Co-Evolutionary Characterization of PBP2 as the Predominant Siderophore Recognizer in Diverse Gram-Positive Bacteria

Linlong Yu[1], Guanyue Xiong[2], Zhiyuan Li[1,2*]

[1]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[2]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

*Correspondence: zhiyuanli@pku.edu.cn (Zhiyuan Li)

Abstract

Iron is a scarce yet essential nutrient for bacteria, which often acquire it by secreting siderophores, diverse small molecules that form high-affinity complexes with iron for microbial uptake. Understanding microbial iron interaction networks requires characterizing the specificity of siderophore recognition. In Gram-positive bacteria, substrate-binding proteins (SBPs) bind iron-siderophore complexes and deliver them to ABC transporters for import. However, the SBPs responsible for selective recognition remain poorly characterized, hindering large-scale data mining and network reconstruction. Here, we systematically analyzed siderophore uptake systems in five representative Gram-positive genera. Through genome mining and co-evolutionary analysis, we characterized Peripla_BP_2 subtype SBPs as the predominant recognizer family, exhibiting strong co-evolutionary signals with siderophore synthetase genes. Unlike their counterparts in Gram-negative bacteria, synthetase and recognizer genes in Gram-positive bacteria are sometimes genomically distant yet display coordinated transcriptional regulation, indicating a spatially decoupled but transcriptionally synchronized mechanism. Our findings underscore key differences between Gram-positive and -negative iron acquisition systems, providing foundational knowledge for large-scale inference of siderophore-mediated microbial interactions.

Keywords: siderophore, Gram-positive bacteria, substrate-binding protein, PBP2, iron acquisition

Introduction

Iron is essential for bacterial growth, survival, and virulence[1], serving as a cofactor in key enzymes involved in DNA replication and electron transfer[2, 3]. Despite its abundance on Earth, iron's bioavailability in natural environments is severely limited due to the oxidation of ferrous iron (Fe^{2+}) to insoluble ferric iron (Fe^{3+}) following the Great Oxidation Event [4-6]. To overcome this scarcity, most bacteria secrete siderophores to the surrounding environment to chelate iron[4]. Siderophore is a class of small, high-affinity iron-chelating secondary metabolites[1, 4, 7, 8], biosynthesized by non-ribosomal peptide synthetases (NRPS) or NRPS-independent siderophore synthetases (NIS) pathways[8-10]. Once secreted, siderophores scavenge iron from the environment to form siderophore-iron complexes, which are then imported back into bacterial cells through specific membrane recognizers[11-13].

The specific recognition between siderophores and their recognizers mediates intricate microbial interactions in iron competition[4]. Previous studies have shown that siderophores are preferentially bound and imported by specific receptors, akin to keys fitting locks[13-16]. For instance, three strains of *Pseudomonas aeruginosa* each produce distinct pyoverdines, a typical *Pseudomonas* siderophore, and primarily utilize only their own, with minimal cross-recognition or uptake[13, 15]. Meanwhile, microbes often possess multiple membrane recognizers to pirate xenosiderophores produced by others. For example, *E. coli* K-12 only synthesizes Enterobactin but also acquires foreign siderophores via receptors like Cir (for Heterobactin) and FhuE (for Coprogen)[17, 18]. Similarly, the Gram-positive *Bacillus subtilis* produces Bacillibactin yet imports additional siderophores, including Enterobactin, Ferrichrome, and Ferrioxamine E[19-21]. This lock-and-key specificity likely evolves from an arms race between producers and cheaters, shaping rich community games: For microbes with matching receptors, a siderophore acts as a shared public good facilitating iron uptake; for those without, it becomes a “public bad” that sequesters iron away[22]. Understanding siderophore-receptor specificity is key to dissecting complex microbial communities[4].

Nevertheless, siderophore uptake mechanisms differ fundamentally between Gram-negative and Gram-positive bacteria, complicating community-scale investigations[23]. Gram-negative bacteria have been more thoroughly studied: siderophores are imported via TonB-dependent outer membrane receptors (TonBDR) that undergo conformational changes upon ligand binding[24]. Structural analyses have pinpointed residues in loop L7 and the plug domain that govern receptor selectivity [25, 26], while recent bioinformatics revealed sequences near the plug domain as key determinants of pyoverdine specificity in *Pseudomonas[27]*. Building on this systematic characterization, our prior work utilizes sequence co-evolution between synthetase and receptor genes to reconstruct pyoverdine-mediated interaction networks in *Pseudomonas[28]*. This method assumes strongest co-evolution signals arise between matched synthetase-receptor pairs, enabling algorithms to infer siderophore production and uptake profiles per strain. The predicted networks aligned

well with experimental validations and uncovered key distinctions between pathogens and non-pathogens.

Gram-positive species lack an outer membrane and instead employ ATP-binding cassette (ABC) transporter systems to import ferric-siderophore complexes[29]. This process begins with substrate-binding proteins (SBPs), often lipoproteins tethered to the membrane, that recognize and bind siderophores with high affinity[30]. Then SBP subsequently deliver the ligand to the permease component of the ABC transporter for ATP-driven translocation into the cytoplasm (Fig. 1a)[31]. Previous research has elucidated recognition mechanisms in model systems, including *Bacillus cereus* YxeB, which targets ferrioxamine-type siderophores via a shuttle mechanism for iron exchange[32]; *Bacillus subtilis* FeuA, which binds enterobactin-type siderophores with limited promiscuity through surface pockets and conformational closure[33]. However, SBPs constitute a huge and heterogeneous superfamily, classified into over 33 families in Pfam and eight structural groups, reflecting adaptations for diverse substrates including siderophores[34, 35]. Under such vast diversity, the precise identities of siderophore-specific SBPs in Gram-positive bacteria remain incompletely characterized, with limited systematic sequence-based comparison across species, hindering broader data-driven studies.

In this work, we developed a genome mining pipeline to systematically investigate siderophore biosynthetic gene clusters (BGCs) across five representative Gram-positive genera: *Bacillus*, *Streptomyces*, *Rhodococcus*, *Corynebacterium*, and *Staphylococcus* (Fig. 1b). Within siderophore BGCs, we identified Peripla_BP_2 subtype SBPs as the most putative recognizer family that determines substrate specificity, exhibiting strong co-evolution with siderophore synthetase genes (Pearson correlation = 0.90). We further employed sequence-based feature analysis to characterize specificity-determining residues and examined transcriptional regulatory mechanisms governing coordinated expression of siderophore production and uptake systems. Collectively, our findings identify PBP2 as a key siderophore receptor in diverse Gram-positive bacteria, providing a foundation for understanding siderophore-receptor interactions and offering potential targets for disrupting iron acquisition in pathogenic Gram-positive bacteria.

Results

1. Overview of siderophore BGCs in five representitive Gram-positive genera

To begin our systematic investigation of siderophore receptors in Gram-positive bacteria, we surveyed siderophore biosynthetic gene clusters (BGCs) across five representative genera: *Bacillus* (including famous model organism *B.subtilis* [36] and pathogens *B. cereus* and *B. anthracis*[37]), *Staphylococcus* (including *S.aureus*, a foodborne pathogen with extensively characterized iron uptake mechanisms[38]), *Streptomyces* (known for secondary metabolic potential[39]), *Rhodococcus* (for their industrial relevance in bioremediation) and *Corynebacterium* (relevant for its amino acid production ability). We collected 5,940 complete genomes that belong to these five Gram-positive bacterial genera from the NCBI RefSeq database (accessed January 21, 2025).

Phylogenetic analysis is consistent with known taxonomic classifications (Fig. 1c): Members of the class Actinomycetes (*Streptomyces*, *Rhodococcus*, and *Corynebacterium*) exhibited close phylogenetic relationships, as did *Bacillus* and *Staphylococcus* belonging to class Bacilli. AntiSMASH v7.0[40] was used to extract 81,923 biosynthetic gene clusters (BGCs) from the 5,940 genomes. Of these, 11,095 (13.5%) were annotated as siderophore BGCs. Notably, 95.9% of the genomes contained at least one siderophore BGC (Fig. 1c), highlighting the ubiquity of siderophore-mediated iron acquisition. Further analysis revealed genus-specific preferences in siderophore biosynthetic pathways. Genus-specific biosynthetic preferences were evident: *Staphylococcus* predominantly employed the NIS pathway, whereas *Rhodococcus* favored NRPS for siderophore production (Fig. 1c).

We next assessed the biosynthetic potential of the five genera (Fig. 1d). *Streptomyces* showed the highest average BGC count (29.9 BGC per genome on average), consistent with previous reports[41], whereas *Corynebacterium* displayed the lowest (5.1 BGC per genome on average). *Corynebacterium* has been known to produce monosodium glutamate (MSG) highly efficiently in industry[42], which suggests a metabolic emphasis on primary pathways at the expense of secondary metabolite biosynthesis.

We evaluated the novelty of the predicted BGCs by comparison against the manually curated Minimum Information about a Biosynthetic Gene cluster (MIBiG) database (Fig. 1e). Notably, the MIBiG database exhibits genus-level biases: Well-studied genera such as *Bacillus* and *Staphylococcus* showed high similarity to the MIBiG database, as over 80% of their siderophore BGCs having similarity scores above 80% to known clusters. In contrast, less extensively studied genera like *Rhodococcus* and *Corynebacterium* had limited representation, with less than 20% of their siderophore BGCs achieving the same similarity threshold. Surprisingly, despite being one of the most extensively studied genera, *Streptomyces* exhibited unexpected

novelty in its siderophore biosynthetic potential, with more than half of its siderophore BGCs showing less than 50% similarity to known clusters in MIBiG. This finding suggests that *Streptomyces* may still hold significant potential for the discovery of novel siderophores.

In our dataset, we identified 16 types of siderophore BGCs with similarity scores above 80%, all of which correspond to experimentally validated siderophores(Table 1). Although most of the remaining siderophore BGCs lack experimental confirmation, several NRPS-derived monomers could be reliably annotated (Fig. S1). When the MIBiG database contains multiple BGC entries for the same siderophore, such as Heterobactin A (BGC0000371.5) and Heterobactin B (BGC0002696.2) that differ only in gene order or exhibit minor sequence variations, we consolidated them into a single siderophore BGC. α-Diversity analysis of all predicted siderophore BGCs revealed that *Streptomyces* exhibited the highest diversity of siderophore BGCs (Fig. 1f).

Table 1. The information of 16 types of siderophore BGC

Siderophore name	MIBiG accession[a]	SIDERITE ID[b]	Biosynthetic pathway	Genus
Bacillibactin	BGC0000309.5 BGC0002695.2 BGC0001185.5	SID00152 SID00195 SID00151 SID00746 SID00747	NRPS	*Bacillus*
Petrobactin	BGC0000942.4	SID00129	NIS	*Bacillus*
Staphyloferrin A	BGC0000944.5	SID00117	NIS	*Staphylococcus*
Staphyloferrin B	BGC0000943.5	SID00123 SID00125	NIS	*Staphylococcus*
Heterobactin	BGC0002696.2; BGC0000371.5	SID00813 SID00594 SID00595 SID00597	NRPS	*Rhodococcus*
Dehydroxynocardamine	BGC0002073.3	SID00598; SID00018	NIS	*Corynebacterium*
Desferrioxamine	BGC0002305.2 BGC0001478.5 BGC0001453.5 BGC0000940.5	SID00012 SID00013 SID00016–SID00018 SID00020–SID00022 SID00027 SID00028 SID00030–SID00033 SID00335 SID00336 SID00338 SID00340 SID00344 SID00362 SID00655–SID00657 SID00772–SID00779	NIS	*Streptomyces*

		SID00727 SID00849		
Coelichelin	BGC0000325.5	SID00305	NRPS	*Streptomyces*
Scabichelin	BGC0000423.5	SID00545	NRPS	*Streptomyces*
Peucechelin	BGC0002466.3	SID00298	NRPS	*Streptomyces*
Fuscachelin	BGC0000359.5	SID00222 SID00223 SID00596	NRPS	*Streptomyces*
Albomycin delta2	BGC0002300.2	SID00560	NRPS	*Streptomyces*
Griseobactin	BGC0000368.5	SID00201	NRPS	*Streptomyces*
Paenibactin	BGC0000401.5	SID00196	NRPS	*Streptomyces*
Qinichelins	BGC0001752.4	SID00588 SID00589 SID00590 SID00591	NRPS	*Streptomyces*
EDHA	BGC0002567.2 BGC0001587.5	NA[c]	NIS-like[d]	*Streptomyces*

a. In the MIBiG database, some siderophores have multiple accessions.

b. Certain siderophores exhibit structural variations due to different modifications, which correspond to multiple IDs in the SIDERT database.

c. This type of siderophore is not yet included in the SIDERT database.

d. This siderophore belongs to the aminopolycarboxylic acid type, and its biosynthetic pathway is NRP-independent; however, it does not contain canonical NIS synthetases [43].

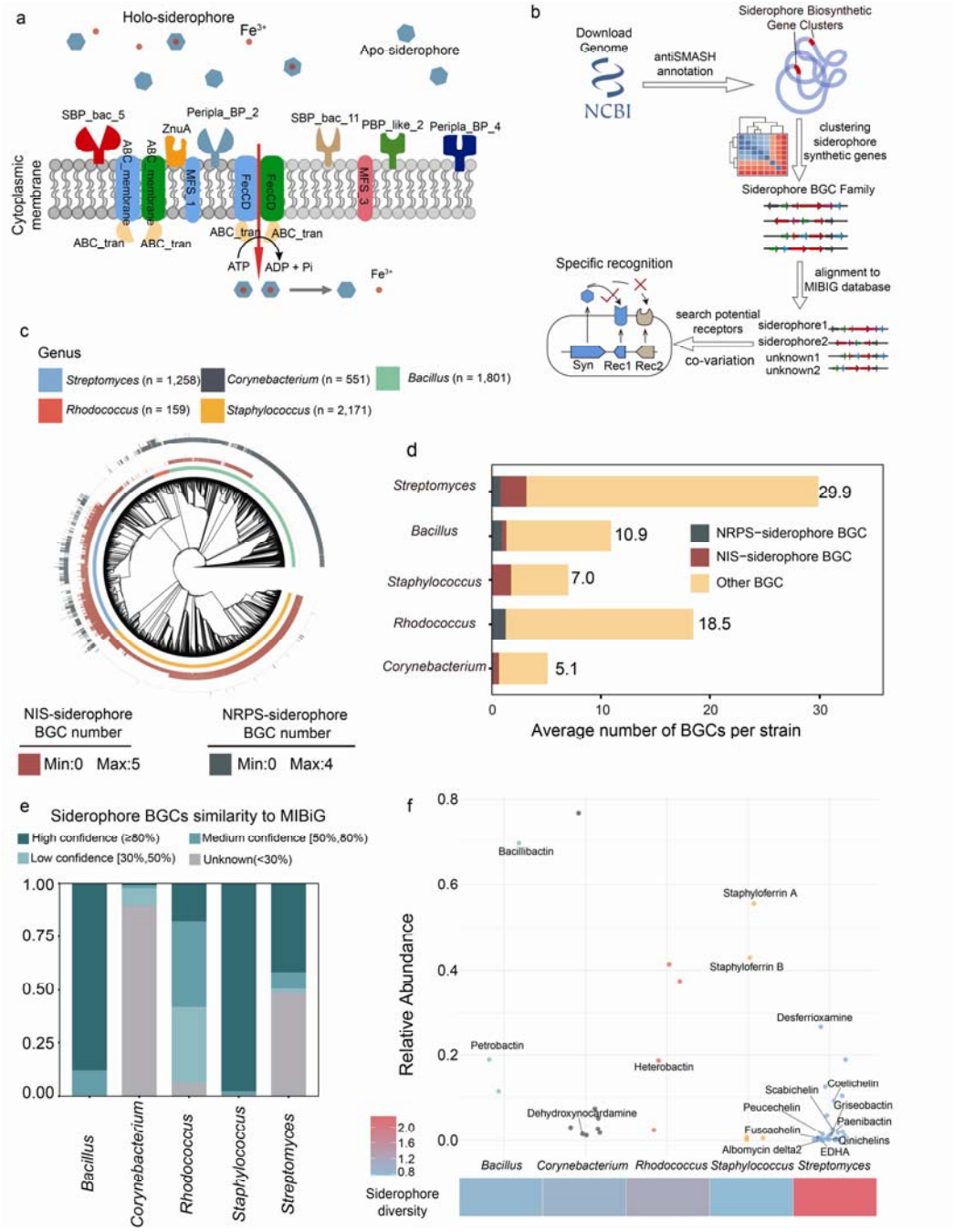

Figure 1. Overview of siderophore BGCs in five representative Gram-positive genera.

a. Cartoon illustration of siderophore-mediated iron acquisition in Gram-positive bacteria. Iron-siderophore complexes are first recognized and bound by substrate-binding proteins (SBPs) on the cytoplasmic membrane, then translocated into the cytoplasm via the ABC transporter complex. Different colors and shapes denote SBP subtypes.

b. Workflow of genome mining for the identification of potential siderophore recognizers.

c. Phylogenetic relationships among the 5,940 strains, based on concatenated alignment of 400 single-copy conserved genes. Colors in the first ring distinguish the five genera; bar heights in the second and third rings indicate the number of NIS-type and NRPS-type siderophore BGCs per strain, respectively.

d. BGC composition. Stacked bars show the average number of BGCs per strain across genera, color-coded as NRPS-type siderophore BGCs (green), NIS-type siderophore BGCs (red), and other BGCs (yellow).

e. Similarity of siderophore BGCs to known BGCs in the MIBiG database. Color-coded bars highlight BGC similarity percentages against the MIBiG database: <30% (gray), [30%,50%) (light green), [50%,80%) (green), ≥80% (dark green).

f. Relative abundance of siderophore BGCs across genera. Labeled dots denote BGCs with ≥80% similarity to MIBiG entries. Siderophore diversity was calculated using Shannon entropy, analogous to alpha-diversity in microbial communities.

2. PBP2-subtype SBPs function as siderophore recognizer in five Gram-positive bacterial genera

Following analysis of siderophore BGCs, we aimed to identify their corresponding recognizers: the membrane proteins that best predict which types of siderophores cells can utilize. Employing LLM-assisted literature mining combined with manual curation, we assembled a dataset comprising 22 experimentally validated siderophore-binding proteins localized in the cytoplasmic membrane from 11 different bacterial genera (Fig. S2a). Protein domain analysis showed that Peripla_BP_2 emerged as the only domain shared across all siderophore-binding proteins (Fig. S2b). The Peripla_BP_2 domain (PFAM: PF01497) functions as a sensor domain in bacterial periplasmic binding proteins and is known to interact with diverse ligands, including Bacillibactin, enterobactin, vitamin B12, and heme[44]. However, due to current data limitations, the identification of SBPs responsible for siderophore recognition requires investigation in larger datasets.

We extended the analysis to the five representative genera we analyzed in the last section. Previous reports showed that the Pfam database contains at least 33 substrate-binding protein (SBP) domains[34, 35]. In this study, genes encoding these domains were collectively defined as “SBP genes”. Among 11,095 siderophore BGCs from five genera, 18 of the 33 SBP domains were detected (Table S1). Further analysis showed that each SBP protein harbors only one SBP-related domain. Accordingly, SBPs sharing the same domain were classified into the same subtype, e.g., genes containing the Peripla_BP_2 domain were designated as “Peripla_BP_2-type SBP genes” (abbreviated as PBP2 genes). Peripla_BP_2, Peripla_BP_3, SBP_bac_1, SBP_bac_5 and the OpuAC are among the most frequently appearing domains.

Co-localization within a BGC was first utilized to assess coevolution (Fig. 2a, left panel), inspired by co-localization of cognate receptors in Gram-negative siderophore BGCs[16, 45]. Focusing on the 16 types of high-fidelity siderophore BGCs from Section 1, we selected up to ten instances from each BGC type to ensure balanced representation, prioritizing those with higher purity (i.e., BGCs predicted to encode siderophores only, as antiSMASH may sometimes merge adjacent but functionally unrelated clusters into a single BGC) and similarity scores (due to limitations in available genome, some BGC types had fewer than ten instances). In total, 128 siderophore BGCs were selected for further analysis.

We extracted all gene classes within these 128 BGCs by their domains, excluding synthetases and calculated their occurrence frequencies (Fig. 2a, left panel). PBP2 genes were the most prevalent class, present in 90.1% of the BGCs (Fig. 2a, right panel). To quantify co-evolutionary strength[16], we calculated Pearson correlations between the sequence distance matrices of each gene class and their co-localized siderophore synthetase genes. PBP2 exhibited a strong correlation with siderophore synthetase genes (Pearson's r = 0.90, Fig. 2a, right panel). This correlation was significantly stronger than that observed between siderophore synthetase genes and bacterial phylogenetic relationships (Pearson's r = 0.58), indicating a specific functional linkage beyond merely phylogenetic association (Fig. 2b).

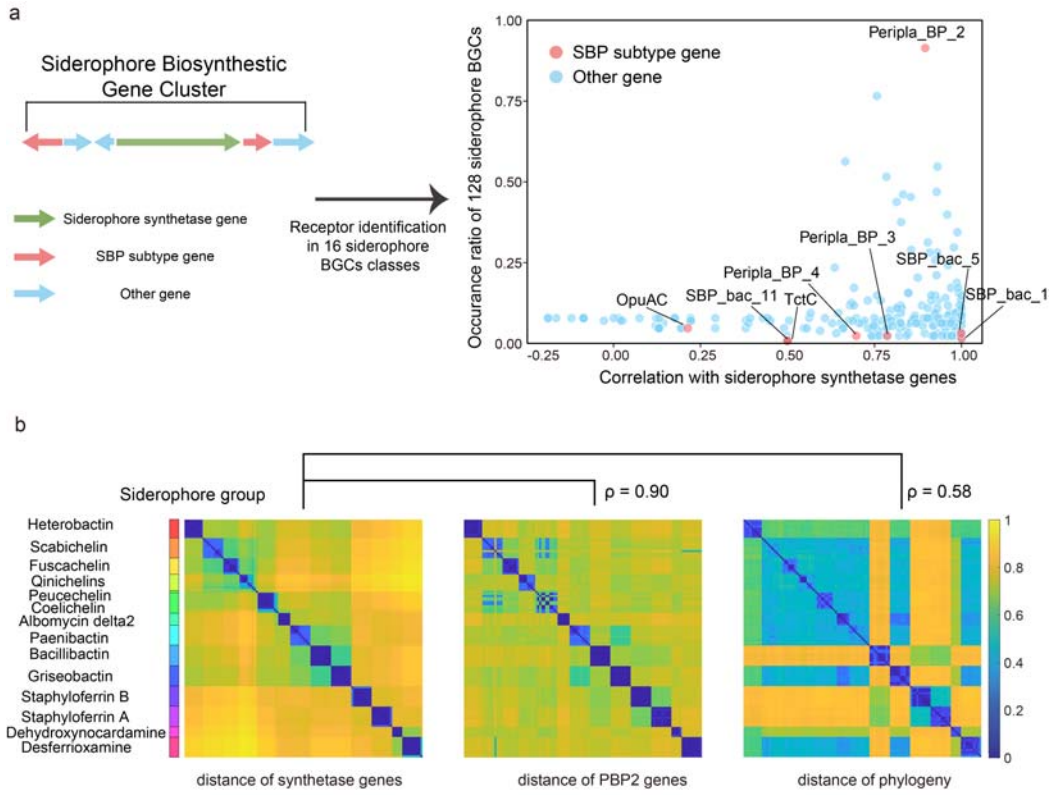

Figure 2. PBP2-subtype SBPs function as siderophore recognizer in five Gram-positive bacterial genera.

a. Identification of potential recognizers within 16 experimentally validated siderophore BGCs. The left panel illustrates the typical configuration of a BGC, with gene functions indicated by colors. In the right panel, each dot represents a gene class, grouped based on shared Pfam domain annotations. The x-values indicate the Pearson correlation between the sequence distance matrices of each gene class and their co-localized siderophore synthetase genes. The y-value reflects the occurrence ratio of each gene class across the siderophore BGCs.
b. Heatmap visualizing sequence distance matrices correlations between siderophore synthetase and PBP2 genes within the same BGCs, and between siderophore synthetase genes and phylogenetic genes. This label was assigned based on high-confidence similarity (≥80%) between predicted siderophore BGCs and known clusters in the MIBiG database.

3. Structural investigation of the most informative residues of siderophore recognizers in *Bacillus*

After identifying PBP2 as a potential siderophore recognizer, we investigated sequence regions contributing to its specificity. To minimize the phylogenetic effect, we focused our analysis on *Bacillus*, where six experimentally validated siderophore-binding PBP2 genes have been reported[21, 46]: FatB, FpuA, and YclQ (all three recognizing Petrobactin); FeuA (for Bacillibactin); YfiY (for schizokinen); and YxeB (for ferrichrome). Notably, while FatB and FpuA both bind Petrobactin, they target distinct moieties: FatB the 3,4-catecholate groups, and FpuA the citrate residues[47]. In contrast, YclQ predominates in non-Petrobactin-producing strains, likely serving for siderophore piracy[46]. Given the sequence divergence from YclQ, FatB, and FpuA, we considered these three Petrobactin recognizers as distinct PBP2 groups in our analysis.

To identify sequence regions most informative for recognizer specificity, we computed mutual information (MI) between each alignment site and the recognizer group label, with higher MI values denoting stronger associations with recognizer type. Mapping these residues to the FeuA crystal structure (PDB: 2WHY), we observed that high-MI residues are distributed across the sequence, indicating no single contiguous region dominates functional importance (Fig. 3a).

Notably, our mutual information (MI) analysis pinpointed high-scoring residues overlapping previously validated functional sites [20]. Four of five mutagenesis-confirmed residues critical for FeuA-Bacillibactin binding (K84, K105, R180, K213) ranked in the top 15% of MI scores, aligning computational predictions with experiments. The fifth (R178) scored lower, consistent with its indirect role modulating binding via R180 interaction rather than direct ligand contact. Based on these observations, we defined the top 15% of MI-ranked residues (43 sites in total) as “feature sites” of the receptor. Among them, eight residues are located within 5 Å of

the ligand in the FeuA-Bacillibactin complex structure, suggesting direct physical interaction (Fig. 3b). The remaining feature sites likely influence binding indirectly through transitive interactions.

We then compared clustering performance using full-length versus feature-site-only sequences. Pairwise sequence distances were computed, and silhouette scores assessed clustering quality against functional labels. Full-length sequences yielded a moderate score of 0.72, whereas feature sites improved it to 0.85. Phylogeny-based clustering produced near-random scores, confirming recognizer grouping reflects functional specialization, not phylogenetic relatedness.

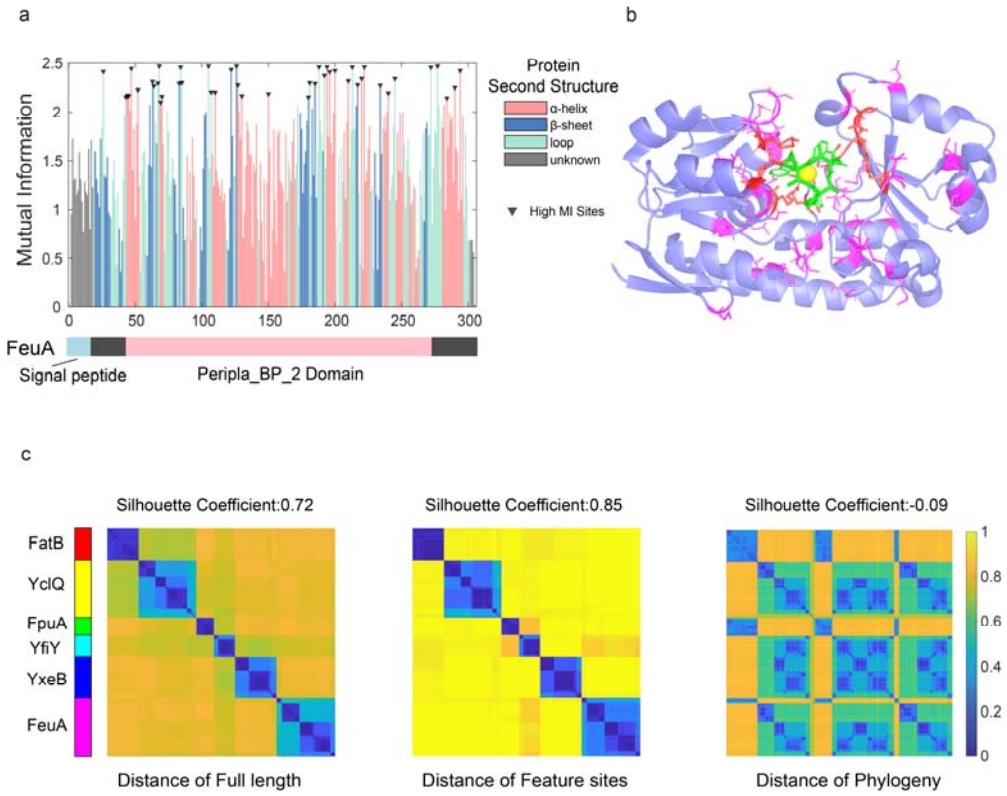

Figure 3. Structural investigation of the most informative residues of siderophore recognizers in *Bacillus*.

a. Mutual information analysis of 889 siderophore recognizer sequences from *Bacillus*. Residues falling within the top 15% of mutual information scores (indicating high correlation with receptor grouping) are marked with black inverted triangles and defined as "feature sites." These residues are mapped to the crystal structure of FeuA from *Bacillus* subtilis bound to Bacillibactin (PDB: 2WHY).

b. Crystal structure of FeuA in complex with Bacillibactin. Iron atoms are shown in yellow, Bacillibactin in green. Feature sites are highlighted in pink and red, with red indicating residues that directly interact with the ligand.

c. Heatmaps depicting the hierarchical clustering of 889 siderophore recognizers based on pairwise distances calculated using full-length sequences (left) and feature sites only (middle). The right heatmap displays the phylogenetic distance between the

corresponding strains based on concatenated alignment of 400 single-copy conserved genes. Strain clustering in all three panels is based on full-length receptor sequences. Silhouette coefficient calculated based on receptors' label.

4. Abundance, distribution, and iron-responsive regulation of PBP2 genes

Confident that PBP2 genes serve as primary siderophore recognizers, we examined their genome-wide distribution across the 5,940 genomes. HMMER screening revealed PBP2 genes exist in all genomes, ranging from 1 to 49 copies per genome, with over 90% harboring 4-12 copies (Fig. 4a). Notably, *Rhodococcus erythropolis* and *Rhodococcus qingshengii* stand out with an exceptionally high average of 46.37 PBP2 genes per genome (Fig. 4a). These two species are commonly found in iron-limited environments such as arsenic contaminated soil, Antarctic soil, and weathered serpentine rocks [48-50]. The abundance of PBP2 genes in these species likely reflects an adaptive strategy to maximize iron acquisition by capturing a broad range of xenosiderophores present in such challenging habitats.

Inspired by receptor-BGC colocalization in Gram-negative bacteria, we analyzed PBP2 genomic positions relative to siderophore BGCs, uncovering three patterns(Fig. 4b): (1) Consistent colocalization, where every BGC of a type includes a PBP2 receptor (e.g., Heterobactin in *Rhodococcus*); (2) Partial colocalization, where only some BGCs contain a PBP2 gene (e.g., ~78% of Desferrioxamine BGCs , and ~15% of Bacillibactin BGCs); and (3) Complete absence, where the BGC completely lack a PBP2 gene within the cluster (e.g., Petrobactin BGCs).

Desferrioxamine, a crucial siderophore in *Streptomyces*, exists as linear Desferrioxamine B and cyclic Desferrioxamine E. Both DesE and CdtB receptors recognize these forms, though DesE exhibits ~100-fold higher affinity for Desferrioxamine E than CdtB[51]. Among 5,940 analyzed genomes, 987 harbor a Desferrioxamine BGC, with 771 also containing *desE* gene. The *desE* gene typically colocalizes with the Desferrioxamine BGC (within 6 kb), with only two genomes showing a distance >1 Mb. In contrast, *cdtB* is encoded in only 384 Desferrioxamine -containing genomes and never clearly colocalizes with the Desferrioxamine BGC (Fig. 4b-c). For Petrobactin BGC, *fatB* and *fpuA* are generally >1 Mb from the BGC (Fig. 4b). Nearly all Petrobactin-containing strains encode *fatB*, whereas only ~75% carry *fpuA* (Fig. 4c). The contrasting genomic distributions of siderophore BGCs and their cognate recognizers prompt investigation into their co-regulation.

Ferric uptake regulator (Fur in *Bacillus*) and divalent metal-dependent regulatory protein 1 (DmdR1 in *Streptomyces*) centrally govern iron homeostasis [52, 53]. Using position weight matrix (PWM) based scanning, we identified potential Fur/DmdR1 binding sites within BGCs and upstream of recognizer genes. For Bacillibactin, 1,356/1,357 BGCs and 1,288 *feuA* loci harbor Fur sites, indicating co-regulation. For

Petrobactin (459 strains), 92% of BGCs and all *fatB* promoters had Fur sites, but only 1% of *fpuA* loci did, suggesting alternative or constitutive control. For Desferrioxamine, *desE* is typically within the BGC, facilitating co-regulation. In the only two distant (>1 Mb) *desE* cases, both *desE* and the Desferrioxamine BGC remain DmdR1-regulated. Conversely, *cdtB* showed weaker DmdR1 association: ~4% co-regulation when *desE* is present. Strikingly, in *desE*'s absence strains, 86% of *cdtB* loci and all Desferrioxamine BGCs are DmdR1-controlled. These findings collectively suggest DesE is the primary Desferrioxamine recognizer, with CdtB serving as an alternative recognizer in *desE*-lacking strains.

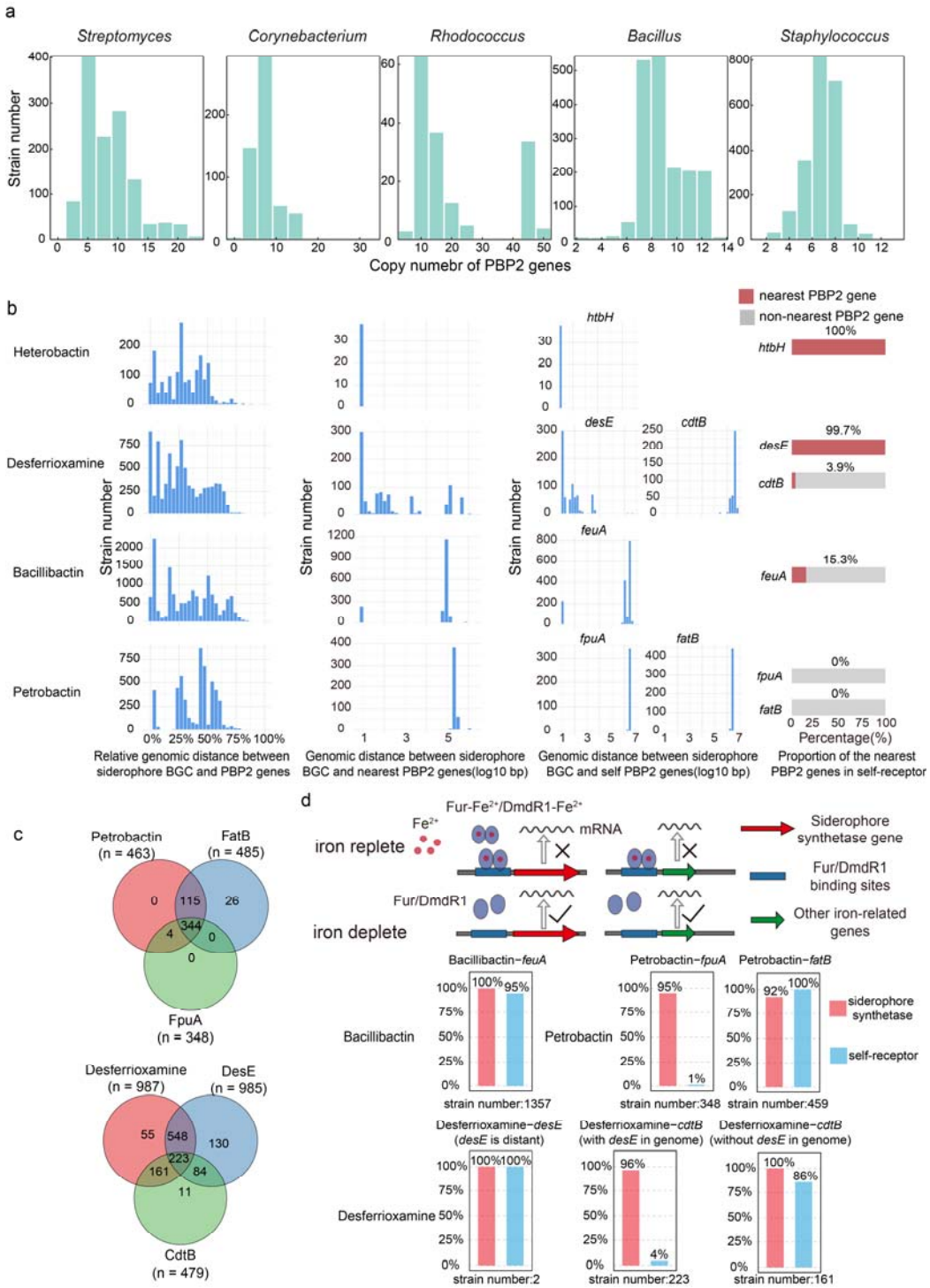

Figure 4. **Abundance, Distribution, and Iron-Responsive Regulation of PBP2 Genes**

a. Distribution of PBP2 gene number among five genera. The x-axis represents the number of PBP2 genes, while the y-axis shows the number of strains.

b. From left to right: (i) genomic distribution of all PBP2 genes with siderophore BGCs set as the reference points; (ii) distribution of the genomically distances

between siderophore BGCs and their nearest PBP2 genes(log10 scale); (iii) distribution of the genomically distances between self-receptors and siderophore BGCs(log10 scale); and (iv) proportion of self-receptors that are also the nearest PBP2 genes to siderophore BGCs.

c. Venn diagram showing the overlap between strains harboring siderophore biosynthetic genes and those harboring receptor genes.

d. Proportion of siderophore synthetase and receptor genes regulated by Fur/DmdR1 protein. The diagram illustrates the Fur/DmdR1 protein regulation mechanism.

Discussion

In an age of rapid genome sequencing and AI-powered tools, we now have powerful ways to uncover microbial interactions at scale, especially in iron acquisition systems that shape community behavior [4]. In Gram-negative bacteria, patterns of co-evolution between siderophore synthetases and receptors have made it possible to build computational models of interaction networks, highlighting differences tied to pathogens[16]. Yet, these advances rely on systematic characterization of TonB-dependent siderophore receptors in *Pseudomonas*[27], which do not exist in Gram-positive bacteria. Instead, Gram-positive bacteria lack an outer membrane and depend on substrate-binding proteins (SBPs) linked to ABC transporters for siderophore uptake, a setup that has left their recognition systems poorly mapped. This gap has slowed efforts to automatically construct iron networks for Gram-positive species, limiting our insight into how real microbial communities interact by iron.

In this study, we mined 5,940 genomes from five key Gram-positive genera to pinpoint Peripla_BP_2 (PBP2) as the main siderophore-binding SBP. It appears in 90.1% of siderophore biosynthetic gene clusters and shows strong co-evolution with synthetase genes (Pearson correlation r=0.91), far beyond what phylogeny alone explains (r=0.58). We also mapped 43 scattered feature sites in Bacillus PBP2 that drive specificity, with just eight directly touching the ligand. Beyond facilitating network reconstruction, these findings illuminate evolutionary divergences between Gram-positive and Gram-negative iron acquisition strategies, prompting deeper reflections on their independent adaptations post-divergence.

Some studies suggest that the single membrane of Gram-positive cells likely resemble that of the last universal common ancestor [54, 55], then the outer membrane of Gram-negative bacteria evolved later as an adaptive trait [56]. Still, this field relies heavily on phylogenetic inference and remains contentious [57, 58]. More consistent analysis based on molecular clocks have suggested that Gram-positive and Gram-negative groups split around 2.2 to 3.2 billion years ago[59], well before widespread oxygen buildup[60]. The Great Oxidation Event, which spiked atmospheric oxygen around 2.4 to 2.5 billion years ago and oxidized iron to less available forms, came after this split[61]. Siderophores likely diversified broadly post-GOE to scavenge iron[62], and the distinct recognizer properties in Gram-positive and Gram-negative systems reflect separate evolutionary paths and selection pressures.

The first notable difference between Gram-positive and Gram-negative siderophore recognition systems lies in the arrangement of specificity-determining sites. In Gram-negative bacteria such as *Pseudomonas*, these sites cluster in the plug domain and loop L7 of TonBDR, governing pyoverdine selectivity through direct ligand interactions and conformational changes [25-27]. In contrast, our analysis of *Bacillus* PBP2 proteins reveals 43 feature sites scattered across the sequence, with only 19% (eight sites) positioned within 5 Å of the ligand for direct contact. From the perspective of iron-net construction, this dispersion complicates co-evolutionary

analyses by hindering the identification of compact motifs for sequence-based pairing of synthetases and receptors[16]. From the perspective of evolution, the absence of a tight binding hotspot in PBP2 may reflect a more flexible recognition mechanism involving allosteric effects or conformational shifts, enabling Gram-positive receptors to accommodate a broader range of siderophores, allowing promiscuity. This interpretation aligns with previous case reports of crossover recognition in Gram-positive bacteria: in *Bacillus subtilis*, FeuA binds and imports both its cognate catecholate siderophore bacillibactin and the structurally similar xenosiderophore enterobactin[20, 33], while in *Bacillus cereus*, FpuA and FatB interact with both apo- and ferric-petrobactin[29, 32].

Another key divergence appears in the genomic organization of siderophore systems. In Gram-negative bacteria like *Pseudomonas*, cognate TonBDRs are nearly always co-localized with their siderophore biosynthetic gene clusters, aiding tight evolutionary coupling [16, 45]. In Gram-positive bacteria, however, PBP2 genes frequently lie far from these clusters, as seen in petrobactin and bacillibactin pathways, yet maintain coordination through shared iron regulators like Fur or DmdR1. One possible explanation for this differential spatial pattern is that Gram-negative bacteria may undergo horizontal gene transfer (HGT) more often than Gram-positive ones[63-65], therefore gram-negative species on clustered genes for stable inheritance. Meanwhile, Gram-positives, with lower HGT, could evolve decoupled systems that depend on robust, iron-responsive transcription to link distant parts. Such spatial flexibility may allow Gram-positive bacteria mix receptors across clusters for added versatility. For example, in desferrioxamine pathways, the main receptor DesE usually sits within the cluster under DmdR1 control, but the backup CdtB—distant and multi-specific for siderophores like desferrioxamine and coelichelin—steps in when DesE is missing, acting as a safety net for strains facing variable iron stress or acting as cheaters[51]. Similarly, petrobactin's FatB and FpuA show species-specific roles, with FpuA vital for growth in Bacillus anthracis under low iron despite weak Fur ties, hinting at layered regulation[66, 67]. This modularity might buffer against environmental shifts, contrasting the more rigid setups in Gram-negatives.

Despite these insights, our study has several limitations. First, we focused on five representative Gram-positive genera, which may not fully capture the diversity across all Gram-positive bacteria, especially underrepresented lineages. Second, the mutual information analysis for feature sites was limited to Bacillus, potentially overlooking genus-specific variations in PBP2 specificity. Third, while our bioinformatics predictions align with known structures and mutagenesis data, experimental validations such as binding assays or crystallography are needed to confirm the roles of scattered feature sites and promiscuity. Finally, transcriptional regulation inferences relied on PWM scanning, which may miss nuanced or species-specific motifs.

Looking ahead, our systematic characterization of PBP2 deepens insights into how Gram-positive and Gram-negative bacteria independently adapted their iron acquisition strategies to post-GOE iron scarcity. It also bridges a critical gap in Gram-positive iron acquisition, enabling genome-based inference of siderophore-mediated interaction networks in broader microbial communities.

Method

Data collection and taxonomy detection

We collected 46,362 complete genomes from the NCBI RefSeq database as of January 21, 2025. To ensure a consistent and genome-informed taxonomic classification, we assigned taxonomy using the Genome Taxonomy Database (GTDB). Among these genomes, 1,341 were classified as *Streptomyces*, 170 as *Rhodococcus*, 556 as *Corynebacterium*, and 2,195 as *Staphylococcus*. The classification of *Bacillus* is more complex. Traditional taxonomy includes species such as *Bacillus subtilis*, *Bacillus thuringiensis*, *Bacillus anthracis*, and *Bacillus cereus* under a single genus. However, GTDB reclassifies these organisms into several distinct genera, many of which retain the prefix "*Bacillus*". To enable a comprehensive analysis of the *Bacillus* group, we aggregated all GTDB-defined genera beginning with the name "*Bacillus*" and collectively referred to them as *Bacillus* in this study, totaling 1,861 genomes.

Construction of phylogeny tree

The phylogenetic tree depicted in Figure.1c was constructed using the PhyloPhlAn3 pipeline. PhyloPhlAn is a comprehensive pipeline that encompasses the identification of phylogenetic markers, multiple sequence alignment, and the inference of phylogenetic trees. In this analysis, we used over 400 universal genes defined by PhyloPhlAn as our selected phylogeny markers. Subsequently, the taxonomic cladogram was generated using the R package "ggtree"

Clustering analysis to group siderophore synthetic gene

Firstly, we extracted all genes annotated as core biosynthetic genes from the siderophore BGCs. We then calculated the pairwise sequence distances using p-distance, defined as the proportion of differing positions between two sequences relative to their total length. Next, we applied hierarchical clustering to group these genes, determining the optimal clustering threshold based on silhouette scores and the number of resulting groups. To account for phylogenetic effects, core biosynthetic genes from each genus were clustered separately.

Genome mining to detect BGCs and annotated siderophore BGC

We performed genome mining using antiSMASH v7.0.0 on annotated genomes to identify secondary metabolite biosynthetic gene clusters (BGCs). BGCs classified as either "NRP-metallophore" or "NIS-siderophore" were considered siderophore BGCs, corresponding to NRPS-type and NIS-type biosynthetic pathways, respectively (Fig.1c, d). To evaluate whether the identified BGCs are likely to produce known compounds, we used the knownclusterblast results provided by antiSMASH. A stringent similarity cutoff of ≥80% was applied to tentatively assign BGCs as encoding known secondary metabolites based on similarity to entries in the MIBiG database (Fig. 1e).

Calculation of the correlation between synthetase and candidate genes located within siderophore BGC sequence distance matrices

To identify potential siderophore receptors, we quantified the coevolutionary relationships between genes located within siderophore BGCs and their corresponding biosynthetic genes. We first performed hierarchical clustering of the siderophore biosynthetic genes based on sequence distance. Using the same order derived from this clustering, we then reordered the pairwise sequence distance matrix (calculated using p-distance) for all candidate genes within the BGCs. Finally, we calculated the Pearson correlation between the two matrices as a measure of coevolutionary association.

Collection and domain analysis of known siderophore SBPs

We manually curated a set of 22 experimentally validated siderophore substrate-binding proteins (SBPs) from diverse genera based on literature review. Protein sequences of these receptors were obtained from the PDB and NCBI databases. Multiple sequence alignment of the 22 protein sequences was performed using Clustal Omega. To infer the phylogenetic relationships among these SBPs, we constructed a phylogenetic tree using the Neighbor-Joining (NJ) method based on pairwise amino acid distances calculated with the WAG substitution model. The resulting tree was midpoint-rooted to better reflect evolutionary directionality. To identify conserved domains, we used HMMER v3.0 to search against the full Pfam HMM database and analyzed the shared domain architecture among these SBPs.

Identification of siderophore receptor, multiple sequence alignment, and detection of feature sites in *Bacillus*

We first collected six experimentally validated siderophore receptors from Bacillus species based on published literature. Using BLAST, we searched for homologous sequences across 1,861 Bacillus genomes. After removing identical sequences, a total of 889 unique receptor sequences were identified, corresponding to the six receptor types. Multiple sequence alignment was performed using Clustal Omega. Mutual information (MI) was then calculated between each alignment position and receptor identity labels to evaluate the contribution of each site to receptor specificity. Secondary structure information was obtained using DSSP based on the crystal structure of FeuA. Sites ranking in the top 15% of MI scores were defined as feature sites (Fig..3a).

Genome-wide identification of PBP2 genes

Based on prior knowledge, we first filtered all protein-coding sequences in the genome to retain those with amino acid lengths between 200 and 500. We then used the hmmscan function from HMMER to search these sequences against the PBP2 domain profile (PF01497) from the Pfam database. Proteins with domain hits exhibiting a score greater than 0 were considered as potential PBP2 genes.

Identification of Fur binding sites upstream of genes

Based on 19 experimentally validated Fur binding sites from *Bacillus subtilis* reported in the literature, we used the MEME suite (5.5.7) to construct a position weight matrix (PWM) representing the Fur binding motif. Using this PWM, we applied FIMO to scan the 150 bp upstream regions of all genes across Bacillus genomes. A total of 91,708 motif occurrences were identified with a p-value threshold of < 0.001.

Acknowledgement

This work was supported by the National Key Research and Development Program of China (No. 2024YFA0919500) and National Natural Science Foundation of China (No. T2321001). LZ was supported in part by the Peking-Tsinghua Center for Life Sciences.

Conflicts of Interest

All authors of this manuscript declare that they have no conflict of interest or financial conflicts to disclose.

Ethics Statement

This article does not contain any studies with human or animal subjects performed by any of the authors.

Supplement

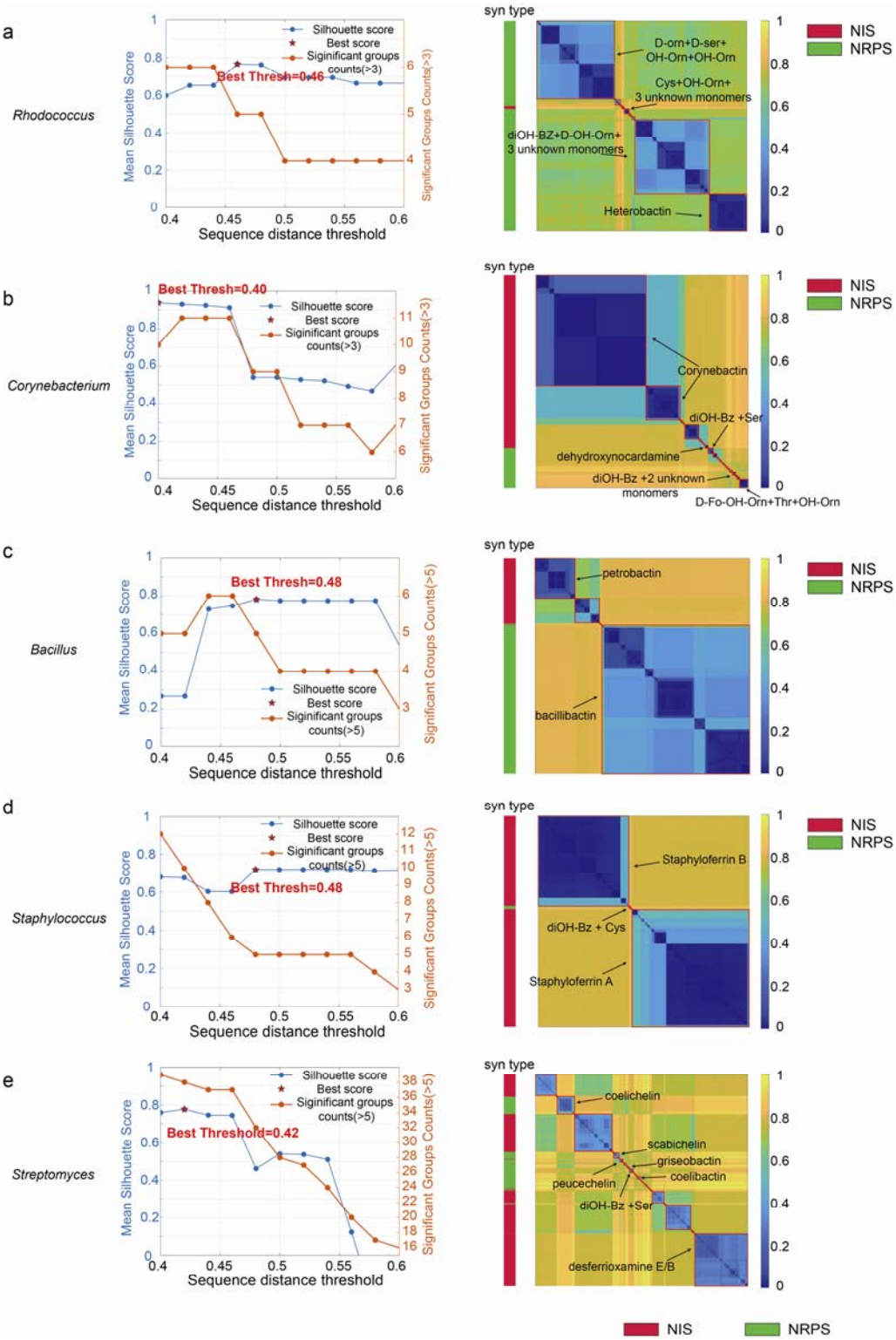

Figure S1. Clustering of Siderophore Biosynthetic Genes across Five Bacterial Genera. Left: Silhouette scores and significant group number at different clustering thresholds.

Right: Siderophore gene clustering based on synthetase gene sequence distances.

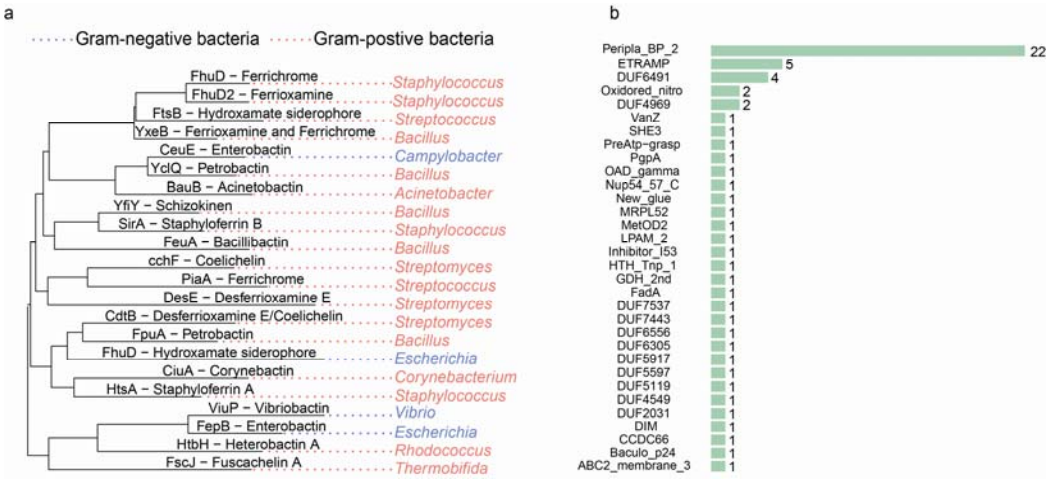

Figure S2. Domain analysis of known siderophore cytoplasmic membrane receptor a. The phylogeny tree of 22 known SBPs. b. The shared domain number among 22 known SBPs.

Table S1：Distribution of SBP domain in siderophore BGCs and their co-occurrence with other domains

SBP domain	Total number of genes containing this domain	Number of genes with domain co-occurrence	Co-occurring domains[a]
NMT1	18	0	na
NMT1_3	33	0	na
OpuAC	148	132	BPD_transp_1(100%)
PBP_like_2	10	0	na
Peripla_BP_1	13	12	Lacl(100%)
Peripla_BP_2	6023	229	ABC_tran(0.4%),FecCD(0.4%), TAT_signal(0.4%),HTH_18(98.7%)
Peripla_BP_3	184	181	GentR(0.6%),HTH_DeoR(4.4%), Lacl(95%)
Peripla_BP_4	114	9	Lacl(100%)
Peripla_BP_5	2	0	na
Peripla_BP_6	14	2	DAO(100%)
SBP_bac_1	148	0	na
SBP_bac_11	17	2	VWA(50%),Aldolase_II(50%)
SBP_bac_3	55	5	Pkinase(100%)
SBP_bac_5	221	1	TAT_signal(100%)
SBP_bac_6	27	0	na
SBP_bac_8	19	0	na
TctC	32	0	na
ZnuA	49	0	na

a. Numbers in parentheses denote the proportion of co-occurrence for each domain.

Table S2：Number of SBPs in siderophore BGCs

SBP Domain	Number (128 unbiased siderophore BGCs)	Number (7,707 high-confidence siderophore BGCs)	Number (11,095 siderophore BGCs in total)
NMT1	0	3	12
NMT1_3	0	0	33
OpuAC	6	128	148
PBP_like_2	0	8	10
Peripla_BP_1	0	8	13
Peripla_BP_2	117	4626	5571
Peripla_BP_3	3	93	161
Peripla_BP_4	3	54	82
Peripla_BP_5	0	2	2
Peripla_BP_6	0	6	14
SBP_bac_1	2	99	142
SBP_bac_11	1	4	17
SBP_bac_3	0	13	54
SBP_bac_5	4	74	187
SBP_bac_6	0	10	20
SBP_bac_8	0	12	18
TctC	1	16	32
ZnuA	0	18	49

Reference

1. Schalk, I.J., *Bacterial siderophores: diversity, uptake pathways and applications.* Nature Reviews Microbiology, 2024. **23**(1): p. 24-40.
2. Coughlan, M.P., *The role of iron in microbial metabolism.* Science Progress (1933-), 1971. **59**(233): p. 1-23.
3. Carter, A., S. Racey, and S. Veuger, *The Role of Iron in DNA and Genomic Instability in Cancer, a Target for Iron Chelators That Can Induce ROS.* Applied Sciences, 2022. **12**(19): p. 10161.
4. Gu, S., et al., *Forging the iron-net: Towards a quantitative understanding of microbial communities via siderophore-mediated interactions.* Quantitative Biology, 2025. **13**(2): p. e84.
5. Jiang, H.-B., et al., *Natural ocean iron fertilization and climate variability over geological periods.* Global Change Biology, 2023. **29**(24): p. 6856-6866.
6. Ilbert, M. and V. Bonnefoy, *Insight into the evolution of the iron oxidation pathways.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2013. **1827**(2): p. 161-175.
7. Guerinot, M.L., *Microbial iron transport.* Annu Rev Microbiol, 1994. **48**: p. 743-72.
8. He, R.L., et al., *SIDERITE: Unveiling hidden siderophore diversity in the chemical space through digital exploration.* Imeta, 2024. **3**(2).
9. Hider, R.C. and X. Kong, *Chemistry and biology of siderophores.* Natural product reports, 2010. **27**(5): p. 637-657.
10. Sarah M Barry , G.L.C., *Recent advances in siderophore biosynthesis.* Current Opinion in Chemical Biology, 2009. **13**(2): p. 205-215.

681 11. Krewulak, K.D. and H.J. Vogel, *Structural biology of bacterial iron uptake.* Biochimica
682 et Biophysica Acta (BBA) - Biomembranes, 2008. **1778**(9): p. 1781-1804.
683 12. Grigg, J.C., et al., *Specificity of Staphyloferrin B recognition by the SirA receptor from*
684 *Staphylococcus aureus.* The Journal of biological chemistry, 2010. **285**(45): p.
685 34579-34588.
686 13. Meyer, J.-M., et al., *Use of Siderophores to Type Pseudomonads: The Three*
687 *Pseudomonas Aeruginosa Pyoverdine Systems.* Microbiology, 1997. **143**(1): p. 35-43.
688 14. Grigg, J.C., et al., *Specificity of Staphyloferrin B recognition by the SirA receptor from*
689 *Staphylococcus aureus.* Journal of Biological Chemistry, 2010. **285**(45): p.
690 34579-34588.
691 15. De Chial, M., et al., *Identification of type II and type III pyoverdine receptors from*
692 *Pseudomonas aeruginosa.* Microbiology, 2003. **149**(4): p. 821-831.
693 16. Gu, S., et al., *Siderophore synthetase-receptor gene coevolution reveals habitat- and*
694 *pathogen-specific bacterial iron interaction networks.* Science Advances, 2025. **11**(3):
695 p. eadq5038.
696 17. Matzanke, B.F., et al., *Escherichia coli iron enterobactin uptake monitored by*
697 *Mössbauer spectroscopy.* Journal of Bacteriology, 1986. **167**(2): p. 674-680.
698 18. Carrano, C.J., et al., *Heterobactins: A new class of siderophores from Rhodococcus*
699 *erythropolis IGTS8 containing both hydroxamate and catecholate donor groups.*
700 Biometals, 2001. **14**(2): p. 119-125.
701 19. Grandchamp, G.M., L. Caro, and E.A. Shank, *Pirated Siderophores Promote*
702 *Sporulation in Bacillus subtilis.* Applied and Environmental Microbiology, 2017. **83**(10):

703 p. e03293-16.

704 20. Peuckert, F., et al., *Structural Basis and Stereochemistry of Triscatecholate*
705 *Siderophore Binding by FeuA.* Angewandte Chemie International Edition, 2009.
706 **48**(42): p. 7924-7927.

707 21. Hotta, K., et al., *Siderophore-mediated iron acquisition in Bacillus anthracis and*
708 *related strains.* Microbiology, 2010. **156**(7): p. 1918-1925.

709 22. Figueiredo, A.R.T., et al., *Siderophores drive invasion dynamics in bacterial*
710 *communities through their dual role as public good versus public bad.* Ecol Lett, 2022.
711 **25**(1): p. 138-150.

712 23. Timofeeva, A.M., M.R. Galyamova, and S.E. Sedykh, *Bacterial Siderophores:*
713 *Classification, Biosynthesis, Perspectives of Use in Agriculture.* Plants (Basel), 2022.
714 **11**(22).

715 24. Schalk, I.J., G.L.A. Mislin, and K. Brillet, *Chapter Two - Structure, Function and*
716 *Binding Selectivity and Stereoselectivity of Siderophore–Iron Outer Membrane*
717 *Transporters*, in *Current Topics in Membranes*, J.M. Argüello and S. Lutsenko, Editors.
718 2012, Academic Press. p. 37-66.

719 25. Greenwald, J., et al., *FpvA bound to non-cognate pyoverdines: molecular basis of*
720 *siderophore recognition by an iron transporter.* Molecular Microbiology, 2009. **72**(5): p.
721 1246-1259.

722 26. Bouvier, B., C. Cézard, and P. Sonnet, *Selectivity of pyoverdine recognition by the*
723 *FpvA receptor of Pseudomonas aeruginosa from molecular dynamics simulations.*
724 Physical Chemistry Chemical Physics, 2015. **17**(27): p. 18022-18034.

725 27. Gu, S., et al., *From sequence to molecules: Feature sequence-based genome mining*
726 *uncovers the hidden diversity of bacterial siderophore pathways.* eLife, 2024. **13**:
727 p.:RP96719.
728 28. Gu, S., et al., *Siderophore-receptor coevolution analysis reveals habitat-and*
729 *pathogen-specific bacterial iron interaction networks.* bioRxiv, 2023: p. 2023.11.
730 05.565711.
731 29. Fukushima, T., et al., *Gram-positive siderophore-shuttle with iron-exchange from*
732 *Fe-siderophore to apo-siderophore by Bacillus cereus YxeB.* Proceedings of the
733 National Academy of Sciences, 2013. **110**(34): p. 13821-13826.
734 30. Nguyen, M.-T., et al., *Lipoproteins in Gram-positive bacteria: abundance, function,*
735 *fitness.* Frontiers in microbiology, 2020. **11**: p. 582582.
736 31. Akhtar, A.A. and D.P. Turner, *The role of bacterial ATP-binding cassette (ABC)*
737 *transporters in pathogenesis and virulence: Therapeutic and vaccine potential.*
738 Microbial pathogenesis, 2022. **171**: p. 105734.
739 32. Fukushima, T., B.E. Allred, and K.N. Raymond, *Direct evidence of iron uptake by the*
740 *Gram-positive siderophore-shuttle mechanism without iron reduction.* ACS Chem Biol,
741 2014. **9**(9): p. 2092-100.
742 33. Peuckert, F., et al., *The Siderophore Binding Protein FeuA Shows Limited Promiscuity*
743 *toward Exogenous Triscatecholates.* Chemistry & Biology, 2011. **18**(7): p. 907-919.
744 34. Cerna-Vargas, J.P., et al., *Sensing preferences for prokaryotic solute binding protein*
745 *families.* Microbial Biotechnology, 2023. **16**(9): p. 1823-1833.
746 35. Chandravanshi, M., S.K. Tripathi, and S.P. Kanaujia, *An updated classification and*

747 *mechanistic insights into ligand binding of the substrate-binding proteins.* FEBS
748 Letters, 2021. **595**(18): p. 2395-2409.
749 36. Heinrichs, D.E., et al., *Staphylococcus, Streptococcus, and Bacillus.* Iron Transport in
750 Bacteria, 2014: p. 387-401.
751 37. Segond, D., et al., *Iron Acquisition in Bacillus cereus: The Roles of IlsA and*
752 *Bacillibactin in Exogenous Ferritin Iron Mobilization.* PLoS Pathogens, 2014. **10**(2).
753 38. Beasley, F.C., et al., *Staphylococcus aureus Transporters Hts, Sir, and Sst Capture*
754 *Iron Liberated from Human Transferrin by Staphyloferrin A, Staphyloferrin B, and*
755 *Catecholamine Stress Hormones, Respectively, and Contribute to Virulence ⍰.*
756 Infection and Immunity, 2011. **79**(6): p. 2345-2355.
757 39. Alam, K., et al., *Streptomyces: The biofactory of secondary metabolites.* Frontiers in
758 Microbiology, 2022. **Volume 13 - 2022**.
759 40. Blin, K., et al., *antiSMASH 7.0: new and improved predictions for detection, regulation,*
760 *chemical structures and visualisation.* Nucleic Acids Research, 2023. **51**(W1): p.
761 W46-W50.
762 41. Mohite, O.S., et al., *Pangenome mining of the Streptomyces genus redefines species'*
763 *biosynthetic potential.* Genome Biology, 2025. **26**: p. 9.
764 42. Cankar, K., N.A. Henke, and V.F. Wendisch, *Functional food additives/ingredients*
765 *production by engineered Corynebacterium glutamicum.* Systems Microbiology and
766 Biomanufacturing, 2023. **3**(1): p. 110-121.
767 43. Spohn, M., et al., *Identification of a novel aminopolycarboxylic acid siderophore gene*
768 *cluster encoding the biosynthesis of ethylenediaminesuccinic acid hydroxyarginine*

769 *(EDHA)†.* Metallomics, 2018. **10**(5): p. 722-734.

770 44. Matilla, M.A., et al., *A catalogue of signal molecules that interact with sensor kinases,*
771 *chemoreceptors and transcriptional regulators.* FEMS Microbiology Reviews, 2021.
772 **46**(1).

773 45. Crits-Christoph, A., et al., *Transporter genes in biosynthetic gene clusters predict*
774 *metabolite characteristics and siderophore activity.* Genome Research, 2021. **31**(2): p.
775 239-250.

776 46. Zawadzka, A.M., et al., *Characterization of a Bacillus subtilis transporter for*
777 *petrobactin, an anthrax stealth siderophore.* Proceedings of the National Academy of
778 Sciences, 2009. **106**(51): p. 21854-21859.

779 47. *Siderophore-Mediated Iron Acquisition Systems in Bacillus cereus: Identification of*
780 *Receptors for Anthrax Virulence-Associated Petrobactin, | Biochemistry.*

781 48. Wang, Z., et al., *Biodegradation of sodium selenite by a highly tolerant strain*
782 *Rhodococcus qingshengii PM1: Biochemical characterization and comparative*
783 *genome analysis.* Current Research in Microbial Sciences, 2025. **9**: p. 100426.

784 49. Khilyas, I.V., et al., *Genomic and phenotypic analysis of siderophore-producing*
785 *Rhodococcus qingshengii strain S10 isolated from an arid weathered serpentine rock*
786 *environment.* Archives of Microbiology, 2021. **203**(2): p. 855-860.

787 50. Wang, X., et al., *Complete genome sequence of Rhodococcus qingshengii strain R*
788 *isolated from Antarctic soil.* Microbiology Resource Announcements, 2025. **14**(6): p.
789 e01316-24.

790 51. Patel, P., L. Song, and G.L. Challis, *Distinct Extracytoplasmic Siderophore Binding*

791 *Proteins Recognize Ferrioxamines and Ferricoelichelin in Streptomyces coelicolor*
792 *A3(2).* Biochemistry, 2010. **49**(37): p. 8033-8042.
793 52. Baichoo, N. and J.D. Helmann, *Recognition of DNA by Fur: a Reinterpretation of the*
794 *Fur Box Consensus Sequence.* Journal of Bacteriology, 2002. **184**(21): p. 5826-5832.
795 53. Flores, F.J., et al., *Functional analysis of two divalent metal-dependent regulatory*
796 *genes dmdR1 and dmdR2 in Streptomyces coelicolor and proteome changes in*
797 *deletion mutants.* The FEBS Journal, 2005. **272**(3): p. 725-735.
798 54. Léonard, R.R., et al., *Was the last bacterial common ancestor a monoderm after all?*
799 Genes, 2022. **13**(2): p. 376.
800 55. Moody, E.R.R., et al., *The nature of the last universal common ancestor and its impact*
801 *on the early Earth system.* Nature Ecology & Evolution, 2024. **8**(9): p. 1654-1666.
802 56. Gupta, R.S., *Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure*
803 *rather than endosymbiosis likely led to the evolution of bacterial cells with two*
804 *membranes.* Antonie Van Leeuwenhoek, 2011. **100**(2): p. 171-182.
805 57. Witwinowski, J., et al., *An ancient divide in outer membrane tethering systems in*
806 *bacteria suggests a mechanism for the diderm-to-monoderm transition.* Nature
807 Microbiology, 2022. **7**(3): p. 411-422.
808 58. Coleman, G.A., et al., *A rooted phylogeny resolves early bacterial evolution.* Science,
809 2021. **372**(6542): p. eabe0511.
810 59. Battistuzzi, F.U. and S.B. Hedges, *A Major Clade of Prokaryotes with Ancient*
811 *Adaptations to Life on Land.* Molecular Biology and Evolution, 2008. **26**(2): p. 335-343.
812 60. Martinez-Gutierrez, C.A., J.C. Uyeda, and F.O. Aylward, *A timeline of bacterial and*

archaeal diversification in the ocean. eLife, 2023. **12**: p. RP88268.

61. Davín, A.A., et al., *A geological timescale for bacterial evolution and oxygen adaptation.* Science, 2025. **388**(6742): p. eadp1853.

62. Wade, J., et al., *Temporal variation of planetary iron as a driver of evolution.* Proceedings of the National Academy of Sciences, 2021. **118**(51): p. e2109865118.

63. Sheinman, M., et al., *Identical sequences found in distant genomes reveal frequent horizontal transfer across the bacterial domain.* eLife, 2021. **10**: p. e62719.

64. Michaelis, C. and E. Grohmann, *Horizontal Gene Transfer of Antibiotic Resistance Genes in Biofilms.* Antibiotics (Basel), 2023. **12**(2).

65. Groussin, M., et al., *Elevated rates of horizontal gene transfer in the industrialized human microbiome.* Cell, 2021. **184**(8): p. 2053-2067.e18.

66. Carlson Jr, P.E., et al., *Genetic analysis of petrobactin transport in Bacillus anthracis.* Molecular Microbiology, 2010. **75**(4): p. 900-909.

67. Lee, J.Y., et al., *Regulation of Petrobactin and Bacillibactin Biosynthesis in Bacillus anthracis under Iron and Oxygen Variation.* PLOS ONE, 2011. **6**(6): p. e20777.