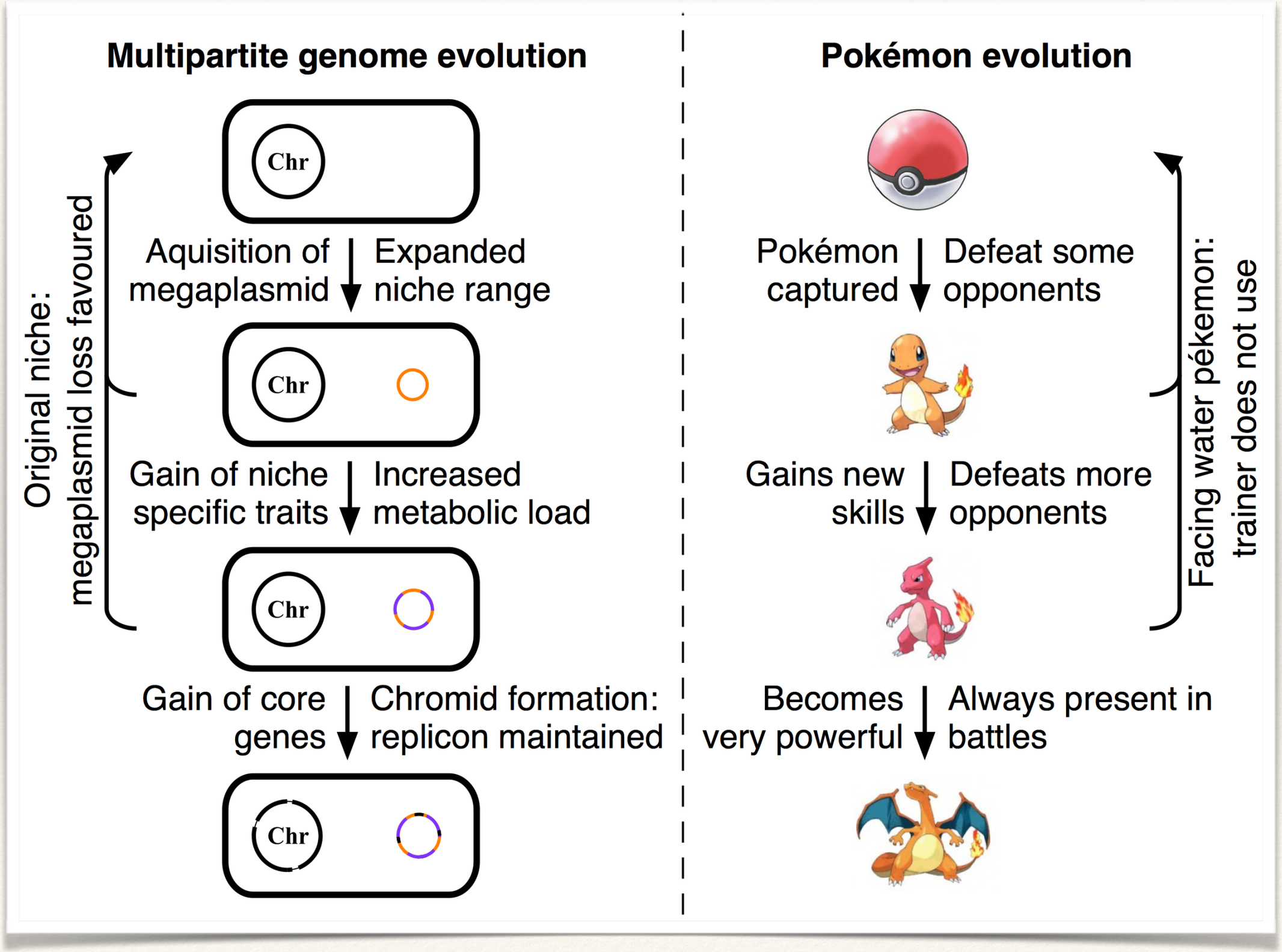
- Coordinated gene regulation
  - Pro: Replicon biases in transcriptional responses
  - Con: Unclear the order of causality
- Adaptation to new niches
  - Pro: Consistent with many features of the divided genomes (e.g. functions)
  - Con: Species without divided genomes occupy same niches



Differential expression in nodule versus free-living *S. meliloti*

(Barnett *et al*. 2004. PNAS. 101: 16636)

Adaptation to novel niches

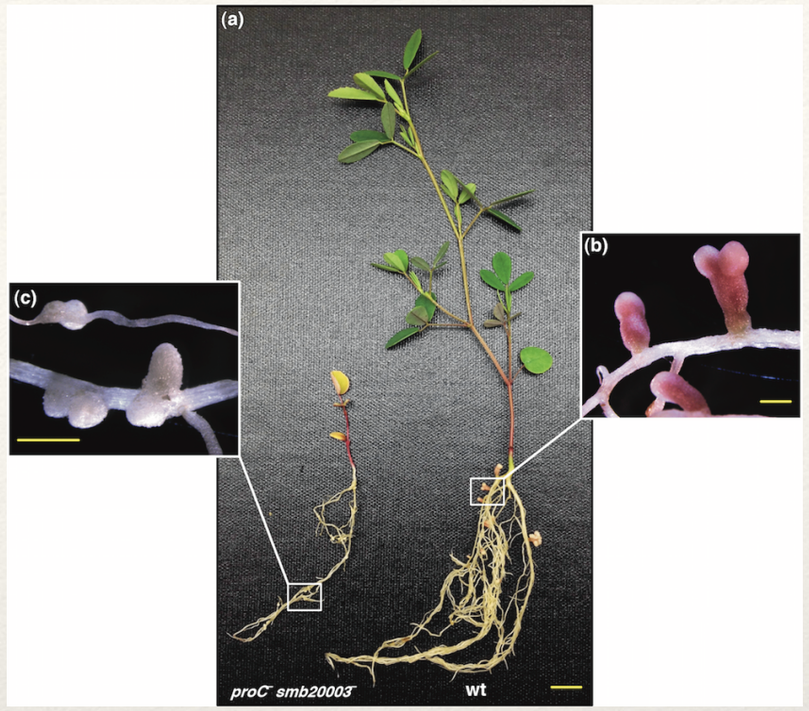


Outline

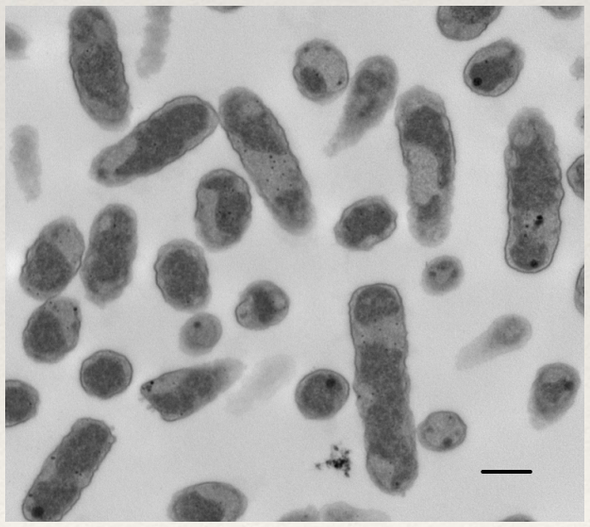
1. Genome organization and introduction to divided genomes
2. Characteristics of divided genomes
3. Formation of divided genomes
4. Possible advantages of divided genomes
5. *Sinorhizobium meliloti*: A case study

Sinorhizobium meliloti

- N_2-fixing symbiont of legumes such as alfalfa and sweet clover
- Complete genome sequence was published in 2001
- Large 6.7 Mb genome split into three replicons: 3.7 Mb (55%) chromosome, 1.7 Mb (25%) pSymB, 1.4 Mb (20%) pSymA



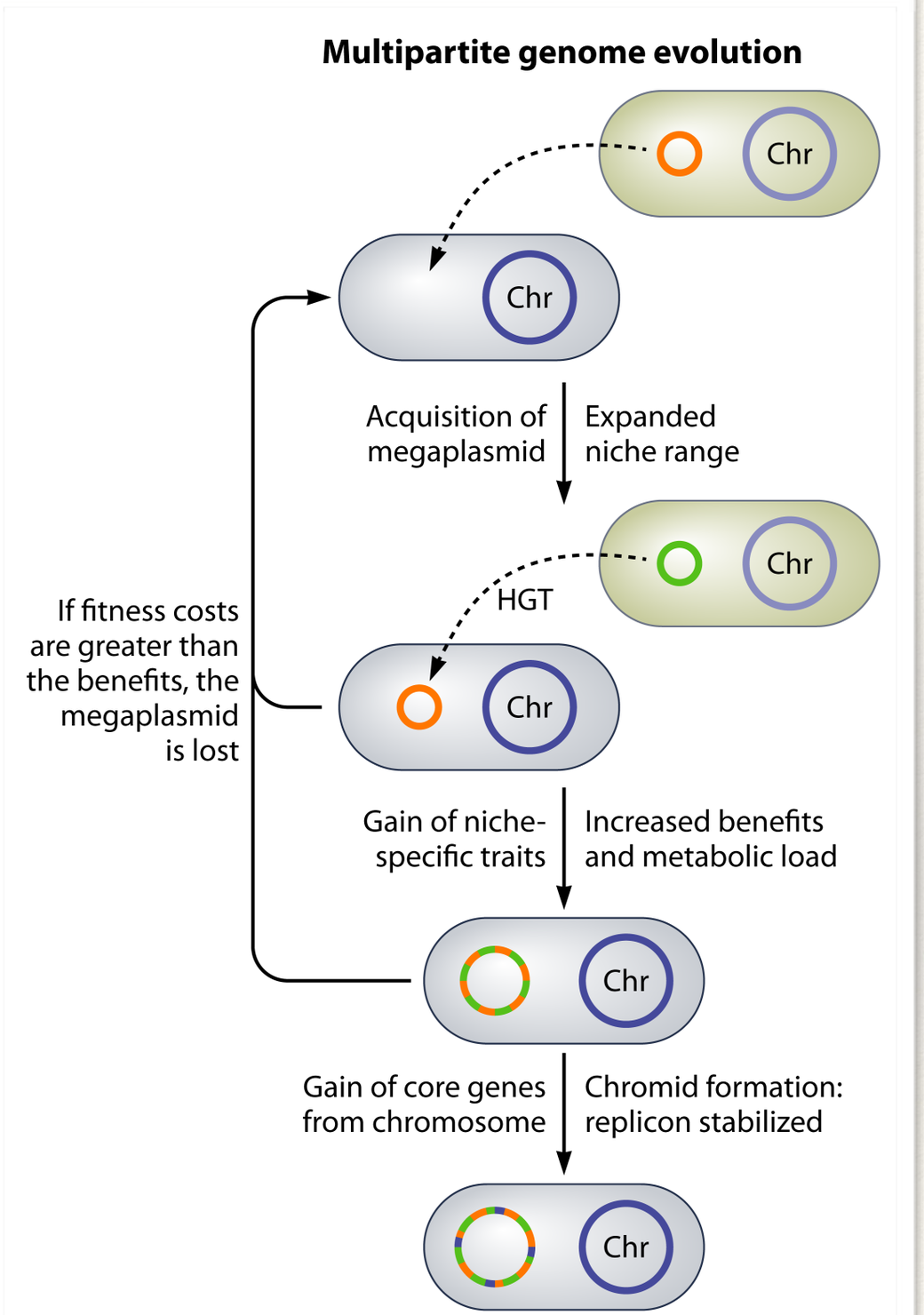
Sweet clover with wild type *S. meliloti* (right) or a proline auxotroph (left)



S. meliloti visualized with TEM

Images from:
diCenzo *et al*. 2015. Microbiology
diCenzo *et al*. 2017. J Bacteriol

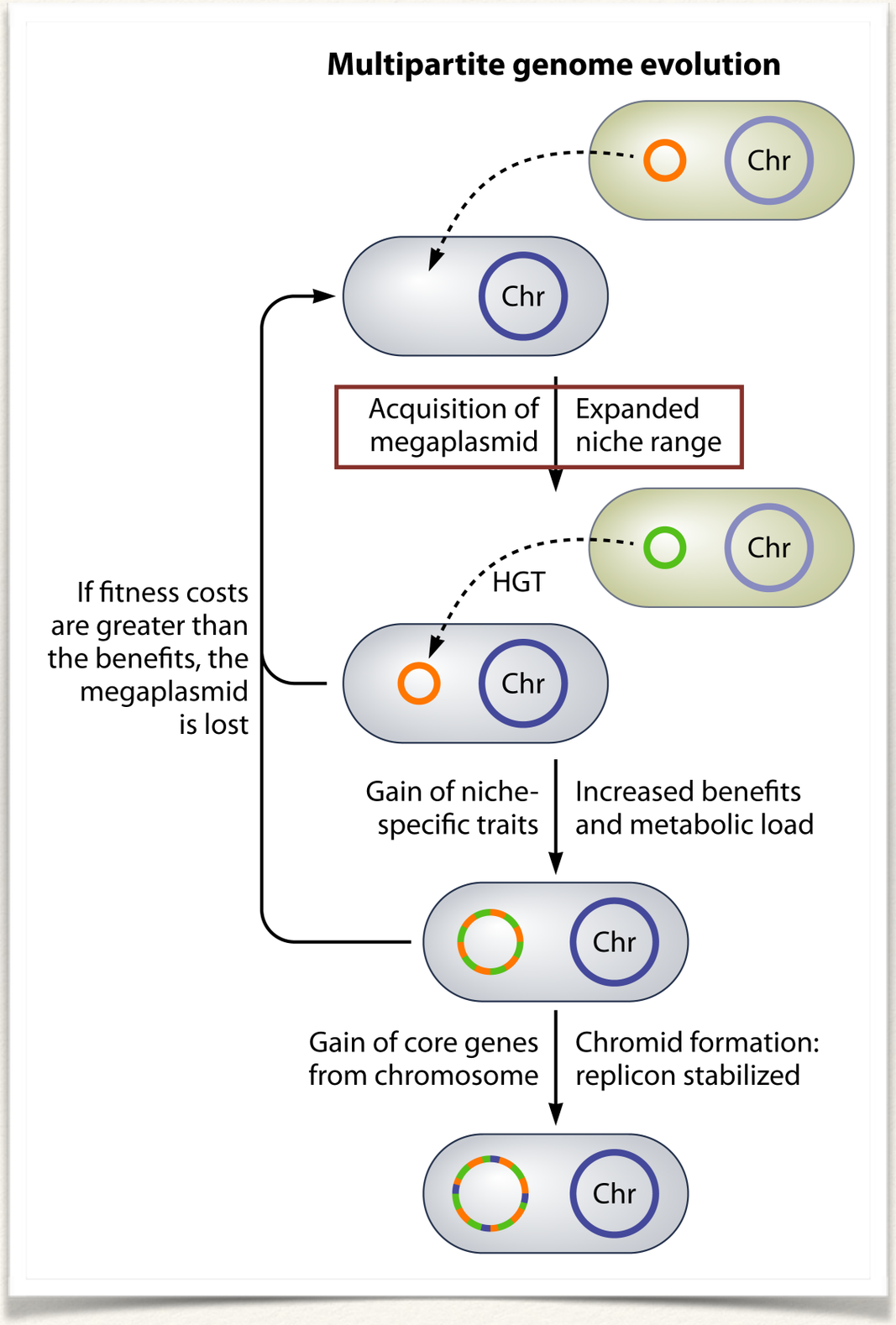
S. meliloti: a niche specialization case study



(diCenzo and Finan. 2017. MMBR. 81: e00019)

- Reminder: *S. meliloti* genome consists of a chromosome, pSymB chromid, pSymA megaplasmid

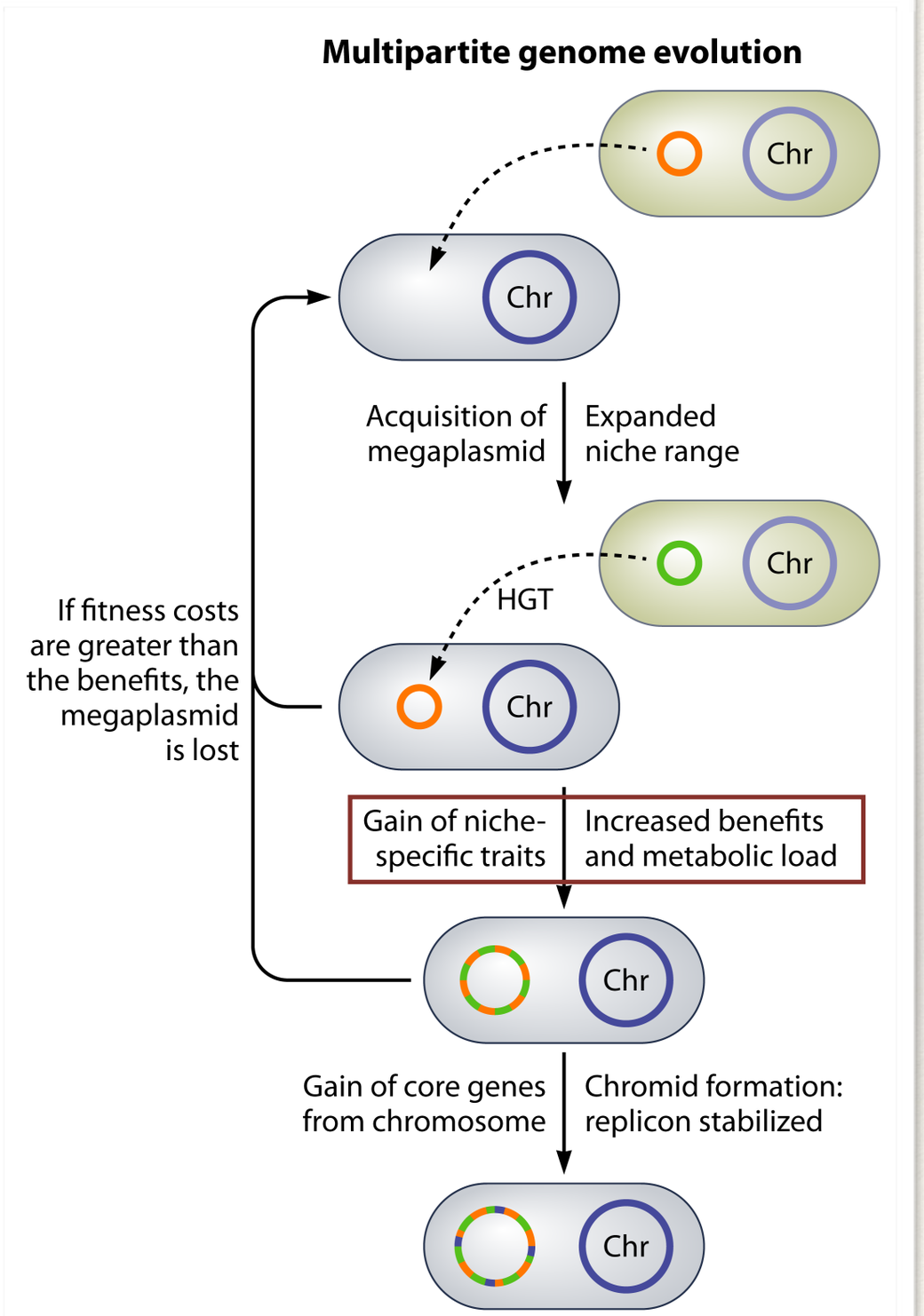
S. meliloti: a niche specialization case study



- pSymA is essential for nodule colonization
- pSymB contains genes important for rhizosphere colonization

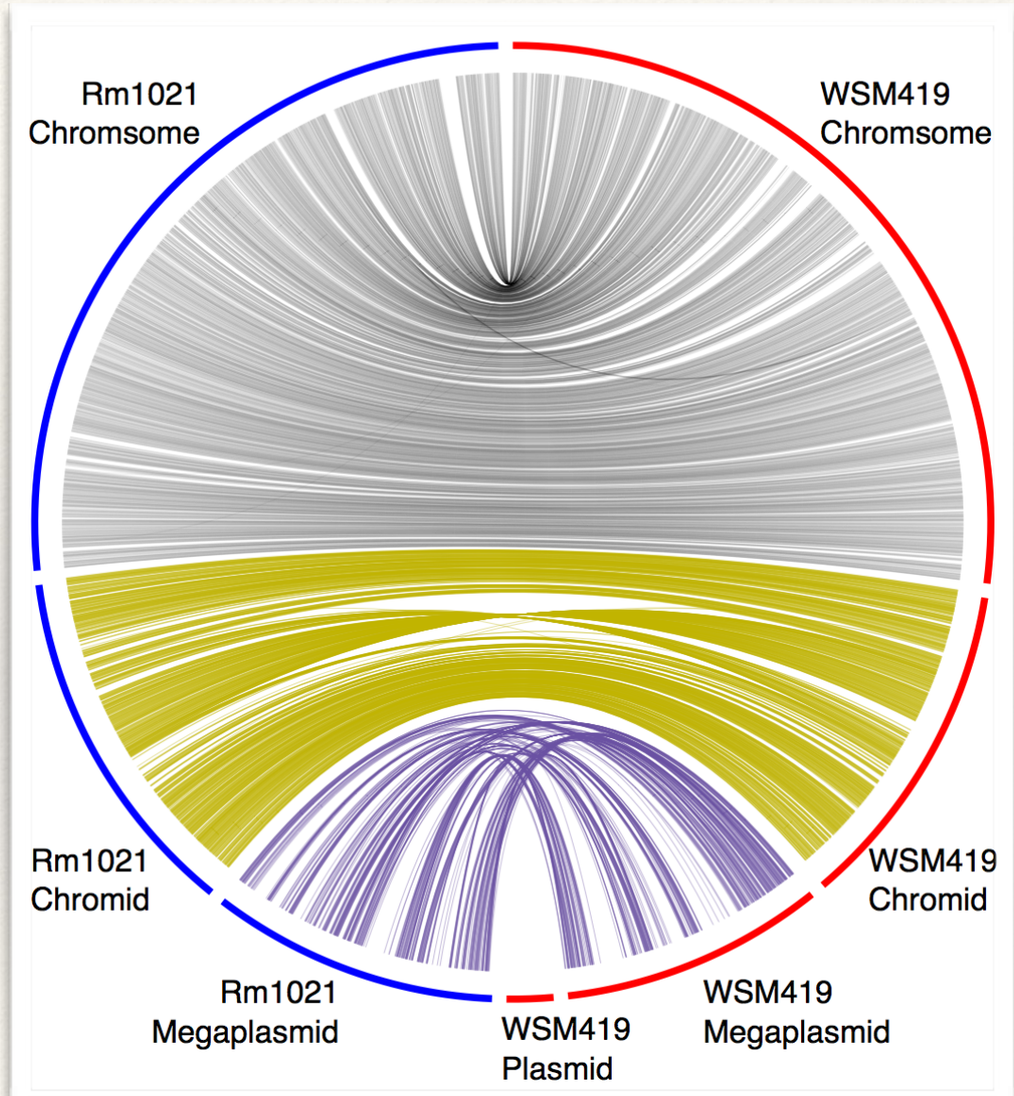
(diCenzo and Finan. 2017. MMBR. 81: e00019)

S. meliloti: a niche specialization case study



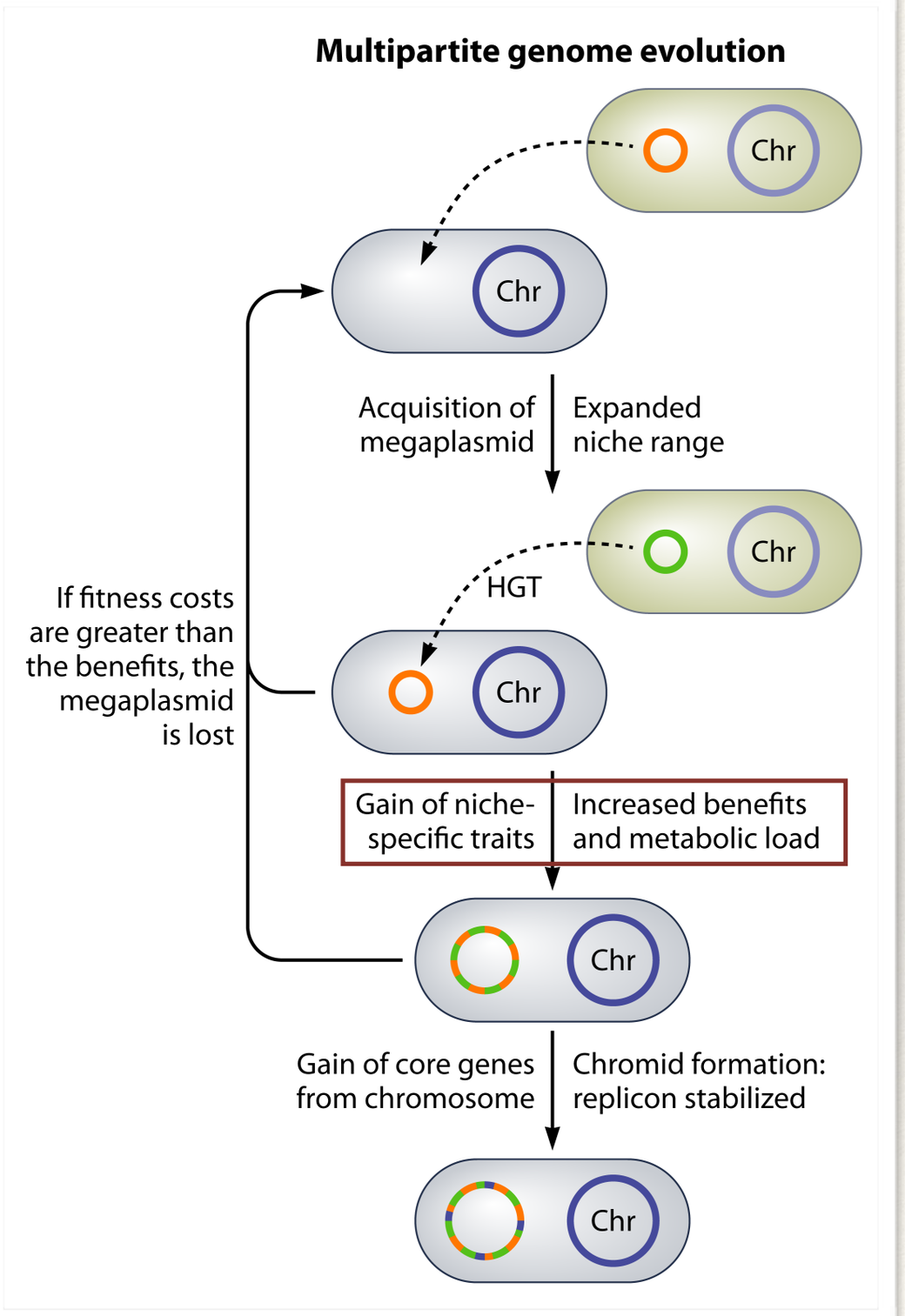
(diCenzo and Finan. 2017. MMBR. 81: e00019)

- pSymA formed through recent HGT, pSymB from ancient HGT



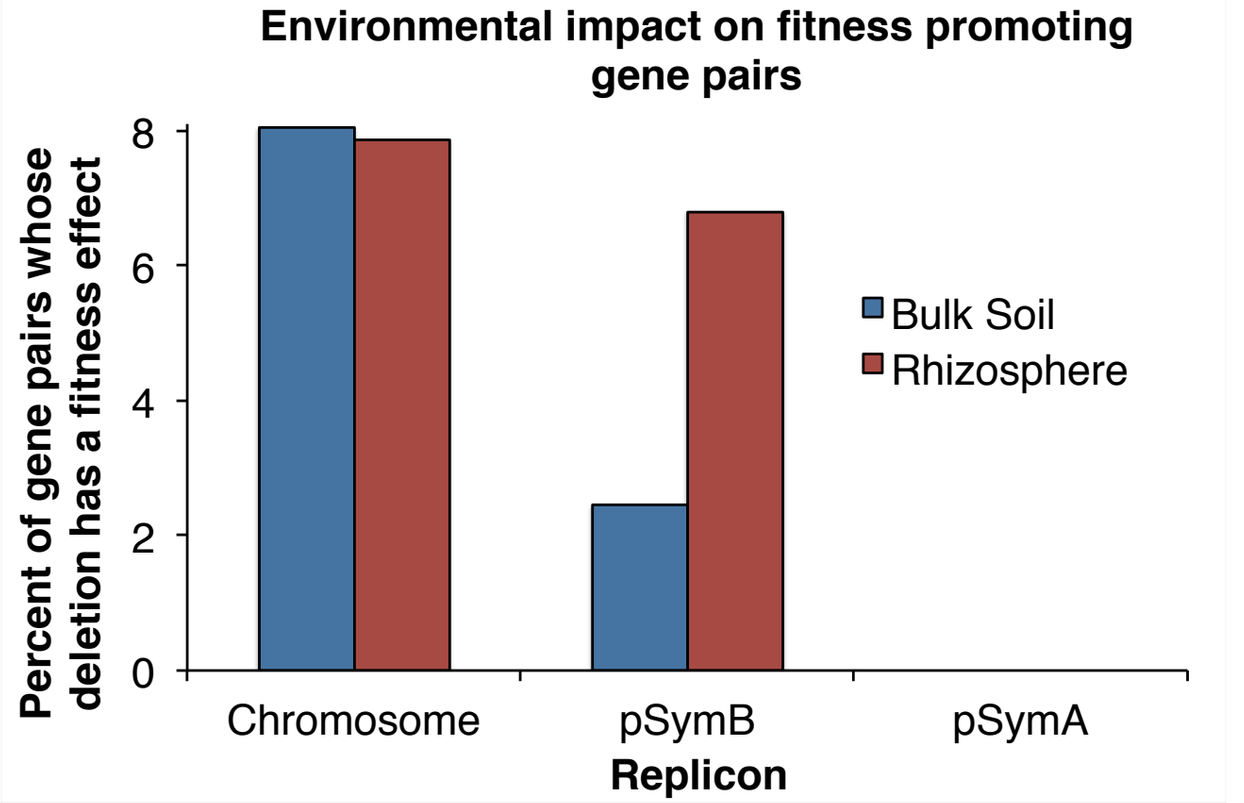
Comparing *S. meliloti* and *S. medicae* (closely related)

S. meliloti: a niche specialization case study



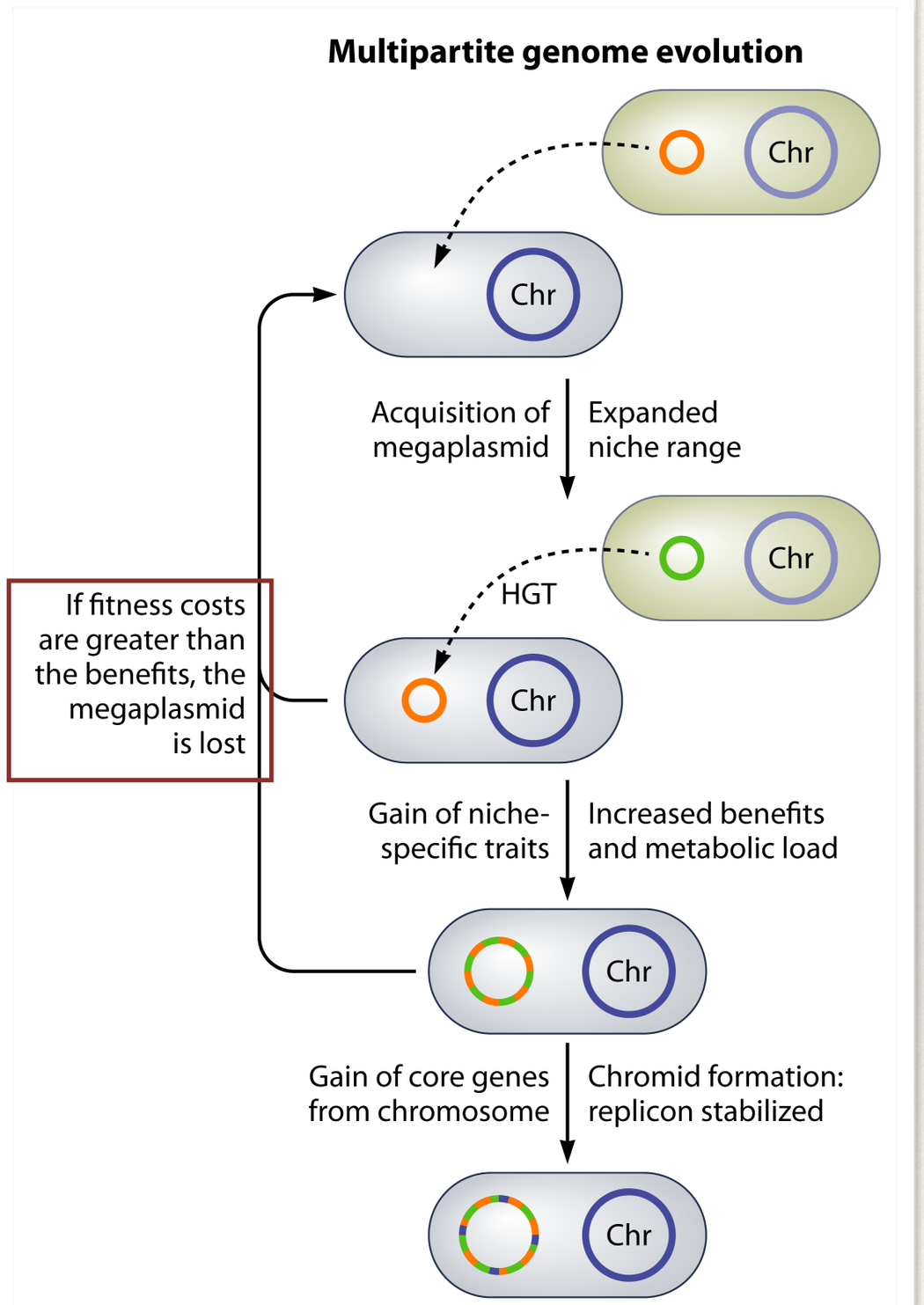
(diCenzo and Finan. 2017. MMBR. 81: e00019)

- Metabolic functions of pSymB are specialized for rhizosphere colonization



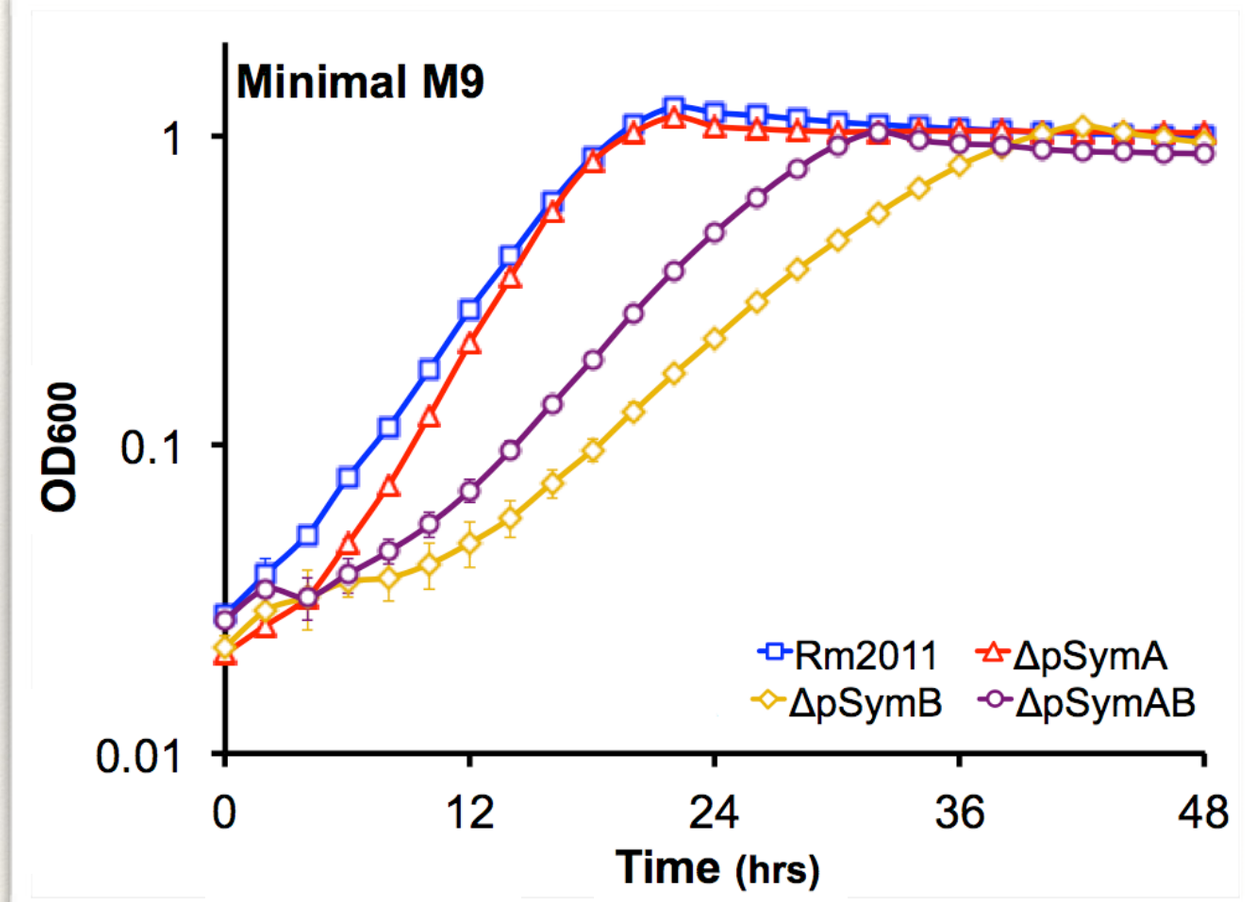
The number of chromosome or pSymB encoded genes predicted to contribute to growth in bulk soil or the rhizosphere based on *in silico* simulations (diCenzo *et al*. 2016. Nature Communications. 7: 12219)

S. meliloti: a niche specialization case study



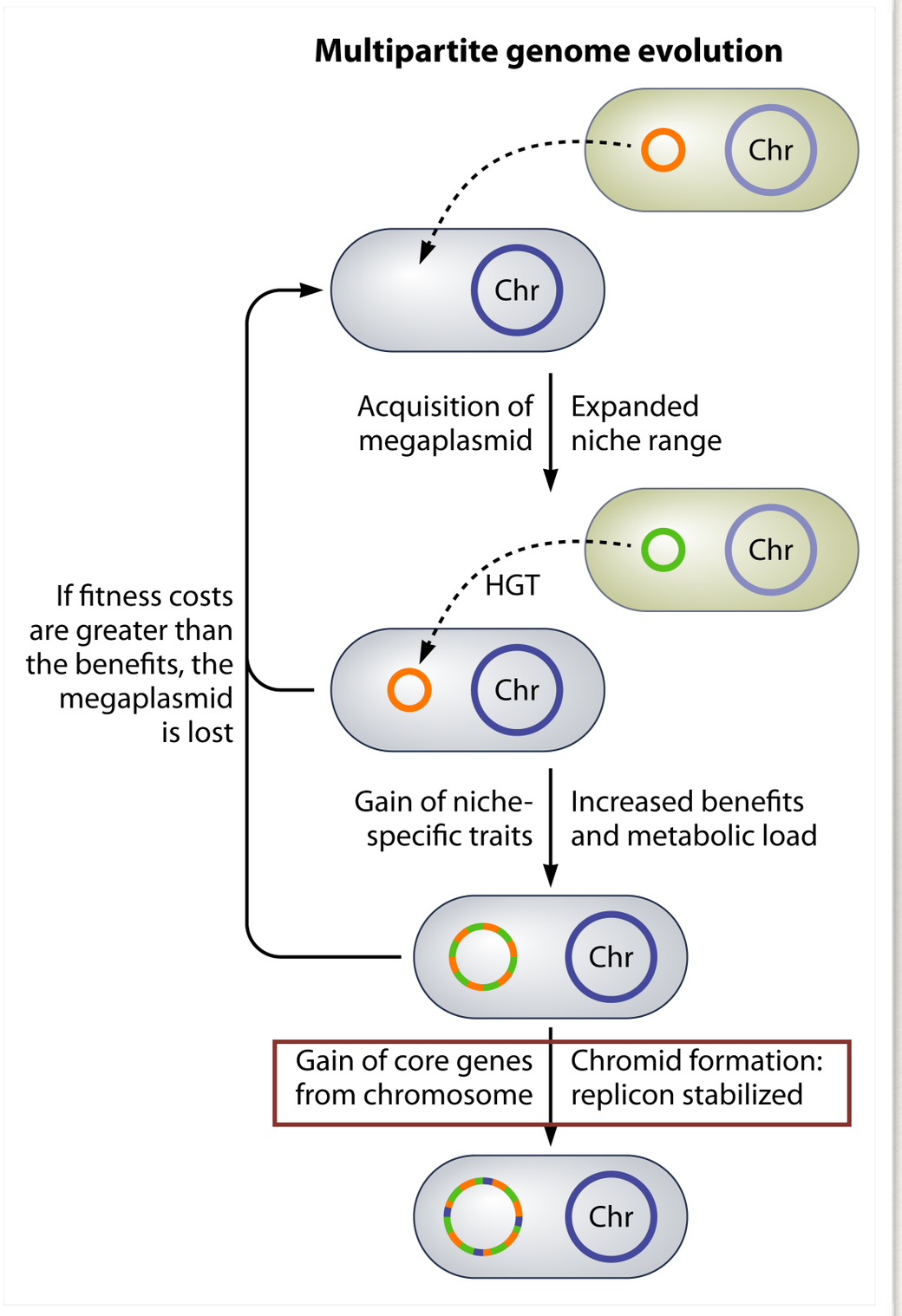
(diCenzo and Finan. 2017. MMBR. 81: e00019)

- *S. meliloti* may grow slightly faster in lab conditions when pSymA is removed



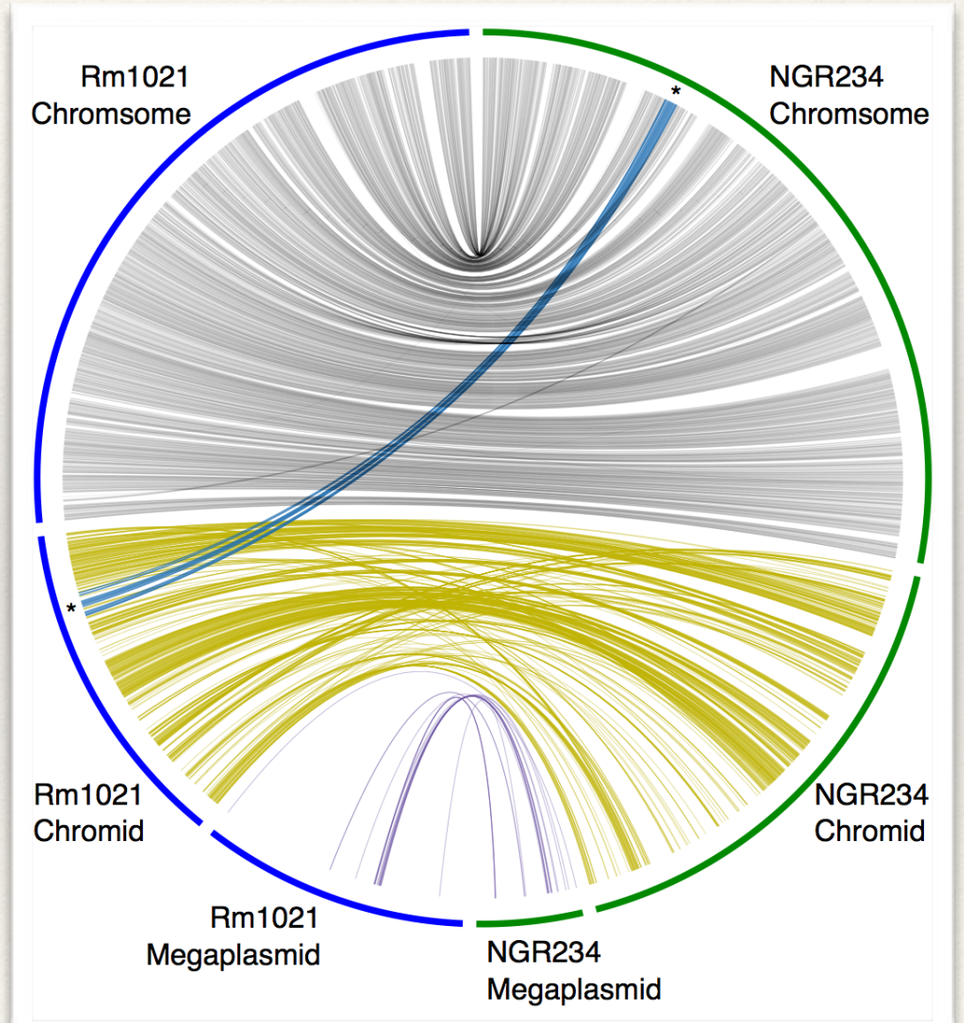
Growth rates of *S. meliloti* strains with or without pSymA and/or pSymB in minimal media (diCenzo *et al*. 2014. PLOS Genet. 10: e1004742)

S. meliloti: a niche specialization case study



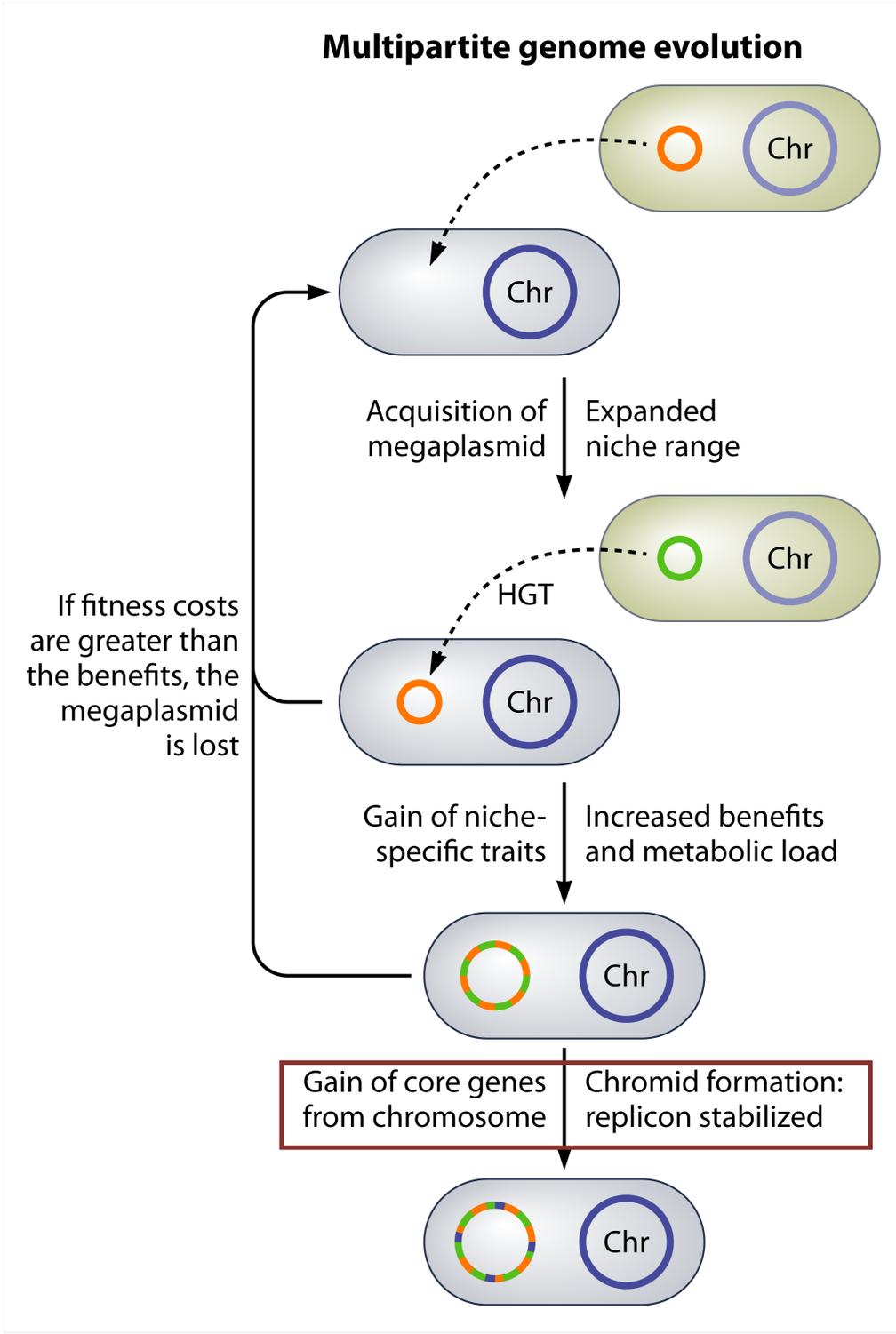
(diCenzo and Finan. 2017. MMBR. 81: e00019)

- A 69-kilobase region moved from the chromosome to pSymB, moving two essential genes



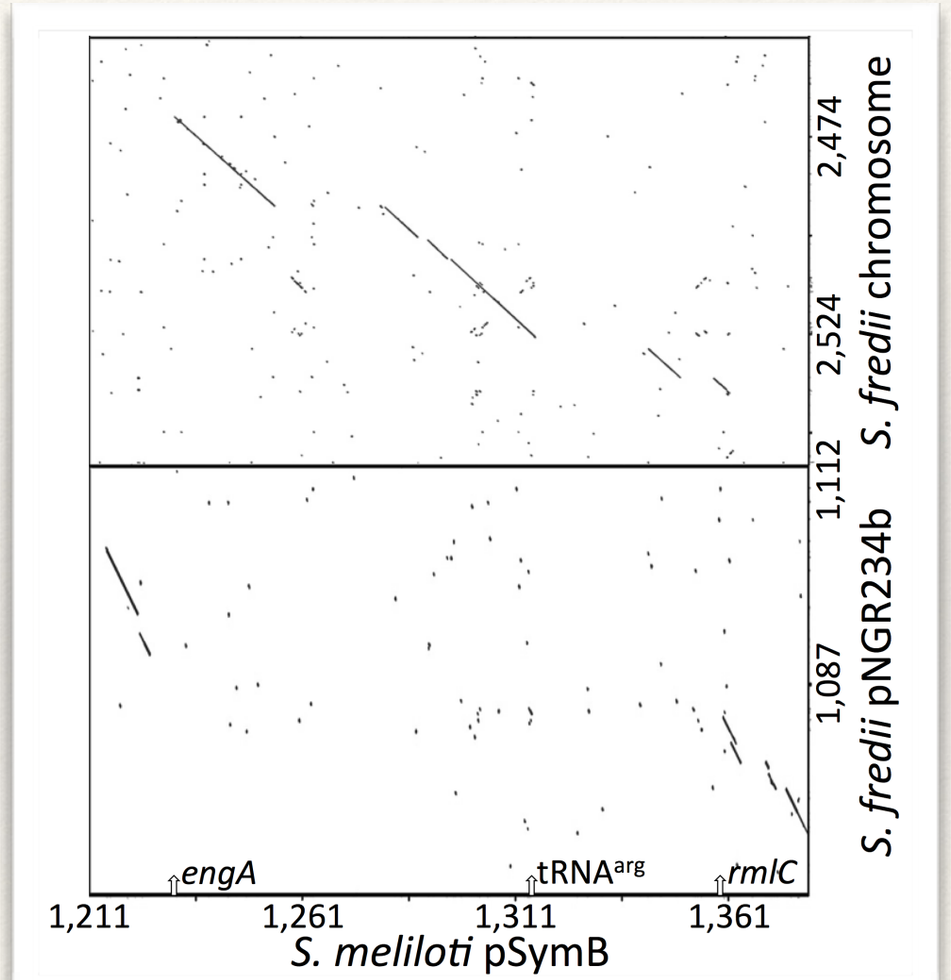
Comparing the *S. meliloti* and *S. fredii* genomes

S. meliloti: a niche specialization case study



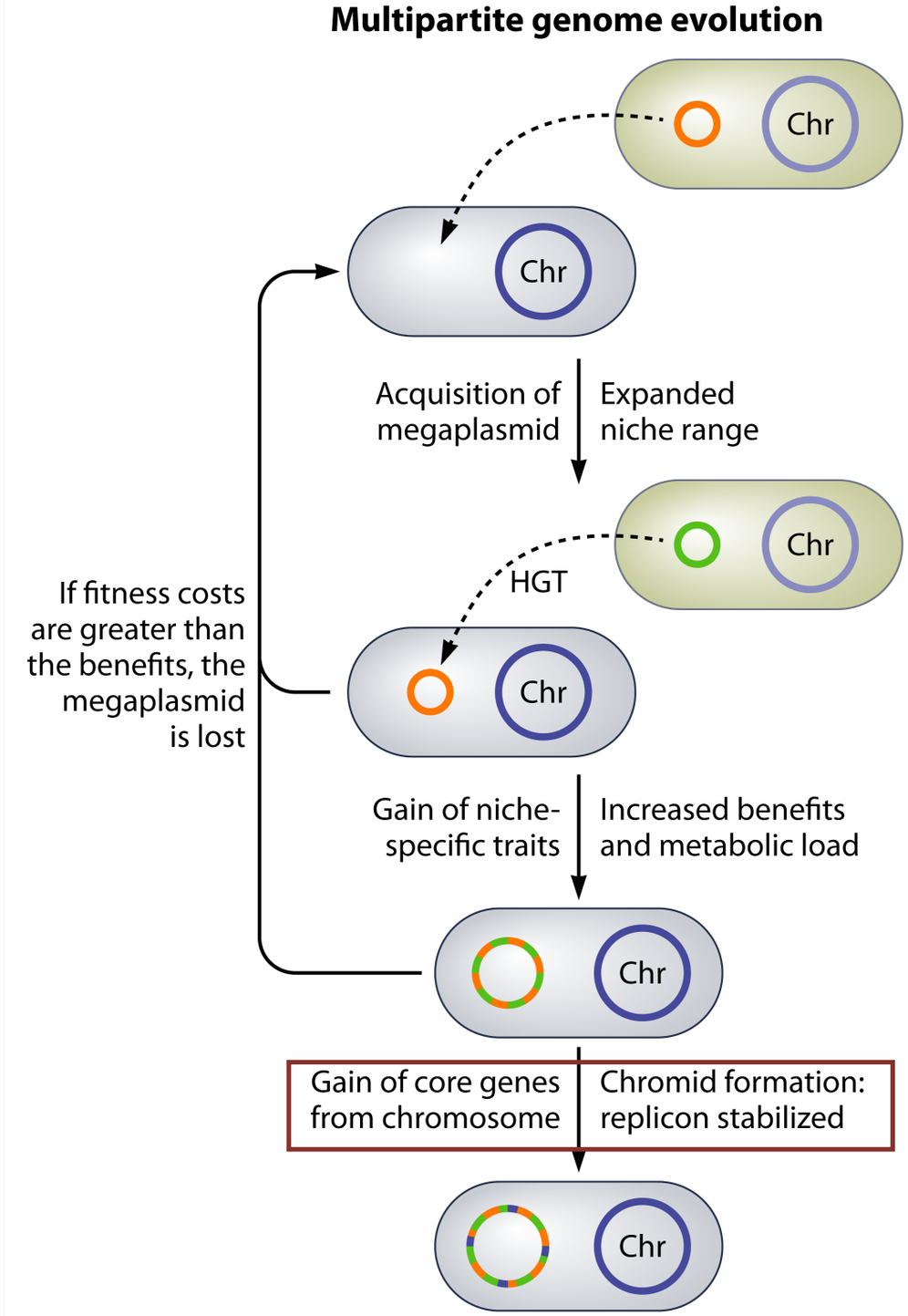
(diCenzo and Finan. 2017. MMBR. 81: e00019)

- A 69-kilobase region moved from the chromosome to pSymB, moving two essential genes



Dot plots showing the region transferred to pSymB (diCenzo *et al.* 2013. J Bacteriol. 195: 202)

S. meliloti: a niche specialization case study



(diCenzo and Finan. 2017. MMBR. 81: e00019)

- pSymB is more integrated into the cell transcriptionally and metabolically than pSymA

"Excuse me, sir,
would you mind getting the door for us?"