Functional & Integrative Genomics (2026) 26:330
https://doi.org/10.1007/s10142-026-02109-2

BRIEF REPORT

# Co-occurring divergent arbitrium receptors outside canonical prophage regions in bacterial genomes

**José L. Balcázar**[1]

Received: 13 August 2026 / Revised: 26 October 2026 / Accepted: 28 October 2026
© The Author(s) 2026

## Abstract

Arbitrium signaling enables peptide-mediated communication among bacteriophages and regulates lysis-lysogeny decisions through diverse genetic architectures. Here, we identified two divergent AimR-like receptors in each of the alkaliphilic bacteria *Alkalihalobacterium alkalinitrilicum* DSM 22532 and *Caldalkalibacillus mannanilyticus* JCM 10596, with within-genome amino acid identities of 30.6% and 27.4%, respectively. All four receptors occurred outside predicted prophage regions, with distinct genomic neighborhoods and a putative *aimP*-like gene adjacent to one receptor. Despite substantial sequence divergence, structural comparisons revealed conservation within aligned regions of the AimR fold. Associations with mobile genetic elements suggest possible genomic mobility, although the evolutionary origins and functional roles of these receptors remain unresolved. These findings highlight the diversity of co-occurring AimR-like receptors in bacterial genomes.

**Keywords** Arbitrium signaling · AimR-like receptors · Sequence divergence · Prophage regions · Genomic organization

Bacteriophages use arbitrium signaling systems to coordinate lysis-lysogeny decisions through peptide-mediated communication (Erez et al. 2017). Initially characterized as a three-component system involving the signaling peptide AimP, the receptor AimR, and the regulatory element AimX, arbitrium signaling is now recognized to encompass more diverse genetic architectures and regulatory mechanisms (Erez et al. 2017; Brady et al. 2023; Kabel et al. 2026). Comparative genomic studies have also identified AimR homologs outside recognizable mobile genetic elements and without associated AimP peptides (Stokar-Avihail et al. 2019; Felipe-Ruiz et al. 2022). Recent experimental evidence has further demonstrated that arbitrium signaling can occur between distinct phage lineages through non-cognate peptide-receptor interactions (Gallego-Del-Sol et al. 2026; Manley et al. 2026; Lomeli-Ortega and Balcázar 2026). However, the genomic organization and structural conservation of co-occurring divergent AimR-like receptors outside recognizable prophage regions remain insufficiently characterized.

To investigate the genomic distribution of arbitrium-related components, protein similarity searches were performed against the NCBI non-redundant protein sequence database (nr), using the AimR (CAB14004.1) and AimP (CAB14003.1) proteins from *Bacillus subtilis* subsp. *subtilis* str. 168 (GenBank accession AL009126.3) as reference sequences. Initial searches were conducted using BLASTP, followed by two iterations of PSI-BLAST for both proteins to identify more divergent homologs, with a maximum of 500 sequences selected for each subsequent iteration. Searches for *aimX* were initially performed using BLASTN, followed by manual inspection of the genomic regions surrounding candidate *aimR*-like genes. Following the identification of two AimR-like homologs within the same bacterial genome, subsequent analyses focused on their sequence divergence, genomic organization, and association with prophages and other mobile genetic elements. Candidate sequences and their surrounding genomic regions were manually inspected. Prophage regions were predicted using PHASTEST (Wishart et al. 2023), whereas putative horizontally transferred regions were identified using Alien Hunter (Vernikos and Parkhill 2006).

✉ José L. Balcázar
jlbalcazar@icra.cat

[1] Catalan Institute for Water Research (ICRA-CERCA), 17003 Girona, Spain

Comparative genomic analysis of *B. subtilis* subsp. subtilis str. 168 and *B. subtilis* KRS015 (NZ_CP125762.1) revealed differences in the genomic organization of arbitrium-associated loci (Fig. 1A). In strain 168, the arbitrium locus was located within a predicted prophage region. In contrast, KRS015 contained two *aimR-aimP-aimX* modules within a 269.2-kb putative horizontally transferred region identified using Alien Hunter. Although two prophage regions were predicted in KRS015 using PHASTEST, neither encompassed these arbitrium modules. The two AimR proteins encoded by these modules shared 100% amino acid sequence identity, providing a reference for comparison with the divergent AimR-like homologs identified in alkaliphilic bacteria.

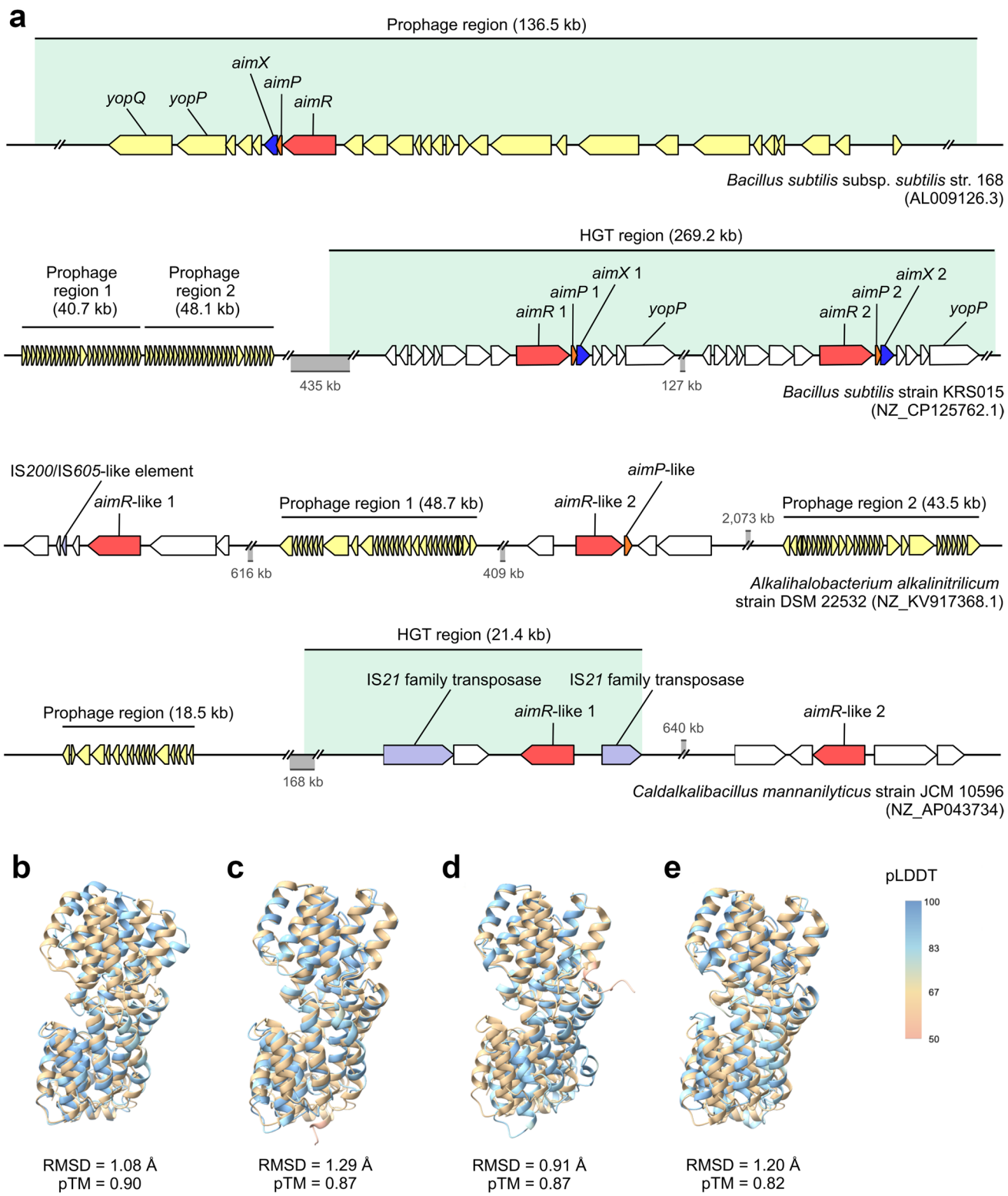
Further analysis identified two AimR-like homologs in each of the alkaliphilic bacteria *Alkalihalobacterium alkalinitrilicum* DSM 22532 (NZ_KV917368.1) and *Caldalkalibacillus mannanilyticus* JCM 10596 (NZ_AP043734) (Fig. 1A). All four homologs were located outside the prophage regions predicted by PHASTEST, although their genomic neighborhoods differed. In *A. alkalinitrilicum*, *aimR*-like 1 was located near an IS*200*/IS*605*-like element, whereas *aimR*-like 2 was immediately adjacent to a putative *aimP*-like gene. The corresponding protein was analyzed using SignalP 6.0 (Teufel et al. 2022), which predicted a Sec/SPI signal peptide with a cleavage site between residues 24 and 25. No recognizable *aimX* gene was identified in the surrounding genomic regions. In *C. mannanilyticus*, *aimR*-like 1 was located within a 21.4-kb putative horizontally transferred region predicted by Alien Hunter, near genes encoding IS*21* family transposases, whereas *aimR*-like 2 was located outside this region. Neither receptor was associated with a recognizable *aimP*-like or *aimX* gene. These observations reveal distinct genomic arrangements among the co-occurring AimR-like homologs, including the presence of a putative *aimP*-like gene adjacent to one receptor, but no recognizable *aimX* genes in their genomic neighborhoods.

Sequence comparisons revealed substantial divergence between the co-occurring AimR-like homologs, with amino acid sequence identities of 30.6% in *A. alkalinitrilicum* and 27.4% in *C. mannanilyticus*. Phylogenetic analysis was performed using MUSCLE for sequence alignment (Edgar 2004) and the neighbor-joining method implemented in MEGA12 (Kumar et al. 2024), with 1,000 bootstrap replicates. The resulting phylogeny illustrated the sequence diversity of the identified AimR-like homologs relative to the reference AimR proteins (Supplementary Figure S1).

To further assess the structural similarity between the identified AimR-like proteins and canonical AimR receptors, three-dimensional structures were predicted using AlphaFold 3 (Abramson et al. 2024) and aligned against the experimentally determined structure of the SPβ AimR receptor (PDB: 5ZW5) (Wang et al. 2018) using the MatchMaker algorithm implemented in UCSF ChimeraX (Pettersen et al. 2021). The predicted structures exhibited high confidence, with mean pLDDT values ranging from 88.84 to 92.21. Despite substantial amino acid sequence divergence, all four AimR-like proteins exhibited an α-helical architecture resembling that of the reference AimR receptor (Fig. 1B–E). Structural superposition yielded RMSD values ranging from 0.91 to 1.29 Å after pruning structurally discordant atom pairs, whereas RMSD values calculated across all initially aligned atom pairs ranged from 3.25 to 4.87 Å (Supplementary Table S1). These results indicate structural similarity within the retained aligned regions, with larger deviations observed when all aligned atom pairs were considered. Protein domain analyses using InterPro (Blum et al. 2025), Pfam (Mistry et al. 2021), CDD (Wang et al. 2023), and HHpred (Zimmermann et al. 2018) further supported their annotation as AimR-like proteins.

The heterogeneous genomic contexts of the identified AimR-like homologs suggest possible associations with

**Fig. 1** Genomic organization and structural comparison of canonical AimR and divergent AimR-like receptors. **a** Genomic organization of arbitrium-associated loci in *Bacillus subtilis* subsp. *subtilis* str. 168, *B. subtilis* KRS015, *Alkalihalobacterium alkalinitrilicum* DSM 22532, and *Caldalkalibacillus mannanilyticus* JCM 10596. In *B. subtilis* 168, the *aimR-aimP-aimX* module is located within a prophage region, whereas KRS015 contains two such modules within a 269.2-kb putative horizontally transferred region. The two AimR proteins in KRS015 share 100% amino acid sequence identity. Both alkaliphilic species harbor two divergent *aimR*-like homologs outside predicted prophage regions. In *A. alkalinitrilicum*, one homolog is located near an IS*200*/IS*605*-like element, whereas the other is adjacent to a putative *aimP*-like gene. In *C. mannanilyticus*, one homolog is located within a 21.4-kb putative horizontally transferred region containing IS*21* family transposases. Prophage regions were predicted using PHASTEST (Wishart et al. 2023), and putative horizontally transferred regions were identified using Alien Hunter (Vernikos and Parkhill 2006). Yellow arrows indicate prophage-associated genes; red arrows, *aimR* or *aimR*-like genes; orange arrows, *aimP* or putative *aimP*-like genes; blue arrows, *aimX*; purple arrows, transposases; and white arrows, neighboring genes. Double slashes indicate omitted genomic intervals, with the corresponding distances indicated in kilobases (kb). Genomic features are shown schematically and are not drawn to scale. **b**–**e** Structural superpositions of the experimentally determined SPβ AimR receptor (PDB: 5ZW5, chain A) with AlphaFold 3-predicted AimR-like proteins from *A. alkalinitrilicum* AimR-like 1b, *A. alkalinitrilicum* AimR-like 2 **c**, *C. mannanilyticus* AimR-like 1 **d**, and *C. mannanilyticus* AimR-like 2 **e**. Structures were aligned using the MatchMaker algorithm implemented in UCSF ChimeraX (Pettersen et al. 2021). The experimentally determined SPβ AimR structure is shown in beige, whereas the predicted AimR-like structures are colored according to their per-residue predicted local distance difference test (pLDDT) scores, using a shared color scale (50–100). Root-mean-square deviation (RMSD) values, calculated after iterative pruning of structurally divergent atom pairs, and predicted template modeling (pTM) scores are indicated for each superposition. The structural comparisons reveal conservation within aligned regions despite substantial sequence divergence (Supplementary Table S1)




mobile genetic elements and horizontal gene transfer. The proximity of some *aimR*-like genes to insertion sequence elements, together with the localization of one homolog within a putative horizontally transferred region, is consistent with the possibility that genomic mobility has contributed to their distribution. However, these associations were not observed for all four receptors, and the available genomic evidence does not establish whether the aimR-like genes underwent horizontal transfer. Furthermore, their location outside prophage regions predicted by PHASTEST does not exclude an

association with degraded prophages or other unrecognized mobile genetic elements.

Previous studies have identified AimR homologs outside recognizable mobile genetic elements, including receptors lacking adjacent AimP-encoding genes (Stokar-Avihail et al. 2019; Felipe-Ruiz et al. 2022). The present findings extend these observations by documenting the coexistence of markedly divergent AimR-like homologs within individual bacterial genomes, despite their structural resemblance to canonical AimR receptors. Although a putative *aimP*-like gene was identified adjacent to one receptor, its predicted signal peptide does not establish a role in arbitrium signaling, and no recognizable *aimX* genes were detected in the genomic neighborhoods analyzed. The possibility that these receptors recognize peptides encoded elsewhere or produced by other microorganisms is consistent with recent evidence of non-cognate peptide-receptor interactions among arbitrium systems (Gallego-Del-Sol et al. 2026; 2026). Nevertheless, the signaling capacity and regulatory functions of the identified receptors remain unknown.

Together, these findings highlight the diversity of genomic arrangements and the structural similarities of co-occurring, sequence-divergent AimR-like receptors outside prophage regions predicted by PHASTEST. Their evolutionary origins and functional significance remain unresolved. Future studies integrating comparative genomics, peptide-binding assays, and transcriptional analyses will be necessary to determine whether these receptors recognize distinct or shared signaling peptides or perform alternative regulatory functions.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10142-026-02109-2.

**Acknowledgements** The author acknowledges support from the CERCA Programme of the Generalitat de Catalunya.

**Author Contributions** J.L.B.: Conceptualization, Investigation, Methodology, Formal analysis, Visualization, Writing – Original Draft Preparation, Writing – Review and editing.

**Funding** Open Access funding provided by CERCA through the CRUE-CSIC agreement with Springer Nature. No funding was received for conducting this study.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Ethics approval** This study did not involve any human participants or animal subjects.

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

Abramson J, Adler J, Dunger J et al (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630:493–500. https://doi.org/10.1038/s41586-024-07487-w

Blum M, Andreeva A, Florentino LC, Chuguransky SR, Grego T, Hobbs E, Pinto BL, Orr A, Paysan-Lafosse T, Ponamareva I, Salazar GA, Bordin N, Bork P, Bridge A, Colwell L, Gough J, Haft DH, Letunic I, Llinares-López F, Marchler-Bauer A, Meng-Papaxanthos L, Mi H, Natale DA, Orengo CA, Pandurangan AP, Piovesan D, Rivoire C, Sigrist CJA, Thanki N, Thibaud-Nissen F, Thomas PD, Tosatto SCE, Wu CH, Bateman A (2025) InterPro: the protein sequence classification resource in 2025. Nucleic Acids Res 53:D444–D456. https://doi.org/10.1093/nar/gkae1082

Brady A, Cabello-Yeves E, Gallego del Sol F, Chmielowska C, Mancheño-Bonillo J, Zamora-Caballero S, Omer SB, Torres-Puente M, Eldar A, Quiles-Puchalt N, Marina A, Penadés JR (2023) Characterization of a unique repression system present in arbitrium phages of the SPbeta family. Cell Host Microbe 31:2023–2037. https://doi.org/10.1016/j.chom.2023.11.003

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, Melamed S, Leavitt A, Savidor A, Albeck S, Amitai G, Sorek R (2017) Communication between viruses guides lysis-lysogeny decisions. Nature 541:488–493. https://doi.org/10.1038/nature21049

Felipe-Ruiz A, Marina A, Rocha EPC (2022) Structural and genomic evolution of RRNPPA systems and their pheromone signaling. Mbio 13:e02514–e02522. https://doi.org/10.1128/mbio.02514-22

Gallego-Del-Sol F, Sin D, Chmielowska C, Mancheño-Bonillo J, Li Y, Zamora-Caballero S, Quiles-Puchalt N, Penadés JR, Marina A (2026) Phages communicate across species to shape microbial ecosystems. Cell 189:3025–3038. https://doi.org/10.1016/j.cell.2026.03.004

Kabel S, Omer Bendori S, Borenstein T, Guler P, Martinez-Alonso C, Mancheño-Bonillo J, Gallego del Sol F, Marina A, Eldar A (2026) A widespread extended arbitrium system controls lysis/lysogeny through antirepression. Cell Host Microbe 34:278–290. https://doi.org/10.1016/j.chom.2026.01.007

Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K (2024) MEGA12: molecular evolutionary genetic analysis version 12 for adaptive and green computing. Mol Biol Evol 41:msae263. https://doi.org/10.1093/molbev/msae263

Lomeli-Ortega CO, Balcázar JL (2026) Arbitrium crosstalk enables shared signaling among phages. Cell Rep (in press)

Manley R, Woodhams R, Arrowsmith TJ, Smith E, Bruce J, Blower TR, Temperton B, Westra ER (2026) Arbitrium phages can

manipulate each other's lysis/lysogeny decisions. Cell 189:3039–3051. https://doi.org/10.1016/j.cell.2026.02.020

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A (2021) Pfam: the protein families database in 2021. Nucleic Acids Res 49:D412–D419. https://doi.org/10.1093/nar/gkaa913

Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci 30:70–82. https://doi.org/10.1002/pro.3943

Stokar-Avihail A, Tal N, Erez Z, Lopatina A, Sorek R (2019) Widespread utilization of peptide communication in phages infecting soil and pathogenic bacteria. Cell Host Microbe 25:746–755. https://doi.org/10.1016/j.chom.2019.03.017

Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025. https://doi.org/10.1038/s41587-021-01156-3

Vernikos GS, Parkhill J (2006) Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella pathogenicity* islands. Bioinformatics 22:2196–2203. https://doi.org/10.1093/bioinformatics/btl369

Wang Q, Guan Z, Pei K, Wang J, Liu Z, Yin P, Peng D, Zou T (2018) Structural basis of the arbitrium peptide-AimR communication system in the phage lysis-lysogeny decision. Nat Microbiol 3:1266–1273. https://doi.org/10.1038/s41564-018-0239-y

Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A (2023) The conserved domain database in 2023. Nucleic Acids Res 51:D384–D388. https://doi.org/10.1093/nar/gkac1096

Wishart DS, Han S, Saha S, Oler E, Peters H, Grant JR, Stothard P, Gautam V (2023) PHASTEST: faster than PHASTER, better than PHAST. Nucleic Acids Res 51:W443–W450. https://doi.org/10.1093/nar/gkad382

Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.