

2016

Bacterial Speciation: Genetic Sweeps in Bacterial Species

Frederick M. Cohan

Wesleyan University, fcohan@wesleyan.edu

Follow this and additional works at: <http://wescholar.wesleyan.edu/biolfacpub>

Recommended Citation

Cohan, Frederick M., "Bacterial Speciation: Genetic Sweeps in Bacterial Species" (2016). *Faculty Scholarship*. 187.
<http://wescholar.wesleyan.edu/biolfacpub/187>

This Article is brought to you for free and open access by the Biology at WesScholar. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of WesScholar. For more information, please contact ljohnson@wesleyan.edu.

8. Aso, Y., Sitaraman, D., Ichinose, T., Kaun, K.R., Vogt, K., Belliard-Guerin, G., Placais, P.Y., Robie, A.A., Yamagata, N., Schnaitmann, C., *et al.* (2014). Mushroom body output neurons encode valence and guide memory-based action selection in *Drosophila*. *eLife* 3, e04580.
9. Oswald, D., Felsenberg, J., Talbot, C.B., Das, G., Perisse, E., Huetteroth, W., and Waddell, S. (2015). Activity of defined mushroom body output neurons underlies learned olfactory behavior in *Drosophila*. *Neuron* 86, 417–427.
10. Sejourne, J., Placais, P.Y., Aso, Y., Siwanowicz, I., Trannoy, S., Thoma, V., Tedjakumala, S.R., Rubin, G.M., Tchenio, P., Ito, K., *et al.* (2011). Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in *Drosophila*. *Nat. Neurosci.* 14, 903–910.
11. Oswald, D., and Waddell, S. (2015). Olfactory learning skews mushroom body output pathways to steer behavioral choice in *Drosophila*. *Curr. Opin. Neurobiol.* 35, 178–184.
12. Hige, T., Aso, Y., Modi, M.N., Rubin, G.M., and Turner, G.C. (2015). Heterosynaptic plasticity underlies aversive olfactory learning in *Drosophila*. *Neuron* 88, 985–998.
13. Cohn, R., Morante, I., and Ruta, V. (2015). Coordinated and compartmentalized neuromodulation shapes sensory processing in *Drosophila*. *Cell* 163, 1742–1755.
14. Lewis, L.P., Siju, K.P., Aso, Y., Friedrich, A.B., Bulteel, A.J., Rubin, G.M., and Grunwald Kadow, I.C. (2015). A higher brain circuit for immediate integration of conflicting sensory information in *Drosophila*. *Curr. Biol.* 25, 2203–2214.
15. Keleman, K., Vrontou, E., Krutner, S., Yu, J.Y., Kurtovic-Kozaric, A., and Dickson, B.J. (2012). Dopamine neurons modulate pheromone responses in *Drosophila* courtship learning. *Nature* 489, 145–149.
16. Dreosti, E., Odermatt, B., Dorostkar, M.M., and Lagnado, L. (2009). A genetically encoded reporter of synaptic activity in vivo. *Nat. Methods* 6, 883–889.
17. Aso, Y., Hattori, D., Yu, Y., Johnston, R.M., Iyer, N.A., Ngo, T.T., Dionne, H., Abbott, L.F., Axel, R., Tanimoto, H., *et al.* (2014). The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* 3, e04577.
18. Kim, Y.C., Lee, H.G., and Han, K.A. (2007). D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in *Drosophila*. *J. Neurosci.* 27, 7640–7647.
19. Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466, 457–462.
20. Martin, J.R., Ernst, R., and Heisenberg, M. (1998). Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learn. Mem.* 5, 179–191.

Bacterial Speciation: Genetic Sweeps in Bacterial Species

Frederick M. Cohan

Department of Biology, Wesleyan University, Middletown, CT 06459, USA

Correspondence: fcohan@wesleyan.edu

<http://dx.doi.org/10.1016/j.cub.2015.10.022>

One theory of bacterial speciation states that bacterial and animal species share the property of cohesion, meaning that diversity within a species is constrained. A new study provides direct evidence that genome-wide sweeps can limit diversity within bacterial species.

“Anything found to be true of *E. coli* must also be true of elephants!” While Jacques Monod was thinking about the unity of biochemistry, evolutionary biologists and ecologists have also sought unifying principles across all of life. Accordingly, some evolutionary ecologists have proposed universal principles for the origin and nature of species [1–3]. One particularly contentious claim is that across all of life, species are each ‘cohesive’, in the sense that the diversity within each species is constrained by some force [1,2]. Bacterial species may be subject to a special force of cohesion, owing to their low frequency of recombination [4]. This force is periodic selection, where natural selection favoring any adaptive mutation purges the genetic diversity, genome-wide, within an

ecologically homogeneous species or ecotype (Figure 1A,B) [5]. However, whether periodic selection occurs widely in nature or is even possible has been debated with a passion unusual for the field of microbiology [6–8]. Contributing to the uncertainty is that genome-wide sweeps were never observed in nature, that is, until the recent work of Bendall *et al.* [9] — using a high-throughput, metagenomic approach to survey a bacterial community over time, these authors charted a genome-wide sweep within one natural bacterial population.

In the days before metagenomics, it never seemed a good bet to search for a periodic selection event by focusing on your favorite organism — this process was thought too infrequent to be observed in one population during

one grant-funding period. However, what is unlikely to be seen within a single focus population becomes eminently palpable *within some population* when you can observe an entire community simultaneously. The magic of metagenomics!

Bendall *et al.* [9] surveyed the bacteria of Trout Bog Lake, Wisconsin, over eight years. From each small volume of lake water sampled, they lysed the cells and produced short DNA sequences representing the community’s diversity. Short sequences were assembled into longer sequences by joining sequences (from different organisms) that were homologous and less than 2% divergent from one another. The researchers thereby clustered the organisms into groups of close relatives, and at the same

time produced a contiguous sequence representing most of the genome for a given cluster. For each cluster of organisms demarcated by 2% divergence, they estimated the sequence diversity in each genome region.

In the case of the cluster *Chlorobium*-111, the assembled closely-related organisms began highly variable in 2005 at most gene loci of the genome (up to the 2% divergence cap), but by 2013 the genome had lost its diversity at nearly all loci. Bendall *et al.* [9] interpreted this as a genome-wide sweep caused by selection within *Chlorobium*-111. The authors also found indirect evidence of genome-wide sweeps within four other clusters, although not as cleanly as in *Chlorobium*-111. These clusters had a particularly low level of variation genome-wide over the entire eight years; the authors interpreted these cases as evidence for genome-wide sweeps of diversity occurring *before their study began*.

What do these genome-wide sweeps say about *Chlorobium*-111 and the other four clusters? The most obvious interpretation, noted by Bendall *et al.* [9], is that recombination was rare enough in these clusters to sustain periodic selection (Figure 1A,B). That is, a single, better-adapted genotype from one cluster was able to outcompete all other organisms within the cluster, and recombination was rare enough for the whole lucky genome to rise to near 100%. We should also explicitly infer one further point about *Chlorobium*-111 and the other four clusters: that these clusters must have each consisted of a single, ecologically homogeneous population (or ecotype), a point noted by the authors. If *Chlorobium*-111 and the others had instead each consisted of many ecotypes, the rise of a better-adapted genotype within just one ecotype would have left much of the cluster's diversity untouched [2]. So, we may conclude that the origination of ecotypes (or speciation [2]) was slow enough in these bacteria for organisms as divergent as 2% from one another to be ecologically homogeneous.

Why would diversification in *Chlorobium* and the other clusters be so slow? For *Chlorobium*, one possibility emerges from its photoautotrophic habit, where organisms engage in photosynthesis and do not consume organic compounds. A photoautotroph's

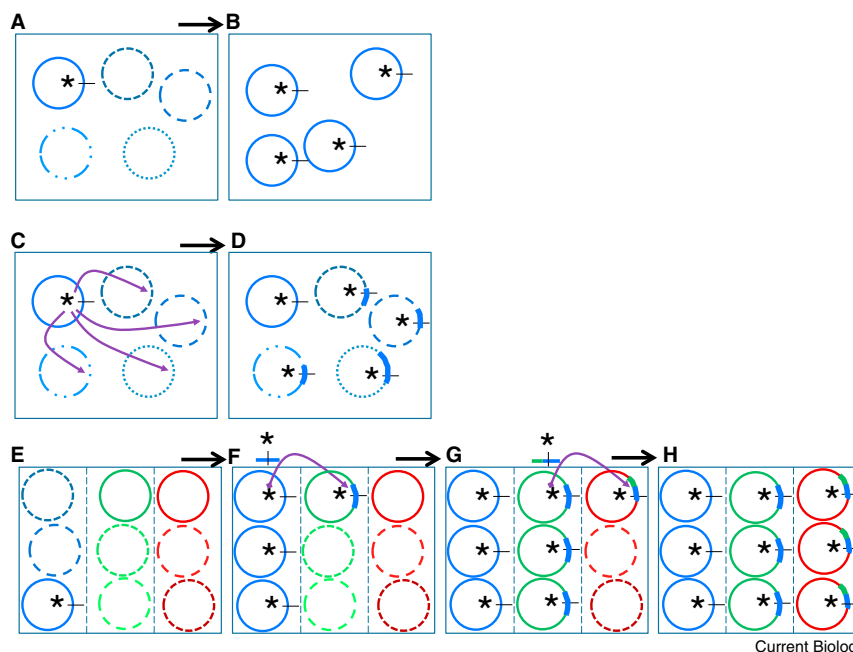


Figure 1. Sweeps by garage broom and whisk broom within a metagenome cluster.

Each row of panels represents a different model of sweep within a metagenome cluster. (A,B) A genome-wide sweep where the metagenome cluster is populated by only a single ecotype, and where recombination is rare enough to allow a genome-wide sweep within an ecotype [2]. (C,D) A narrow sweep (homogenizing only the chromosomal region near the adaptive mutation), where again the metagenome cluster is populated by a single ecotype, but here recombination (indicated by purple arrows) is frequent enough to prevent a genome-wide sweep within an ecotype (model favored by Bendall *et al.* [9]). (E–H) A narrow sweep where the metagenome cluster is populated by many ecotypes (in this case, three), and recombination is rare enough to allow genome-wide sweeps within an ecotype but frequent enough to allow an adaptive mutation to recombine (it need happen only once!) between one ecotype and another [8]. In each row, the wide horizontal arrows represent the course of time. In each panel the rectangle represents one metagenome cluster, and each circle represents a single organism. The asterisk represents an adaptive mutation, which allows its carrier to outcompete other organisms in the same ecotype, but not organisms from other ecotypes. In panels A through D, the metagenome cluster is ecologically homogeneous, and in panels E through H, the metagenome cluster is ecologically heterogeneous and represents three ecotypes, separated by the vertical dashed lines; the different ecotypes are coded by blue, green, and red. The sequence diversity within an ecotype is represented by different shades of the ecotype color and by different styles of line (dotted, dashed, solid). In the case of low recombination rates (A,B and E–H), the adaptive mutation causes a genome-wide sweep within the ecotype containing the mutation. In panels E through H, the adaptive mutation is potentially beneficial in different ecotypes and can transfer on a short chromosomal segment to another ecotype, where it precipitates a new genome-wide sweep within its new ecotype.

disdain for environmental organic compounds would limit its metabolic plasticity, such that opportunities for ecological diversification may be infrequent, as found in hot spring *Synechococcus* [10]. The other clusters appearing to undergo genome-wide sweeps may also have had few opportunities for diversification. *Gallionella*-3415 and *Desulfobulbus*-2922, for example, are both from genera that are very limited in the carbon they consume, and so may speciate infrequently. Another low-diversity cluster, TM7-1225, was classified to a phylum whose metabolism is not well

understood but appears to be limited in its carbon sources [11]. The fourth low-diversity cluster, *Rickettsia*-755, was classified to a genus of obligate intracellular parasites. While not limited in its carbon sources, the lifestyle of this genus may possibly result in infrequent speciation. We may hypothesize that low metabolic plasticity predicts slow speciation and the possibility of genome-wide sweeps within sequence clusters with up to 2% nucleotide diversity.

While the genomic diversity within *Chlorobium*-111 was swept nearly exhaustively, as if with a garage broom, a

whisk broom appears to have swept six other clusters in a more targeted fashion. Each of these clusters lacked diversity in only one short chromosomal region, but elsewhere in the genome they were highly divergent. Bendall *et al.* [9] inferred that a sweep occurred in a single chromosomal region within each cluster before their survey began. They ascribed the narrowness of these sweeps to high recombination rates within populations. Like narrow sweeps in the highly recombining animals, they argued that an adaptive gene could recombine into many genetic backgrounds before the gene reached 100% frequency (Figure 1C,D).

This idea has been floating in the bacterial literature in recent years [6,12]. However, bacterial recombination appears rare enough in every taxon studied [4] to sustain genome-wide sweeps [8,13,14]. Moreover, Bendall *et al.* [9] provided no evidence for differences in recombination rates among the clusters incurring genome-wide versus narrow sweeps. What, then, might distinguish clusters swept genome-wide from those swept only at a single chromosomal region?

One possibility is that each cluster of organisms undergoing a single-gene sweep was ecologically heterogeneous, consisting of many ecotypes. Decades ago, we developed the ‘Adapt Globally Act Locally’ model [15], in which a generally adaptive mutation — adaptive for many ecotypes — causes a genome-wide sweep within its own ecotype, and then a small chromosomal region containing the adaptive mutation recombines into another ecotype and causes a genome-wide sweep there (Figure 1E–H). The mutation may then recombine into a series of ecotypes and cause sweeps in each one, so long as the adaptive mutation is beneficial within each ecotype [13]. Thus, all the ecologically diverse strains within a cluster could become nearly identical in the chromosomal region flanking the adaptive mutation but remain otherwise heterogeneous [8], yielding a narrow sweep.

Bendall *et al.* [9] dismissed the Adapt Globally Act Locally explanation as unparsimonious. They deemed this model unlikely because inter-ecotype recombination would have to be frequent enough to allow adaptive mutations to pass between ecotypes, while intra-

ecotype recombination would have to be rare enough to allow genome-wide sweeps within ecotypes. However, successful passage of an adaptive gene between two ecotypes requires only a single recombination event (because natural selection fixes the adaptive gene in its new ecotype), while prevention of a genome-wide sweep within an ecotype requires recurrent recombination.

The Adapt Globally Act Locally explanation requires clusters undergoing single-gene sweeps to be ecologically heterogeneous, as a result of rapid diversification. For example, these clusters could be generalist heterotrophs, which can change their ecology by simply adding another carbon resource to their consumption repertoire [5]. As predicted, several clusters showing single-gene sweeps in Bendall *et al.* were generalist heterotrophs: *Polynucleobacter*-239 [16], *Holophagales*-254 [17], and two clusters of *Opitutae* [18]. Two clusters showing single-gene sweeps were not generalist heterotrophs, although they may nevertheless undergo rapid diversification. Clusters *Methylobacter*-2062 and *Methylophilaceae*-913 were both classified to taxa that are very limited in carbon sources, but extremely close relatives in these taxa have elsewhere been shown to differentiate by depth [19] and by utilizing different C1 molecules [20], respectively. We may hypothesize that ecologically heterogeneous clusters are most likely to incur single-gene sweeps, while ecologically homogeneous clusters are most likely to incur whole-genome sweeps.

With thanks to metagenomics and Bendall *et al.* [9], we can now see cohesion by genome-wide sweeps within broadly defined (with up to 2% diversity), ecologically homogeneous clusters that contain a single ecotype. However, the low recombination rates in bacteria predict genome-wide sweeps in all ecotypes, not just those in broadly defined, ecologically homogeneous clusters [14]. Can we find direct evidence for cohesion by periodic selection within the phylogenetically tiny ecotypes within ecologically diverse clusters? Unfortunately, this is beyond the current resolution of metagenomics. It is too soon to know for sure whether all life’s species share the principle of cohesion [2].

REFERENCES

1. Templeton, A. (1989). The meaning of species and speciation: a genetic perspective. In *Speciation and its Consequences*, D. Otte, and J. Endler, eds. (Sunderland MA: Sinauer Assoc), pp. 3–27.
2. Cohan, F.M. (2011). Are species cohesive?—A view from bacteriology. In *Bacterial Population Genetics: A Tribute to Thomas S. Whittam*, S. Walk, and P. Feng, eds. (Washington, DC: American Society for Microbiology Press), pp. 43–65.
3. Cadillo-Quiroz, H., Didelot, X., Held, N.L., Herrera, A., Darling, A., Reno, M.L., Krause, D.J., and Whitaker, R.J. (2012). Patterns of gene flow define species of thermophilic Archaea. *PLoS Biol.* 10, e1001265.
4. Vos, M., and Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *ISME J.* 3, 199–208.
5. Kopac, S., Wang, Z., Wiedenbeck, J., Sherry, J., Wu, M., and Cohan, F.M. (2014). Genomic heterogeneity and ecological speciation within one subspecies of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 80, 4842–4853.
6. Shapiro, B.J., and Polz, M.F. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol.* 22, 235–247.
7. Doolittle, W.F., and Zhaxybayeva, O. (2009). On the origin of prokaryotic species. *Genome Res.* 19, 744–756.
8. Cohan, F.M. (2016). Prokaryotic species concepts. In *Encyclopedia of Evolutionary Biology*, R.M. Kliman, ed. (Oxford: Elsevier).
9. Bendall, M.L., Stevens, S.L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W., Martin, J., Pati, A., Bushnell, B., *et al.* (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations revealed by time-series metagenomics. *ISME J.* <http://dx.doi.org/10.1038/ismej.2015.241>.
10. Becraft, E.D., Wood, J.M., Rusch, D.B., Köhl, M., Jensen, S.I., Bryant, D.A., Roberts, D.W., Cohan, F.M., and Ward, D.M. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. *Front. Microbiol.* 6, 590.
11. Kantor, R.S., Wrighton, K.C., Handley, K.M., Sharon, I., Hug, L.A., Castelle, C.J., Thomas, B.C., and Banfield, J.F. (2013). Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *MBio.* 4, e00708–00713.
12. Papke, R.T., Zhaxybayeva, O., Feil, E.J., Sommerfeld, K., Muise, D., and Doolittle, W.F. (2007). Searching for species in haloarchaea. *Proc. Natl. Acad. Sci. USA* 104, 14092–14097.
13. Cohan, F.M. (2005). Periodic selection and ecological diversity in bacteria. In *Selective Sweep*, D. Nurminsky, ed. (Georgetown, Texas: Landes Bioscience), pp. 78–93.

14. Wiedenbeck, J., and Cohan, F.M. (2011). Origins of bacterial diversity through horizontal gene transfer and adaptation to new ecological niches. *FEMS Microbiol. Rev.* 35, 957–976.
15. Majewski, J., and Cohan, F.M. (1999). Adapt globally, act locally: the effect of selective sweeps on bacterial sequence diversity. *Genetics* 152, 1459–1474.
16. Jezbera, J., Jezberova, J., Brandt, U., and Hahn, M.W. (2011). Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* results from ecological diversification. *Environ. Microbiol.* 73, 922–931.
17. Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu, M., Xie, G., Haft, D.H., Sait, M., Badger, J., *et al.* (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl. Environ. Microbiol.* 75, 2046–2056.
18. Choo, Y.J., Lee, K., Song, J., and Cho, J.C. (2007). *Puniceicoccus vermicola* gen. nov., sp. nov., a novel marine bacterium, and description of *Puniceicoccaceae* fam. nov., *Puniceicoccales* ord. nov., *Opiritaceae* fam. nov., *Opiritales* ord. nov. and *Opiritae* classis nov. in the phylum ‘Verrucomicrobia’. *Int. J. Syst. Evol. Microbiol.* 57, 532–537.
19. Biderre-Petit, C., Jezequel, D., Dugat-Bony, E., Lopes, F., Kuever, J., Borrel, G., Viollier, E., Fonty, G., and Peyret, P. (2011). Identification of microbial communities involved in the methane cycle of a freshwater meromictic lake. *FEMS Microbiol. Ecol.* 77, 533–545.
20. Kalyuzhnaya, M.G., Lapidus, A., Ivanova, N., Copeland, A.C., McHardy, A.C., Szeto, E., Salamov, A., Grigoriev, I.V., Suci, D., Levine, S.R., *et al.* (2008). High-resolution metagenomics targets specific functional types in complex microbial communities. *Nat. Biotechnol.* 26, 1029–1034.

Plant Reproduction: Self-Incompatibility to Go

Alejandro Tovar-Mendez and Bruce McClure*

Department of Biochemistry, University of Missouri, 117 Schweitzer Hall, Columbia, MO 65211, USA

*Correspondence: mcclureb@missouri.edu

<http://dx.doi.org/10.1016/j.cub.2015.12.011>

In a new study, the *Papaver rhoeas* (poppy family) self-incompatibility system has been transferred into *Arabidopsis thaliana*, a distantly related plant with a very different floral structure. The simple poppy self-incompatibility system may finally make it possible to introduce this potentially valuable trait into any plant.

“It is an extraordinary fact that with many species, flowers fertilised with their own pollen are either absolutely or in some degree sterile; if fertilised with pollen from another flower on the same plant, they are sometimes, though rarely, a little more fertile; if fertilised with pollen from another individual or variety of the same species, they are fully fertile...”

Charles Darwin, 1876 [1]

The phenomenon Darwin describes is self-incompatibility (SI), or the ability of some plants to reject their own pollen. In fact, SI occurs widely, by some accounts, in half of flowering plant species [2]. So, it is not extraordinary, but it is unfamiliar because, with some notable exceptions, most domesticated plants were selected for the opposite trait — self-compatibility (SC). As convenient as SC is for plant breeding, the superior performance of outcrossed hybrids is well known, and plant breeders and biotechnologists alike

have long sought easy ways to make hybrids, with limited success. Since SI is a natural outcrossing system, it may be an attractive alternative, but until now a portable SI system seemed out of reach. Now, in a recent paper in *Science*, Lin *et al.* report successfully transferring the *Papaver rhoeas* SI system into the distantly related and SC *Arabidopsis thaliana*, suggesting that this type of SI may work in a wide variety of species [3].

Maize is the most familiar example of a hybrid crop. Most maize is produced as F₁ hybrids obtained by crossing inbred lines. These hybrids are produced through the simple, albeit laborious, process of detasseling the female inbred and planting it close to the inbred used as the male parent. While the superior performance of F₁ hybrids certainly applies to other crops, most crops have perfect flowers (i.e., the flowers contain both male and female reproductive organs). Consequently, F₁ hybrids often have to be made by carefully pollinating one flower at a time. Alternatives have been devised, for example, engineering or

employing controlled sterility (cytoplasmic male sterility, use of a gametocide to inhibit undesirable pollen, or conditional sterility), but these alternatives all have practical limitations [4–7]. SI could be an attractive addition to these methods since it does not involve either male or female sterility. This potential application of SI was one rationale for unraveling its genetic and molecular basis, although there are also more than enough basic-research reasons.

SI has been investigated in detail in Brassicaceae, Solanaceae, and Papaveraceae [8]. In all these well-understood systems, compatibility is determined by S-haplotype — one gene controls stigma-side function, and one or more others control the pollen-side. However, beyond the convention of naming the specificity genes as S-genes, the underlying SI mechanisms diverge drastically. A kinase-based signaling mechanism activated in the stigma prevents self-pollen growth in SI Brassicaceae, while stylar ribonucleases