

## 'Wild Type'

Isabel Askenasy, Jemima E. V. Swain, Pok-Man Ho, Rahan Rudland Nazeer, Amelie Welch, Éva Bernadett Bényei, Leonardo Mancini, Sivan Nir, Pinyu Liao and Martin Welch\*

### Abstract

In this opinion piece, we consider the meaning of the term 'wild type' in the context of microbiology. This is especially pertinent in the post-genomic era, where we have a greater awareness of species diversity than ever before. Genomic heterogeneity, *in vitro* evolution/selection pressures, definition of 'the wild', the size and importance of the pan-genome, gene–gene interactions (epistasis), and the nature of the 'wild-type gene' are all discussed. We conclude that wild type is an outdated and even misleading phrase that should be gradually phased out.

## BACKGROUND

One of the most widely used (and possibly abused) expressions in the modern discipline of microbiology is the term 'wild type'. Although originally popularized in the early 20th century by Hunt Morgan and colleagues to describe the phenotype and genotype of fruit flies, the term has now become pervasive across the whole of biology. But what does 'wild type' actually mean, especially in the context of microbiology? This is worth revisiting, if only because almost every microbiology paper uses the term. Moreover, the meaning of 'wild type' has not been critically reviewed in the microbiology literature since the 1960s [1] – well before the current transformative era of genomics. Even back then, and as Demerec *et al.* noted, the wild type was somewhat arbitrarily defined [1].

Microbiology is a broad discipline and the term 'wild type' means different things to different people. For some, 'wild type' conjures up characteristics or properties associated with environmental or clinical isolates (i.e. the organism 'in the wild'), whereas for others the term is simply used to define e.g. the parent/progenitor strain of a mutant or even just a widely used domesticated laboratory strain. For yet others, and to paraphrase Holmes, 'wild type' describes an individual organism or allele deemed 'normal' or typical for the species [2]. However, given the wealth of genomic data now available to us, we believe that the expression 'wild type' needs reappraisal and that there should be a transition towards more meaningful terminology. To be clear, our goal is not to say what the wild type should be, but rather to highlight the limitations of the term – particularly for those tempted to draw species-wide generalizations based on the behaviour of domesticated laboratory strains or small numbers of clinical/environmental isolates. The focus of our commentary is on bacteria, and for reasons of space limitation we have selected exemplars from just a handful of widely studied species. Nevertheless, the general arguments we raise likely apply to all microbial species.

## SOME EXAMPLES OF WILD TYPES

First, a little bit about the background to some well-known 'wild types'. Perhaps the most widely used wild type on the planet is *Escherichia coli* K12. Interestingly, K12 was originally isolated in 1922 from a patient recovering not from a classic enteric infection, but from diphtheria [3]. K12 is essentially avirulent and grows well on laboratory media, although after over a century of domestication (and bombardment with mutagens to cure it of its F-plasmid [4]) it could be argued that the organism has little in common with either its environmental or clinical cousins, or its progenitor [5]. K12 and its descendants also harbour a frameshift (just a single base deletion) in the *rph* operon, leading to polar effects on the downstream *pyrE* ORF and subsequent slower growth during pyrimidine limitation [6]. The ubiquitous *E. coli* DH strains, such as DH5α and DH1, are K12 derivatives engineered by Douglas Hanahan to be more transformable. For example, DH5α carries mutations in *endA1* and *recA1*, and is suitable for blue–white selection due to the *lacZΔM15* deletion (which removes residues 11–41 of LacZ, rendering the product

Received 02 July 2024; Accepted 14 August 2024; Published 30 August 2024

**Author affiliations:** <sup>1</sup>Department of Biochemistry, Hopkins Building, Tennis Court Road, Cambridge, CB2 1QW, Cambridge, UK.

**\*Correspondence:** Martin Welch, mw240@cam.ac.uk

**Keywords:** *Escherichia coli*; evolution; laboratory strain; *Pseudomonas aeruginosa*; wild-type.

**Abbreviations:** pwCF, person with cystic fibrosis.

001495 © 2024 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

active only in the presence of the *lacZ* alpha fragment). Other widely used laboratory strains, such as BL21, are not K12 derivatives, but instead derive from *E. coli* B [7]. The pedigree of the latter is somewhat murky, although it is clear that it was isolated by Félix d'Herelle (the co-discoverer, along with Twort, of bacteriophage) at the Institut Pasteur some time around 1918 [8]. Strain B and its derivatives have greater membrane permeability than K12, and are a better chassis when it comes to studies of bacterial evolution.

A particular favourite of the current authors is *Pseudomonas aeruginosa*, which claims two popular domesticated wild types; PAO1 and UCBPP-PA14. The former is a spontaneous chloramphenicol-resistant mutant of an Australian isolate ('PAO'), originally sourced from an infected wound by Bruce Holloway in 1954 [9]. (The chloramphenicol resistance derives from a loss-of-function mutation in *mexS*. This leads to constitutive expression of a chloramphenicol efflux pump, MexEF-OprN [10, 11]). PAO1 was described as growing 'luxuriantly' on nutrient agar, and was shown to be genetically amenable [9]. By contrast, UCBPP-PA14, more commonly known simply as PA14, was isolated in the 1970s from a patient in Pennsylvania (USA) with a burn wound [12]. PA14 is far more virulent than PAO1, likely due (in part) to its acquisition of a mutation in *ladS* [13]. PA14 is also resistant to the antibiotic rifampicin [14]. We will come back to the significance of genetic integrity later on. Unlike PAO1, which is a representative of a minor clonal lineage (designated 'clone W'), PA14 appears to be representative of the most common circulating *P. aeruginosa* clone worldwide (clone A) [15]. The phylogenetic tree of the species has recently been further refined based on high-resolution genomic analyses [16].

The brief overview above serves to illustrate two things. First, our existing, widely used wild types have sometimes been selected for very specific reasons, and may not be representative of the population as a whole. The *P. aeruginosa* clone A strain, PA14, is an exception to this, although somewhat incongruously, PAO1 (clone W strain) remains far more widely used as a laboratory model. Second, a common feature of many widely used laboratory wild types is that they grow rapidly, display some easily assayable phenotypes, and/or are genetically amenable (or have been engineered to be genetically amenable). These are not necessarily features typical of the species as a whole. Indeed, many field- or clinic-derived isolates are often rather slow growing and genetically rather intractable.

## THE HETEROGENEITY PROBLEM

The usual interpretation of 'wild type' – note that the term is hyphenated only when used as an adjective – is that it represents a fit genetic configuration of an organism in its natural environment. However, two decades' worth of genomic analyses have revealed that strains<sup>1</sup> of a given species isolated from 'the wild' often display considerable diversity, both in terms of genetic integrity (i.e. heterogeneity within the core genome<sup>2</sup>) and accessory genome content<sup>3</sup> [17, 18]. This is well illustrated by *P. aeruginosa*, which has a habit of colonizing diseased human airways, such as those in persons/people with cystic fibrosis (pwCF). Cross-sectional genomic analyses show that although the infecting strain in a given pwCF is usually genomically homogenous, it often undergoes a burst of adaptive radiation shortly after the initial colonization event [19, 20]. Furthermore, the colonizing populations continue to evolve over time, presumably in response to ongoing challenges, such as onslaught from the immune system, aggressive antibiotic intervention, and the presence of a variety of competing microbes [21]. Moreover, even within a given pwCF, we see spatial heterogeneity too, in that different lobes of the lung support genomically distinct sub-populations [22]. Interestingly, though, and in spite of their genomic divergence, these isolates display a remarkably convergent transcriptomic programme [23]. These data throw into question the notion of a single optimally fit *genomic* configuration; rather, it is more a case of 'many roads lead to Rome'. In the light of this, which one of the (potentially) hundreds or even thousands of different variants in a single pwCF is 'wild type'? Choose a different patient, and you'll have hundreds more unique derivatives to choose from. And so on.

## WHERE AND WHAT EXACTLY IS 'THE WILD'?

This is pertinent because, for many species, it is rather difficult to define what 'the wild' actually is – especially for organisms that are capable of occupying a range of disparate environments. For example, and as many people find out the hard way, pathogenic variants of *Escherichia coli* are just as at home in our waterways as they are in our GI tract, yet the former environments are usually cold and rather nutrient-limited, whereas the latter are warm and nutrient-replete. It follows that selection pressures in the two environments are likely to be radically different. Furthermore, the vehicle linking transmission between these two environments might be something as innocuous as a bean sprout [24], and the ability of the organism to survive on such substrata before eventually being consumed potentially layers on additional adaptive requirements. As noted above, the co-habitant species in each environment also play a part, and given that some of these co-habitants are likely to compete with the pathogen, this introduces an additional, if poorly defined, ecological selection pressure.

<sup>1</sup> We have tried to avoid entering into the controversial taxonomic world that aims to define strains, clones and species, although these terms have been discussed by Dijkshoorn et al. in their very readable 2000 article [56].

<sup>2</sup> Genes that are common between most lineages of the species. The core genome is usually described as being subject to purifying selection, leading to conservation.

<sup>3</sup> Genes, usually associated with horizontally acquired genomic islands, that likely confer an advantage in specific environmental scenarios.

There are also a number of misconceptions about 'the wild'. For example, for decades, researchers have made statements along the lines of '*Pseudomonas aeruginosa* is a ubiquitous micro-organism, often found on the surfaces of plants and animals, and in soil and water' or suchlike. However, although it can and has been isolated from all these sources, it turns out that it is actually rather hard to find outside of the built environment [25]. Consequently, some of our assumptions about microbial ecology are flawed and/or perpetuate earlier mistruths; 'the wild' can be a difficult thing to pin down.

## HORIZONTALLY ACQUIRED DNA

It could be argued that for pathogenic species such as *E. coli*, the initial infecting strain is closest to the wild type, but wild type for what? The roadside ditch or the human gut? And given that *E. coli* likely cycles between these different environments, is there such a thing as an 'initial infecting strain'? Unfortunately, bacteria do not come with a logbook detailing their individual natural histories, but if they did, it would likely tell a colourful story. More often than not, the bacteria's solution to such cyclic lifestyle changes is either to dump genetic elements and become a specialist in one or other niche, or to collect additional 'genetic baggage', often (but not exclusively [26]) in the form of horizontally acquired accessory genomic elements. To obtain a sense of the diversity conferred by these genomic elements, a recent study revealed that across ~1300 *P. aeruginosa* isolates, the core (conserved) genome comprised 665 genes, whereas the accessory 'pangenome' comprised >53000 genes [27]. Similarly, an analysis of 307 *E. coli* genome sequences revealed ~780 core genes and some 23000 pan-genes [28]. Clearly, our obsession with just a handful of wild types has come at the cost of understanding what the pan-genes do.

Horizontally acquired DNA has the potential to radically alter bacterial physiology, and there is evidence to suggest that the 'regulatory scars' left behind as DNA is successively acquired and then lost during the evolutionary history of an organism are not without consequences. A vivid example of this is seen in *Serratia marcescens* [29]. *Sma* 12 is a 'wild-type' clinical isolate, whereas *Sma* 274 is a 'wild type' veterinary isolate (originally obtained in 1922 from a sample of milk in the Netherlands). *Sma* 12 is non-pigmented but carries a functional quorum sensing system (i.e. is QS<sup>+</sup>). By contrast, *Sma* 274 is pigmented but is QS<sup>-</sup>. Phage-mediated transfer of the functional QS system from *Sma* 12 into *Sma* 274 led to pigment production in the latter coming under QS control. This remarkable 'genomic memory' of one-time QS regulation was also seen when the pigment biosynthetic cluster from *Sma* 274 was introduced into the non-pigmented but QS<sup>+</sup> strain (*Sma* 12). Here, pigment production in the recipient came under control of the *Sma* 12 QS system. This is a nice example of how horizontally acquired DNA can have large-scale and highly consequential effects on gene expression.

Indeed, bacteriophage may influence physiology substantially beyond their simple capacity to act as vehicles for the transfer of genetic material. The temperate phage of *P. aeruginosa* are reported to affect quorum sensing, biofilm formation and interactions with the host immune system [30], all of which can be modulated by curing the host of its phage [31].

More generally, strain-specific accessory genomic elements provide the organism with a toolkit that it can dip into when circumstances demand. The contents of this toolkit are often highly variable, but they clearly do confer a fitness advantage [32, 33]. Interestingly, individual clones of *P. aeruginosa* appear to prefer a specific repertoire of accessory elements, suggesting that the core and accessory genome segments are not randomly assembled [15]. Furthermore, in both *E. coli* and *P. aeruginosa*, epistatic interactions between core genes and horizontally acquired genomic elements are now known to radically alter gene essentiality [34, 35]. This is all pertinent because another feature of many well-established domesticated wild type strains is that they carry rather little in the way of accessory genome, or have been deliberately engineered to reduce their accessory genome content to a minimum. In this regard, while they are great experimental chassis for functional analysis of the core genome, they are very unlike most 'true' wild types.

## LABORATORY WILD TYPES ARE IN A CONTINUAL STATE OF EVOLUTION

Even once we have isolated a strain, designated it 'wild type', sequenced its genome, and frozen or lyophilized aliquots for archiving and dissemination, we still face problems. Genomes are never static, and biochemical 'slop', as well as the grinding inevitability of chemical tautomerism (reviewed in [36]), means that lesions in DNA are a fact of life. Although most of these lesions are repaired before they become fixed as single-nucleotide polymorphisms (SNPs) or indels, some inevitably slip through the statistical net. Of course, this low probability is offset by the vast numbers of cells in a rapidly dividing culture, such that after a few hours, there will be hundreds of different variants present, even assuming that the initial inoculum was a single, genomically homogeneous bacterial cell.

The unrelenting nature of mutation brings us to a third, and altogether more insidious, problem; that our supposedly well-defined, domesticated laboratory wild types are in a continual state of evolution. For example, in 2010, Klockgether *et al.* used a combination of physical mapping and next-generation sequencing to compare isolates of the common laboratory strain of *P. aeruginosa*, PAO1. Specifically, they compared the PAO1 that was used to obtain the reference genome sequence (PAO1-UW) with that of PAO1 sourced from a culture collection (PAO1-DSM) and a lineage of PAO1 that has been used to generate a widely used community resource, the University of Wisconsin two-allele transposon mutant bank (MPAO1) [37]. This analysis revealed that

the lineages differ due to a plethora of SNPs, indels, large-scale gene cluster duplications, and a 2.2 Mbp inversion. Importantly, these differences had measurable impacts on the fitness, virulence, and antimicrobial resistance of each lineage. Remarkably, Klockgether *et al.* also reported that certain sublines of PAO1-UW have acquired a Pfl-like prophage (RGP42) that is not present in the reference genome [37]. This prophage can only have been acquired during domestic passaging. As noted in the previous section, the presence of prophage can potentially have a large impact on the physiology of the organism. More recently, Chandler *et al.* carried out a similar genotypic/phenotypic analysis of PAO1 lineages donated from a selection of major research laboratories across North America [38]. Their conclusion, like that of Klockgether *et al.*, is that PAO1 is undergoing continual microevolution in the laboratory setting, and that this microevolution is not without functional consequences.

Dorman and Thomson came to a similar conclusion when investigating the ongoing evolution of *Vibrio cholerae* strain NCTC 30, originally isolated from a British soldier convalescing in Egypt in 1916 [39]. They note that NCTC 30 (which, interestingly, encodes a functional  $\beta$ -lactamase, even though the strain was isolated long before  $\beta$ -lactam antibiotics were introduced into the clinic) has since clearly undergone a good deal of evolution *in vitro* [40]. These authors also draw parallels with the ongoing laboratory evolution of several domesticated *E. coli* strains. A corollary of all this is that, when requesting a mutant from a collaborating laboratory, it is probably a good idea to also request the corresponding wild type from which it was derived and/or sequence the genome of both before proceeding.

## IN VITRO SELECTION PRESSURES

Another corollary of ongoing laboratory microevolution is that most well-established wild-type strains are now adapted to growth in common laboratory media such as lysogeny broth (LB). (In this regard, we now know that progression of *E. coli* K-12 through the exponential part of the growth curve in LB is far more complex than previously thought [41]). However, to capture physiologically relevant phenotypes, more exotic media are often required, such as artificial wound medium [42] or artificial sputum medium (ASM) [43]. Interestingly, we found that when a PAO1 derivative was introduced into ASM, many genes displayed a pronounced signature – based on their  $d_N/d_S$  value – of negative selection *cf.* growth in LB [44]. The most likely reason for this is that after decades of passaging in LB, PAO1 has become thoroughly adapted to this medium, such that a change in medium composition imposes a strong selection pressure. Consequently, we recommend prolonged passaging of wild types in any new medium prior to extensive further experimentation.

## WHAT IS A ‘WILD-TYPE GENE’?

The discussion above leads us to our fourth, somewhat more philosophical, point, and that is, ‘how do we define a wild-type gene?’ Even in the absence of a selection pressure, genes drift and a lineup of the amino acid sequences of any given open reading frame (ORF) will likely reveal a good deal of intra-species variation. But which of these variants (alleles) is wild type? Furthermore, not all genes are expressed in laboratory growth conditions; presumably, and based on the principle of use-it-or-lose-it, such genes may be susceptible to degradation. As noted above, this may be particularly relevant for domesticated lineages, which are generally maintained in a rather narrow set of growth conditions. By way of example, when the transcriptome and proteome of *E. coli* REL606 (an *E. coli* strain B derivative) was assessed across 34 different commonly used laboratory growth conditions, no evidence could be found for expression of 186 of the 4379 genes encoded [45]. Presumably, this subset of non-expressed genes will be subject (in the long term) to progressive loss of function in laboratory growth conditions.

## GENETIC CONTEXT AND EPISTASIS

To further confuse the situation, a common assumption is that most ORFs in a given domesticated wild type are functional, although this is not always the case. Indeed, some important phenotypic characteristics associated with widely used wild types are determined (in part) by loss-of-function mutations, e.g. *ladS* and PA14 [13] or *luxS* in DH5 $\alpha$  [46]. Moreover, and pertinent to the discussion above about the genome integrity of wild-type strains within and between different laboratories, lineages of PAO1 also display a marked tendency towards acquiring loss-of-function mutations in *mexT* [10, 47]. This is important, and not only because *mexT* is an important regulator in its own right; it also has proven epistatic interactions with other global regulatory genes, such as the quorum sensing master regulator, *lasR* [48–50]. (In this regard, it is worth noting that both genes represent mutational hotspots in the clinic and in the laboratory [47, 51, 52], and similar hotspotting of mutations in domesticated laboratory lineages has also been reported for other species [40].) Indeed, such gene epistatic interactions are likely a very widespread phenomenon (and can be very useful when untangling mechanistic interactions [53]), although we are only just beginning to scratch the surface of the problem.

## CLOSING COMMENTS

We hope that the discussion above serves to highlight the inadequacy of the term ‘wild type’ in microbiology. Our current domesticated strains have served (and are serving) us well in terms of understanding the basic biology of organisms – remaining



useful backgrounds for comparison against isogenic mutants and, ideally, permitting inter-laboratory reproducibility. However, whole-genome sequencing and the introduction of massively orthogonal investigative tools are forcing us to re-evaluate just how much these models actually represent the species as a whole [34].

We suggest that the expression ‘wild type’ is now so redundant – and indeed even misleading in the context of many studies – that its use in many circumstances should be actively discouraged, for all the reasons outlined above. Given the proven laboratory-to-laboratory variation in reference lineages, we further suggest that more care is taken in defining *exactly* which progenitor strain is being used in a given study. For example, if the MW team makes a mutant in, say, PAO1, that PAO1 should be reported in the accompanying publication as PAO1<sub>MW</sub> (subscript indicating that this is the PAO1 lineage routinely used in the Welch laboratory). If we receive a mutant from Another Laboratory, the experimental reference strain for that would be PAO1<sub>AL</sub>. And so on. This approach is not only good scientific practice in terms of pedigree tracing; it also reinforces the notion that an observed phenotype is relative to the named/defined progenitor, without making any presumption about the contribution of that phenotype to survival or growth in ‘the wild’.

We are not commenting here on *which* experimental reference strains should be used, although efforts to identify potentially more suitable ones are ongoing (e.g. [54]). In this regard, Fontana *et al.* have been implementing multi-omics-based approaches to define the optimal reference strains representing gut bifidobacterial species [55], and there is no reason why similar approaches should not be applied to other organisms. We also note that advances on this front will likely need to be accompanied by the parallel development of a genetic toolbox for each designated reference.

A second recommendation is that whichever strains are selected for downstream study, their genome should be sequenced on a moderately regular basis (and the sequence made publicly available), and care should be taken to minimise passaging of primary stocks on laboratory media. It would probably be a good idea to sequence the genome of any newly received experimental reference strain upon receipt, and best practice to report any changes *cf.* the sequence obtained by the provider in subsequent publications. That would certainly make pedigree records easier to construct and curate.

#### Funding information

I.A. is supported by the Leverhulme Trust. J.E.V.S. is supported by a studentship from the BBSRC and by the UK CF Trust. R.R.N. is supported by a studentship from the Benn W. Levy Trust. A.W. is supported by M.W. E.B.B. is supported by a studentship from the Oliver Gatty Trust. L.M. is supported by a post-doctoral fellowship from the Herchel Smith Fund. S.N. is supported by a post-doctoral fellowship from the Blavatnik Family Foundation. P.L. is supported by a visiting studentship from Stanford University (USA).

#### Author contributions

The idea for this commentary was suggested by conversations between M.W. and J.E.V.S. and A.W. I.A. and M.W. drafted the main body text, and all authors contributed towards editing and reviewing the content.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- Demerec M, Adelberg EA, Clark AJ, Hartman PE. A proposal for a uniform nomenclature in bacterial genetics. *Genetics* 1966;54:61–76.
- Holmes T. The wild type as concept and in experimental practice: a history of its role in classical genetics and evolutionary theory. *Stud Hist Philos Biol Biomed Sci* 2017;63:15–27.
- Bachmann BJ. Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol Rev* 1972;36:525–557.
- Salisbury V, Hedges RW, Datta N. Two modes of “curing” transmissible bacterial plasmids. *J Gen Microbiol* 1972;70:443–452.
- Smith HW. Is it safe to use *Escherichia coli* K12 in recombinant DNA experiments? *J Infect Dis* 1978;137:655–660.
- Jensen KF. The *Escherichia coli* K-12 “wild types” W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. *J Bacteriol* 1993;175:3401–3407.
- d’Herelle F. Sur le rôle du microbe filtrant bactériophage dans la dysentérie bacillaire. *Comptes Rendus de l’Académie des Sciences* 1918;167:970–972.
- Daegelen P, Studier FW, Lenski RE, Cure S, Kim JF. Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21(DE3). *J Mol Biol* 2009;394:634–643.
- Holloway BW. Genetic recombination in *Pseudomonas aeruginosa*. *J Gen Microbiol* 1955;13:572–581.
- Lee S, Gallagher L, Manoil C. Reconstructing a wild type *Pseudomonas aeruginosa* reference strain PAO1. *J Bacteriol* 2021;203:e0017921.
- Sobel ML, Neshat S, Poole K. Mutations in PA2491 (*mexS*) promote MexT-dependent *mexEF-oprN* expression and multidrug resistance in a clinical strain of *Pseudomonas aeruginosa*. *J Bacteriol* 2005;187:1246–1253.
- Mathee K. Forensic investigation into the origin of *Pseudomonas aeruginosa* PA14 – old but not lost. *J Med Microbiol* 2018;67:1019–1021.
- Mikkelsen H, McMullan R, Filloux A. The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in *ladS*. *PLoS One* 2011;6:e29113.
- Singh VK, Almpanti M, Maura D, Kitao T, Ferrari L, *et al.* Tackling recalcitrant *Pseudomonas aeruginosa* infections in critical illness via anti-virulence monotherapy. *Nat Commun* 2022;13:5103.
- Wiehmann L, Wagner G, Cramer N, Siebert B, Gudowius P, *et al.* Population structure of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 2007;104:8101–8106.
- Freschi L, Vincent AT, Jeukens J, Emond-Rheault J-G, Kukavica-Ibrulj I, *et al.* The *Pseudomonas aeruginosa* pan-genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity. *Genome Biol Evol* 2019;11:109–120.
- Wolfgang MC, Kulasekara BR, Liang X, Boyd D, Wu K, *et al.* Conservation of genome content and virulence determinants among

- clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 2003;100:8484–8489.
18. Kung VL, Ozer EA, Hauser AR. The accessory genome of *Pseudomonas aeruginosa*. *Microbiol Mol Biol Rev* 2010;74:621–641.
  19. Chung JCS, Becq J, Fraser L, Schulz-Trieglaff O, Bond NJ, et al. Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients. *J Bacteriol* 2012;194:4857–4866.
  20. Yang L, Jelsbak L, Marvig RL, Damkiær S, Workman CT, et al. Evolutionary dynamics of bacteria in a human host environment. *Proc Natl Acad Sci U S A* 2011;108:7481–7486.
  21. Winstanley C, O'Brien S, Brockhurst MA. *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Trends Microbiol* 2016;24:327–337.
  22. Jorth P, Staudinger BJ, Wu X, Hisert KB, Hayden H, et al. Regional isolation drives bacterial diversification within cystic fibrosis lungs. *Cell Host Microbe* 2015;18:307–319.
  23. Kordes A, Preusse M, Willger SD, Braubach P, Jonigk D, et al. Genetically diverse *Pseudomonas aeruginosa* populations display similar transcriptomic profiles in a cystic fibrosis explanted lung. *Nat Commun* 2019;10:3397.
  24. Tuffs A. Bean sprouts are identified as cause of *E. coli* outbreak. *BMJ* 2011;342:d3737.
  25. Crone S, Vives-Flórez M, Kvich L, Saunders AM, Malone M, et al. The environmental occurrence of *Pseudomonas aeruginosa*. *APMIS J* 2020;128:220–231.
  26. West PT, Chanin RB, Bhatt AS. From genome structure to function: insights into structural variation in microbiology. *Curr Opin Microbiol* 2022;69:102192.
  27. Diggle SP, Whiteley M. Microbe profile: *Pseudomonas aeruginosa*: opportunistic pathogen and lab rat. *Microbiology* 2020;166:30–33.
  28. Tantoso E, Eisenhaber B, Kirsch M, Shitov V, Zhao Z, et al. To kill or to be killed: pangenome analysis of *Escherichia coli* strains reveals a tailocin specific for pandemic ST131. *BMC Biol* 2022;20:146.
  29. Coulthurst SJ, Williamson NR, Harris AKP, Spring DR, Salmond GPC. Metabolic and regulatory engineering of *Serratia marcescens*: mimicking phage-mediated horizontal acquisition of antibiotic biosynthesis and quorum-sensing capacities. *Microbiology* 2006;152:1899–1911.
  30. Secor PR, Burgener EB, Kinnersley M, Jennings LK, Roman-Cruz V, et al. Pf bacteriophage and their impact on *Pseudomonas* virulence, mammalian immunity, and chronic infections. *Front Immunol* 2020;11:244.
  31. Schmidt AK, Schwartzkopf CM, Pourtois JD, Burgener EB, Faith DR, et al. Targeted deletion of Pf prophages from diverse *Pseudomonas aeruginosa* isolates has differential impacts on quorum sensing and virulence traits. *J Bacteriol* 2024;206:e0040223.
  32. Hao W. Extensive genomic variation within clonal bacterial groups resulted from homologous recombination. *Mob Genet Elements* 2013;3:e23463.
  33. Wawire SA, Reva ON, O'Brien TJ, Figueroa W, Dinda V, et al. Virulence and antimicrobial resistance genes are enriched in the plasmidome of clinical *Escherichia coli* isolates compared with wastewater isolates from western Kenya. *Infect Genet Evol* 2021;91:104784.
  34. Rousset F, Cabezas-Caballero J, Piastra-Facon F, Fernández-Rodríguez J, Clermont O, et al. The impact of genetic diversity on gene essentiality within the *Escherichia coli* species. *Nat Microbiol* 2021;6:301–312.
  35. Poulsen BE, Yang R, Clatworthy AE, White T, Osmulski SJ, et al. Defining the core essential genome of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 2019;116:10072–10080.
  36. On YY, Welch M. The methylation-independent mismatch repair machinery in *Pseudomonas aeruginosa*. *Microbiology* 2021;167:001120.
  37. Klockgether J, Munder A, Neugebauer J, Davenport CF, Stanke F, et al. Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains. *J Bacteriol* 2010;192:1113–1121.
  38. Chandler CE, Horspool AM, Hill PJ, Wozniak DJ, Schertzer JW, et al. Genomic and phenotypic diversity among ten laboratory isolates of *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 2019;201:e00595-18.
  39. Dorman MJ, Kane L, Domman D, Turnbull JD, Cormie C, et al. The history, genome and biology of NCTC 30: a non-pandemic *Vibrio cholerae* isolate from World War One. *Proc R Soc B* 2019;286:20182025.
  40. Dorman MJ, Thomson NR. "Community evolution" – laboratory strains and pedigrees in the age of genomics. *Microbiology* 2020;166:233–238.
  41. Sezonov G, Joseleau-Petit D, D'Ari R. *Escherichia coli* physiology in Luria-Bertani broth. *J Bacteriol* 2007;189:8746–8749.
  42. Pouget C, Dunyach-Remy C, Bernardi T, Provot C, Tasse J, et al. A relevant wound-like *in vitro* media to study bacterial cooperation and biofilm in chronic wounds. *Front Microbiol* 2022;13:705479.
  43. Aiyer A, Manos J. The use of artificial sputum media to enhance investigation and subsequent treatment of cystic fibrosis bacterial infections. *Microorganisms* 2022;10:1269.
  44. On YY, Figueroa W, Fan C, Ho P-M, Bényei ÉB, et al. Impact of transient acquired hypermutability on the inter- and intra-species competitiveness of *Pseudomonas aeruginosa*. *ISME J* 2023;17:1931–1939.
  45. Caglar MU, Houser JR, Barnhart CS, Boutz DR, Carroll SM, et al. The *E. coli* molecular phenotype under different growth conditions. *Sci Rep* 2017;7:45303.
  46. Surette MG, Miller MB, Bassler BL. Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proc Natl Acad Sci U S A* 1999;96:1639–1644.
  47. LoVullo ED, Schweizer HP. *Pseudomonas aeruginosa* *mexT* is an indicator of PAO1 strain integrity. *J Med Microbiol* 2020;69:139–145.
  48. Liu Y, Ahator SD, Wang H, Feng Q, Xu Y, et al. Microevolution of the *mexT* and *lasR* reinforces the bias of quorum sensing system in laboratory strains of *Pseudomonas aeruginosa* PAO1. *Front Microbiol* 2022;13:821895.
  49. Kostylev M, Kim DY, Smalley NE, Salukhe I, Greenberg EP, et al. Evolution of the *Pseudomonas aeruginosa* quorum-sensing hierarchy. *Proc Natl Acad Sci U S A* 2019;116:7027–7032.
  50. Oshri RD, Zrihen KS, Shner I, Omer Bendori S, Eldar A. Selection for increased quorum-sensing cooperation in *Pseudomonas aeruginosa* through the shut-down of a drug resistance pump. *ISME J* 2018;12:2458–2469.
  51. Luján AM, Moyano AJ, Segura I, Argaraña CE, Smania AM. Quorum-sensing-deficient (*lasR*) mutants emerge at high frequency from a *Pseudomonas aeruginosa* *mutS* strain. *Microbiology* 2007;153:225–237.
  52. Weimann A, Dinan AM, Ruis C, Bernut A, Pont S, et al. Evolution and host-specific adaptation of *Pseudomonas aeruginosa*. *Science* 2024;385:eadi0908.
  53. Yamaguchi S, Aizawa S-I, Kihara M, Isomura M, Jones CJ, et al. Genetic evidence for a switching and energy-transducing complex in the flagellar motor of *Salmonella typhimurium*. *J Bacteriol* 1986;168:1172–1179.
  54. Mahenthalingam E, Weiser R, Floto RA, Davies JC, Fothergill JL. Selection of relevant bacterial strains for novel therapeutic testing: a guidance document for priority cystic fibrosis lung pathogens. *Curr Clin Micro Rpt* 2022;9:33–45.
  55. Fontana F, Alessandri G, Tarracchini C, Bianchi MG, Rizzo SM, et al. Designation of optimal reference strains representing the infant gut bifidobacterial species through a comprehensive multi-omics approach. *Environ Microbiol* 2022;24:5825–5839.
  56. Dijkshoorn L, Ursing BM, Ursing JB. Strain, clone and species: comments on three basic concepts of bacteriology. *J Med Microbiol* 2000;49:397–401.