Key Points

- Microbial genomics studies the genomes of microorganisms like bacteria, archaea, fungi, and viruses, starting with the first virus genome in 1977.

- The first bacterial genomes, *Haemophilus influenzae* and *Mycoplasma genitalium*, were sequenced in 1995, marking the beginning of bacterial genomics.

- Key advancements include the pangenome concept in 2005, next-generation sequencing in the mid-2000s, and metagenomics in 1998.

- Surprisingly, many bacterial species have open pangenomes, meaning they can acquire new genes indefinitely, affecting their evolution and adaptation.

History of Bacterial Genomics

Bacterial genomics began with the sequencing of *Haemophilus influenzae* and *Mycoplasma genitalium* in 1995, published in *Science* ([Whole-genome random sequencing and assembly of Haemophilus influenzae Rd](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49772/)) and ([The minimal gene complement of Mycoplasma genitalium](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49773/)). These milestones allowed scientists to understand bacterial genetics at a whole-genome level. Over time, technologies like next-generation sequencing, reviewed in 2008 ([Next-generation DNA sequencing](https://doi.org/10.1038/nbt.1491)), and concepts like the pangenome, introduced in 2005 ([Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae](https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0010027)), expanded our knowledge. Metagenomics, starting in 1998 ([Molecular biological access to the chemistry of unknown soil microbes](https://doi.org/10.1016/S1074-5521(98)90166-8)), and single-cell genomics in 2013 ([Insights into the phylogeny and coding potential of microbial dark matter](https://doi.org/10.1038/nature12352)), further revolutionized the field by studying unculturable bacteria and microbial communities.

Context: Other Microbes, Single-Celled Organisms, and Viruses

For context, the first virus genome, bacteriophage ϕX174, was sequenced in 1977 ([Nucleotide sequence of bacteriophage ϕX174 DNA](https://doi.org/10.1038/265687a0)). Archaea followed with *Methanococcus jannaschii* in 1996 ([Complete genome sequence of the methanogenic archaeon, Methanococcus

jannaschii](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1852588/)). Fungi started with *Saccharomyces cerevisiae* in 1996 ([Life with 6000 genes](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858218/)), and protists with *Plasmodium falciparum* in 2002 ([Genome sequence of the human malaria parasite Plasmodium falciparum](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730924/)). These developments highlight the broad scope of microbial genomics beyond bacteria.

A Comprehensive Survey of Microbial Genomics History

Introduction

Microbial genomics, the study of genomes from microorganisms including bacteria, archaea, fungi, and viruses, has revolutionized our understanding of microbial biology, evolution, and ecology. This survey traces the history of microbial genomics, with a primary focus on bacterial genomics, while providing context through the genomics of other microbes, single-celled organisms, and viruses. It emphasizes technological advancements and key discoveries, aligning with the request for a detailed historical overview.

Historical Development of Bacterial Genomics

The field of bacterial genomics commenced with the sequencing of the first complete bacterial genomes in 1995. Specifically, *Haemophilus influenzae* (1.83 Mbp) and *Mycoplasma genitalium* (0.58 Mbp) were sequenced, published in *Science* by Fleischmann et al. ([Whole-genome random sequencing and assembly of Haemophilus influenzae Rd](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49772/)) and Fraser et al. ([The minimal gene complement of Mycoplasma genitalium](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49773/)), respectively. These achievements were pivotal, enabling the inference of minimal gene sets and laying the foundation for comparative genomics.

Following these initial successes, the field saw rapid growth. The concept of the pangenome, introduced by Tettelin et al. in 2005 ([Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae](https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0010027)), highlighted the dynamic nature of bacterial genomes, comprising core genes shared by all strains and accessory genes varying between strains. This was significant, as it revealed that many bacterial species, such as *Streptococcus agalactiae*, have open pangenomes, capable of acquiring new genes indefinitely through mechanisms like horizontal gene transfer.

Horizontal gene transfer, a key driver of bacterial evolution, was extensively reviewed by Ochman et al. in 2000 ([Lateral gene transfer and the nature of bacterial innovation](https://doi.org/10.1038/35005000)), emphasizing its role in bacterial innovation and adaptation. This process was crucial for understanding bacterial pathogenicity and antibiotic resistance, areas where genomics has provided deep insights.

Technological advancements further propelled bacterial genomics. The development of next-generation sequencing (NGS) technologies, detailed by Shendure and Ji in 2008 ([Next-generation DNA sequencing](https://doi.org/10.1038/nbt.1491)), reduced sequencing costs and increased throughput, enabling the sequencing of thousands of bacterial genomes. Metagenomics, pioneered by Handelsman et al. in 1998 ([Molecular biological access to the chemistry of unknown soil microbes](https://doi.org/10.1016/S1074-5521(98)90166-8)), allowed the study of microbial communities without culturing, revealing the vast diversity of unculturable bacteria. More recently, single-cell genomics, exemplified by Rinke et al. in 2013 ([Insights into the phylogeny and coding potential of microbial dark matter](https://doi.org/10.1038/nature12352)), provided insights into the phylogeny and coding potential of microbial dark matter, expanding our understanding of bacterial diversity.

Context: Genomics of Other Microbes, Single-Celled Organisms, and Viruses

To provide a comprehensive context, the history of genomics for other microorganisms is essential. The first genome sequenced was that of bacteriophage ϕX174, a virus, in 1977 by Sanger et al. ([Nucleotide sequence of bacteriophage ϕX174

DNA](https://doi.org/10.1038/265687a0)), marking the beginning of genomic studies in viruses.

Archaea, another domain of single-celled organisms, saw its first genome sequenced in 1996 with *Methanococcus jannaschii* by Bult et al. ([Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1852588/)). This was followed by additional archaeal genomes, contributing to our understanding of extremophiles and their evolutionary relationships.

Fungi, including single-celled yeasts and multicellular molds, had their first genome sequenced with *Saccharomyces cerevisiae* in 1996 by Goffeau et al. ([Life with 6000 genes](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858218/)). This eukaryotic model organism provided insights into fungal genetics and biotechnology applications.

Protists, eukaryotic microorganisms not classified as animals, plants, or fungi, had their first genome sequenced with *Plasmodium falciparum*, the malaria parasite, in 2002 by Gardner et al. ([Genome sequence of the human malaria parasite Plasmodium falciparum](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730924/)). This was followed by other protist genomes, such as *Entamoeba histolytica* in 2003 and *Tetrahymena thermophila* in 2004, expanding our knowledge of eukaryotic microbial diversity.

Key Discoveries and Applications

Bacterial genomics has led to significant discoveries, such as the role of horizontal gene transfer in antibiotic resistance, the identification of virulence factors, and the understanding of metabolic pathways. The pangenome concept, for instance, has shown that bacterial genomes are highly dynamic, with many species exhibiting open pangenomes, a surprising detail that challenges traditional views of genome stability. Applications include the development of vaccines, diagnostics, and strategies for combating antibiotic resistance, with metagenomics and single-cell genomics enhancing our ability to study microbial ecosystems and unculturable bacteria.

Landmark Publications and Dates

The following table summarizes key landmark publications in microbial genomics, focusing on bacterial genomics and providing context for other microbes:

Year	Milestone	Reference
1977	First virus genome sequenced (*ϕX174*)	Sanger et al., Nature, 1977 ([Nucleotide sequence of bacteriophage ϕX174 DNA](https://doi.org/10.1038/265687a0))
1995	First bacterial genomes sequenced (*H. influenzae*, *M. genitalium*)	Fleischmann et al., Science, 1995 ([Whole-genome random sequencing and assembly of Haemophilus influenzae Rd](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49772/)); Fraser et al., Science, 1995 ([The minimal gene complement of Mycoplasma genitalium](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49773/))
1996	First archaeal genome sequenced (*M. jannaschii*)	Bult et al., Science, 1996 ([Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1852588/))
1996	First fungal genome sequenced (*S. cerevisiae*)	Goffeau et al., Science, 1996 ([Life with 6000 genes](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858218/))
2002	First protist genome sequenced (*P. falciparum*)	Gardner et al., Nature, 2002 ([Genome sequence of the human malaria parasite Plasmodium falciparum](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730924/))
2005	Pangenome concept introduced	Tettelin et al., PLOS Genetics, 2005 ([Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae](https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0010027))
2000	Review on horizontal gene transfer	Ochman et al., Nature, 2000 ([Lateral gene transfer and the nature of bacterial innovation](https://doi.org/10.1038/35005000))

| 2008 | Review on next-generation sequencing | Shendure & Ji, Nature Biotechnology, 2008 ([Next-generation DNA sequencing](https://doi.org/10.1038/nbt.1491)) |

| 1998 | Early metagenomics paper | Handelsman et al., Chemistry & Biology, 1998 ([Molecular biological access to the chemistry of unknown soil microbes](https://doi.org/10.1016/S1074-5521(98)90166-8)) |

| 2013 | Single-cell genomics for microbial dark matter | Rinke et al., Nature, 2013 ([Insights into the phylogeny and coding potential of microbial dark matter](https://doi.org/10.1038/nature12352)) |

Conclusion

The history of microbial genomics, particularly bacterial genomics, reflects a trajectory of technological innovation and scientific discovery. From the first bacterial genome sequences in 1995 to the advent of metagenomics and single-cell genomics, the field has provided profound insights into microbial diversity and function. The context of other microbes, including viruses, archaea, fungi, and protists, underscores the broad impact of genomics across microbial life, with ongoing efforts promising further advancements.

Key Citations

- [Whole-genome random sequencing and assembly of Haemophilus influenzae Rd](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49772/)

- [The minimal gene complement of Mycoplasma genitalium](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC49773/)

- [Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae](https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0010027)

- [Lateral gene transfer and the nature of bacterial innovation](https://doi.org/10.1038/35005000)

- [Next-generation DNA sequencing](https://doi.org/10.1038/nbt.1491)

- [Molecular biological access to the chemistry of unknown soil microbes](https://doi.org/10.1016/S1074-5521(98)90166-8)

- [Insights into the phylogeny and coding potential of microbial dark matter](https://doi.org/10.1038/nature12352)

- [Nucleotide sequence of bacteriophage ϕX174 DNA](https://doi.org/10.1038/265687a0)

- [Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1852588/)

- [Life with 6000 genes](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858218/)

- [Genome sequence of the human malaria parasite Plasmodium falciparum](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730924/)