PaliDIS: A tool for fast discovery of novel insertion sequences

Authors

Victoria R. Carr[1,2,*], Solon P. Pissis[3,4], Peter Mullany[5], Saeed Shoaie[2,6], David Gomez-Cabrero[2,7,8], David L. Moyes[2]

[1]Parasites and Microbes, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

[2]Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, SE1 9RT, UK

[3]Centrum Wiskunde en Informatica, Amsterdam, NL

[4]Vrije Universiteit, Amsterdam, NL

[5]Department of Microbial Diseases, Eastman Dental Institute, University College London, 256 Gray's Inn Road, London, WC1X 8LD, UK

[6]Science for Life Laboratory, KTH – Royal Institute of Technology, Stockholm, SE-171 21, Sweden

[7]Bioscience Program, Bioengineering Program, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

[8]Translational Bioinformatics Unit, Navarrabiomed, Complejo Hospitalario de Navarra (CHN), Universidad Pública de Navarra (UPNA), IdiSNA, Pamplona, Spain

*Corresponding author

Abstract

The diversity of microbial insertion sequences, crucial mobile genetic elements in generating diversity in microbial genomes, needs to be better represented in current microbial databases. Identification of these sequences in microbiome communities presents some significant problems that have led to their underrepresentation. Here, we present a software tool called PaliDIS that recognises insertion sequences in metagenomic sequence data rapidly by identifying inverted terminal repeat regions from mixed microbial community genomes. Applying this software to 266 human metagenomes identifies 11,681 unique insertion sequences. Querying this catalogue against a large database of isolate genomes reveals evidence of horizontal gene transfer events of clinically relevant antimicrobial resistance genes between classes of bacteria. We will continue to apply this tool more widely, building

the Insertion Sequence Catalogue, a valuable resource for researchers wishing to query their microbial genomes for insertion sequences.

Keywords

Insertion sequences, transposon, metagenome, horizontal gene transfer, mobile genetic element, antimicrobial resistance, software

Abbreviations

ARG – antimicrobial resistance gene
bp – base pairs
ISC –Insertion Sequence Catalogue
IS – insertion sequence/unit transposon
ITR – inverted terminal repeat
MEM – maximal exact match
PaliDIS - Palindromic Detection of Insertion Sequences

Data Summary

1. The PaliDIS software is available here: github.com/blue-moon22/palidis
2. The Insertion Sequence Catalogue is available to download here: https://github.com/blue-moon22/ISC
3. The raw reads from the Human Microbiome Project can be retrieved using the download links provided in Supplementary Data 1
4. The 21 contig files can be retrieved using the download links provided in Supplementary Data 3

Impact Statement

Insertion sequences are a class of transposable element that play an important role in the dissemination of antimicrobial resistance genes. However, it is challenging to completely characterise the transmission dynamics of insertion sequences and their precise contribution to the spread of antimicrobial resistance. The main reasons for this are that it is impossible to identify all insertion sequences based on limited reference databases and that *de novo* computational methods are ill-equipped to make fast or accurate predictions based on incomplete genomic assemblies. PaliDIS is a new software tool that is generating a larger, more comprehensive catalogue of insertion sequences based on a fast algorithm harnessing

genomic diversity in mixed microbial communities. This catalogue will enable genomic epidemiologists and researchers to annotate genomes for insertion sequences more extensively and advance knowledge of how insertion sequences contribute to bacterial evolution in general and antimicrobial resistance spread across microbial lineages in particular. This will be useful for genomic surveillance, and for development of microbiome engineering strategies targeting inactivation or removal of important transposable elements carrying antimicrobial resistance genes.

Introduction

Swapping genetic information between members of a microbial community, a mechanism referred to as horizontal gene transfer (HGT), is a key process in the microbiome. It allows for the spread of new genes and functionality throughout the community. The result of HGT can be acquisition of a new gene, duplication of an existing gene or even interruption of a current genes. The mechanisms that support HGT have been well described and involve the transfer of mobile genetic elements (MGEs). MGEs are best defined as broadly as possible, as any genetic element that can mediate its own transfer from one part of a genome to another or between different genomes. The most complex elements are conjugative plasmids and Integrative Conjugative Elements (ICEs) which can mediate their transfer between bacterial cells[1]. The simplest and most abundant MGEs are the insertion sequences which only contain enough genetic information for their own transposition. MGEs are best thought of as a continuum ranging from the relatively simple insertion sequences right up to conjugative elements and everything in between[2]. MGEs are crucially important in bacterial evolution as a result of the extensive diversity they generate, an aspect of this is their central role in the spread of antimicrobial resistance genes (ARGs) between microbial genomes.

Insertion sequences are short transposable elements between 700-2,500 bp in length containing genes that code for the proteins involved in their own transposition they are found in both chromosomes, ICEs and plasmids[3]. Most insertion sequences contain one or sometimes two genes encoding transposases, the most ubiquitous genes in prokaryotic and eukaryotic genomes[4]. Insertion sequences and transposons (transposons are defined at genetic elements that can transpose from one part of the genome to another but carry sequences other than those involved in transposition, unlike insertion sequences which just encode the genetic information for their own translocation) can be broadly classified by the amino acids in their

transposase, commonly DDE (aspartic acid, aspartic acid and glutamic acid), DEDD or HUH (two histidine residues separated by any large hydrophobic amino acid) motifs, and their mechanism of transposition (either conservative or replicative)[5]. Common DDE insertion sequences contain two inverted terminal repeats (ITRs) at each end of a 10-50 bp size DNA sequence that are reverse complement sequences of each other. Some insertion sequences are flanked by unique shorter direct repeat sequences, also known as target site duplications (TSDs), which are formed by the duplication of the insertion sequence target site upon insertion[3]. Unit transposons are a similar type of transposable element to insertion sequences containing a pair of ITRs but can also carry ARGs as well as transposases. For simplicity, the abbreviation "IS" will be used hereafter to mean insertion sequence or unit transposon. ARGs can also be carried by composite transposons that are bounded by two copies of two different insertion sequences which can move together in a single unit. A composite transposon can contain one or more passenger genes, such as ARGs, flanked by two insertion sequences and with two TSDs at both ends.

Microbial genomes can be annotated for ISs by querying reference databases of known transposable elements, such ISfinder[6], but these databases are small and do not represent many transposable elements in nature. As transposable elements are the most ubiquitous and abundant MGE, it is a continual effort to catalogue them all using common methods. Novel ISs containing ITRs can be detected using computational tools, such as EMBOSS[7], that search for palindromic sequences representing ITRs[8]. However, transposable elements in isolated genomes that are assembled from short reads can be misassembled or incomplete, since assembly algorithms struggle to resolve repeated elements[9]. Additionally, ITR pairs are not typically exact reverse complements, and algorithms that only detect perfect palindromes may fail to identify many insertion sequences. Alternatively, novel ISs can be identified by manually searching for ITRs or flanking regions of interest (such as ARGs) using a genome browser, but this can be a difficult and tedious process. Alternatively, Hidden Markov Models (HMMs) have been used to identify transposases within these elements, include those without ITRs[8]. However, the presence of a transposase is not sufficient evidence for a transposition event to have occurred.

In this paper, we present a tool called PaliDIS (Palindromic Detection of Insertion Sequences) that finds ISs using an efficient maximal exact matching algorithm to identify ITRs across different genomic loci in reads sequenced from mixed microbial communities.

These ISs can then be pooled and clustered to create a non-redundant catalogue of ISs. PaliDIS can also predict the origins of these ISs by querying the them against ISfinder or a COmpact Bit-sliced Signature (COBS) index[10] of 661,405 microbial genomes[11]. Here, we present the theory and implementation of this tool on 266 short read metagenomes to generate 11,681 unique ISs included in the first release of the Insertion Sequence Catalogue (ISC). Beyond this paper, PaliDIS will continue generating ISs to expand ISC.

Theory and Implementation

PaliDIS is implemented as a Nextflow pipeline with all dependency software packaged in one container image. The input file of PaliDIS is a tab-delimited manifest text file that contains information on the read file IDs, the file paths to the read fastq.gz files, sample ID and file paths to the assemblies. The output files are a FASTA file of ISs and accompanying tab-delimited file of information. The following steps are also illustrated in Figure 1.

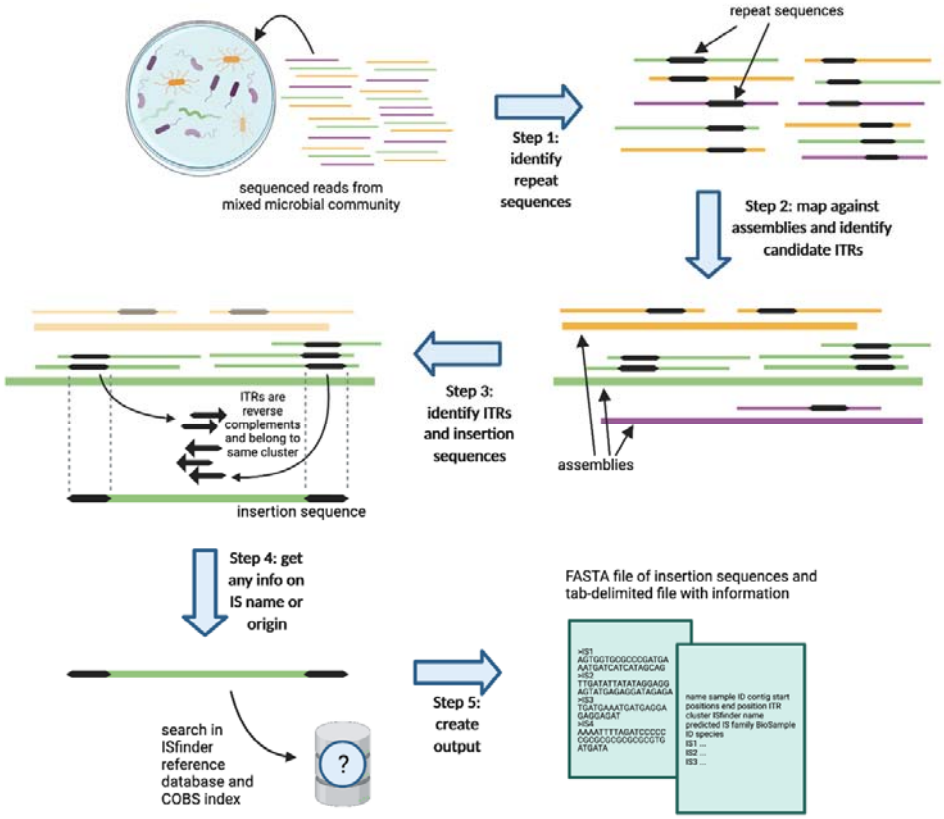


Figure 1. Steps summarising the PaliDIS software. Step 1: Reads from mixed microbial communities are pre-processed and run through an algorithm to identify reads containing repeat sequences. Step 2: Reads containing repeat sequences are mapped against the assemblies to find their positions and proximity filters applied to identify candidate ITRs. Step 3: Candidate ITRs are clustered. ISs are identified by ITRs that are of the same

cluster and are reverse complements of each other. Step 4 ISs are queried against existing databases to identify known ISs and predict their origin. Step 5: Final outputs of a FASTA file with insertion sequences and tab-delimited file with information are created.

Step 1: Reads from mixed microbial communities are pre-processed and run through an algorithm to identify reads containing repeat sequences

Firstly, the FASTQ files are converted to FASTA files with headers prepended with their sequence order (e.g. Seq1, Seq2 etc.). A software tool, called pal-MEM (https://github.com/blue-moon22/pal-mem), was developed and applied an efficient maximal exact matching algorithm[12] to identify repeat sequences that may represent ITRs. A maximal exact match (MEM) between two strings is an exact match (i.e. an exact local alignment), which cannot be extended on either side without introducing a mismatch (or a gap).

Preparing the reference and query data structures

pal-MEM creates a reference hash table from the sequences for some integer $k>0$ defined by the user, in which k-mers are the keys and the corresponding occurring positions are their values. The nucleotides of k-mers are encoded as unique combinations of two bits (0 and 1), (where A is 00, C is 01, G is 10 and T is 11), reducing memory requirements. In addition, it is not required for all k-mers to be stored, reducing the demand on memory further. A k-mer is stored only when it has a position that is a multiple of $(L - k) + 1$ (where k is the length of the k-mer and L is the minimum ITR length), i.e.

(eq. 1) $b_r \leq j((L - k) + 1) \leq e_r - k + 1$

where b_r and e_r are the start and end positions of a maximal exact match (MEM) and $j \geq 1$. The sequences are then also used to create a query data structure of unsigned 64-bit integers representing blocks of 32 nucleotides where each nucleotide is represented by two bits (A is 00, C is 01, G is 10 and T is 11). Random 20-bit sequences are stored between the array of reads define their boundaries. The start and end positions for each read and random sequence are stored in another data structure.

Applying the algorithm to find repeat sequences

Each k-mer from the query read is looked up against the reference hash table to retrieve a matching k-mer. The first k-mer window starts from the beginning of the query and continues to shift every two bits, but skips the positions within the random sequences. These matching

k-mers are then extended in both directions to make larger sequence matches until mismatches disrupt the extension, making a MEM. The algorithm performs this process using an interval halving approach. The sequence is extended to the left end position of the shortest of the two sequences. If there is no match, the extension is halved until a match is made. The extension is elongated by one nucleotide at a time until no more exact matches can be made. This is repeated on the right side. A repeat sequence is found once a MEM has a length greater than or equal to the minimum ITR length and less than or equal to the maximum ITR length as defined by the user. If a repeat sequence is found, pal-MEM will move on to the next read in the query, given it is expected that a read from short-read sequencing would contain only one ITR.

Dealing with technical repeats from amplified read libraries

Read libraries are dominated by technical as well as biological repeated sequences that are the result of sequencing amplified regions. To reduce the frequency of technical repeats being identified as biological repeats, MEMs are also excluded if their start or end positions are within a buffer length of 20 nucleotides (40 bits) from either end of the read. This model represents an alignment of the prefix or suffix of a read typical of a technical repeat.

Step 2: Reads containing repeat sequences are mapped against the assemblies to find their positions and proximity filters applied to identify candidate ITRs

Reads containing repeat sequences identified in Step 1 are mapped using Bowtie2[13] against their associated assemblies. A Python script uses the output of Bowtie2 to identify mapped reads with candidate ITRs where the positions of the repeats are located between the minimum and maximum IS length as defined by the user.

Step 3: Candidate ITRs are clustered and ISs are identified by ITRs that are of the same cluster and are reverse complements of each other

The candidate ITRs are clustered using CD-HIT-EST[14] where nucleotide sequences that meet a 1) sequence identity threshold *c*, 2) a global *G* 1 or local alignment *G* 0, 3) alignment coverage for the longer sequence *aL*, 4) alignment coverage for the shorter sequence *aS* and 5) minimal alignment coverage control for the both sequences *A* (that can be specified by the user). The ISs are generated in a FASTA format with an accompanying tab-delimited file containing the sample ID, assembly name, start and end positions of the ITRs and their

cluster using a Python script. The ISs must contain ITRs that 1) belong to the same cluster, 2) are within the minimum and maximum specified ITR length, 3) are within the minimum and maximum IS length, and 4) are reverse complements of each other where the two sequences aligned using BLASTn[15] (with parameters *-task blastn -word_size 4*) have "*Strand=Plus/Minus*" and "*Identities"* greater than or equal to the specified minimum ITR length.

Step 4: ISs are queried against existing databases to identify known ISs and predict their origin

ISs are queried against a non-redundant database of ISs from ISfinder[6] in 2020 using BLASTn (as documented here: https://github.com/blue-moon22/PaliDIS/tree/master/db/ISfinder-sequences). An IS that is a match with an ISfinder sequence is assigned as being a complete homolog if the alignment has an identity and a coverage of at least 99 %. Otherwise, the IS is assigned a predicted IS family. The origin of these ISs can be found by searching using cobs query[10] against a COBS index of microbial genomes with NCBI BioSample IDs. The taxonomy of those genomes containing these ISs are found by querying the BioSample IDs using a metadata retrieval tool, ffq[16] (https://github.com/pachterlab/ffq).

Step 5: Final output

A Python script generates a FASTA file of ISs and a tab-delimited file of information including their header name, sample ID, contig, start and end positions of their ITRs on their contigs, ITR cluster, ISfinder name, predicted IS family from ISfinder, BioSample ID and origin species.

First release of the Insertion Sequence Catalogue using PaliDIS

A catalogue of insertion sequences was generated using PaliDIS applied to 266 human oral and gut metagenomic reads from the Human Microbiome Project (Supplementary Data 1)[17]. The reads were quality controlled, filtered and assembled as previously described[18]. A total of 25,650 ISs were identified from all these samples with PaliDIS v2.9.1 using default parameters (*--min_itr_length 25 --max_itr_length 50 --kmer_length 15 --min_is_len 500 --*

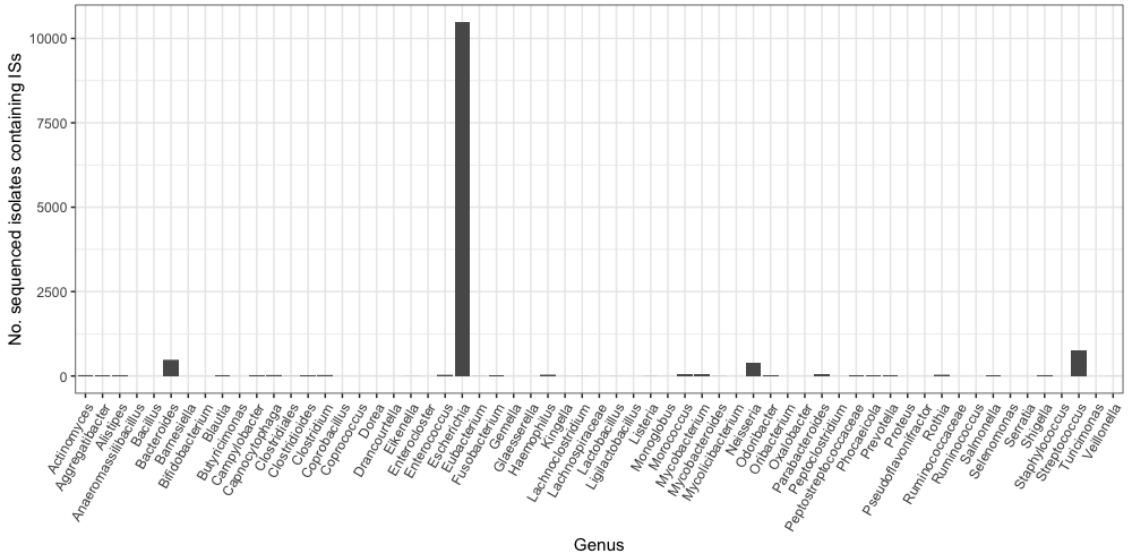
max_is_len 3000 --cd_hit_G 0 --cd_hit_c 0.9 –cd_hit_G 0 -cd_hit_aL 0.0 --cd_hit_aS 0.9 --cobs_threshold 1 --e_value 1e-50) and a COBS index (specified by --*cobs_index*) of 661,405 bacterial genomes created from European Nucleotide Archive (ENA) in 2018 (http://ftp.ebi.ac.uk/pub/databases/ENA2018-bacteria-661k/661k.cobs_compact)[11].

The ISs were then clustered using CD-HIT-EST v4.8.1 (with a sequence identity threshold *-c 0.99* and default parameters) to create the Insertion Sequence Catalogue (ISC). ISC contains a FASTA file of 11,681 unique ISs (https://github.com/blue-moon22/ISC) that were found in 10,810 contigs across 253 (out of 266) samples (Supplementary Data 2).

In PaliDIS, the ISs were queried against existing databases, ISfinder and the COBS index, to identify known ISs and predict their origin. Only 8 ISs were found in ISfinder (ISBvu3, ISBf3, ISBf8, ISLh1, ISVesp1, ISBvu4, IS1249 and ISBuba1). Another 164 ISs were predicted to belong to 17 families in ISfinder (IS1182, ISAs1, IS110, IS5, IS630, ISLre2, IS256, IS200/IS605, IS30, IS3, ISL3, IS1595, IS982, IS1380, IS4, IS66 and IS481). 722 ISs were located in 16,803 unique sequenced sources (NCBI BioSample IDs). 505 of these ISs were found in 11,516 microbial isolate genomes with known taxonomy consisting of 61 genera (Figure 2a) and 120 known species. 58 and 70 ISs originate from more than one genus (Figure 2b) and known species (not labelled *sp.*), respectively.

The IS shared across most genera is IS_cluster_192991_length_544 that was found within 18 genera and 27 species in 290 unique biological sources (NCBI BioSample IDs) and was not identified in ISfinder. 21 assemblies out of 290 BioSamples were publicly available and downloaded (Supplementary Data 3). Despite being found in all 21 samples' reads, IS_cluster_192991_length_544 was only found in 17 assemblies (using blastn v2.13.0 with e-value cut-off 1e-10) (Supplementary Data 4). These 17 assemblies were then annotated with prokka v1.14.5 to find functional genes and genomic features[19]. The clinically relevant tetracycline-resistant gene *tet(O)* was found upstream of IS_cluster_192991_length_544 in 15 assemblies across different classes: *Clostridium perfringens, Enterococcus gallinarum, Streptococcus agalactiae, Streptococcus azizii* and *Streptococcus suis* (Supplementary Figure 1a, b, d-g, i-q). This IS may therefore have a role in the HGT of *tet(O)*, probably by mediating the transposition of *tet(O)* to conjugative elements. Use of PaliDIS and the associated ISC can thus reveal new and important ISs that function in the spread of AMR across species and lineages.

a



b

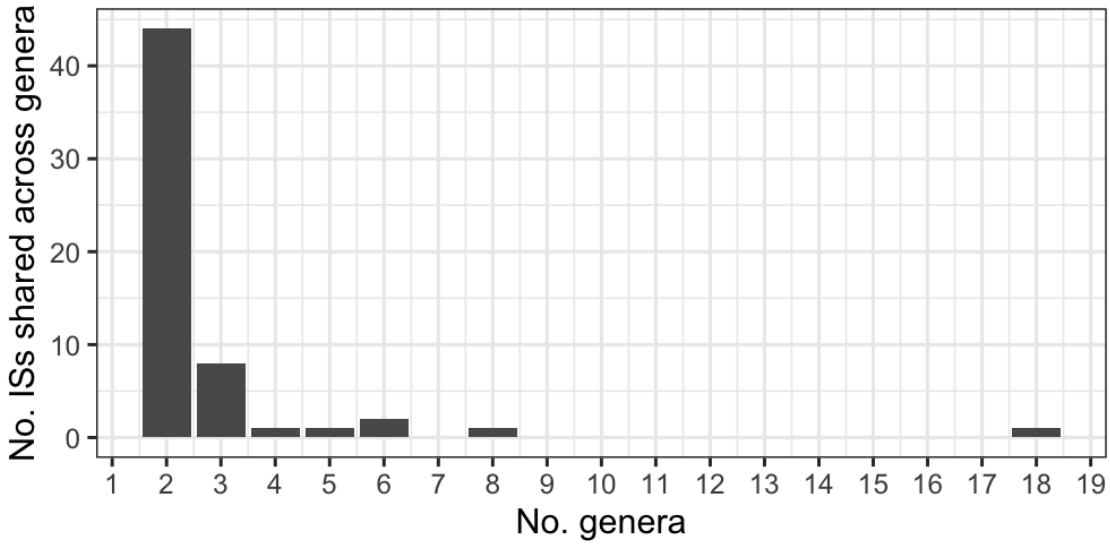


Figure 2. a) Number of sequenced isolates (NCBI BioSample IDs) containing an IS found in a particular genus. **b)** Number of ISs shared across more than one genus.

Discussion

Identification of transposable elements, including insertion sequences, in metagenomic datasets is critical in our ability to accurately define the profile of mobile genetic elements. In turn, accurate and complete characterisation of mobile genetic elements (i.e. the mobilome) of a community is central to understanding the spread and epidemiology of different genes in microbial communities, such as virulence genes and antimicrobial

resistance genes. Here, we describe a tool and subsequent catalogue that enables this to proceed. PaliDIS is a tool that discovers novel ISs from mixed microbial communities by applying a fast maximal exact matching algorithm to identify ITRs. As a result, we have released the first version of ISC, a catalogue containing 11,681 ISs. Already, this is a valuable resource for researchers to search for ISs in isolated genomes. However, since PaliDIS was only applied to metagenomes sequenced from the healthy human oral cavity and stool samples, it is recommended ISC is used as a reference for annotating isolates sourced from human oral and stool samples.

The main limitation of the current ISC is that it only contains common DDE types of ISs with ITRs, although these mobile genetic elements make up a large proportion of ISs. PaliDIS is currently only equipped with discovering ISs with ITRs. We are planning to include other databases into the catalogue, such as ISfinder, and we invite the research community to contribute and submit ISs to the catalogue. Another limitation is that the catalogue currently contains ISs with ITRs that are 25 or greater nucleotides in length as generated by PaliDIS, although ITRs can be as short as 10 nucleotides in length. It is possible to run PaliDIS with a lower minimum ITR length threshold and smaller *k*-mer length, but at these smaller sizers, it becomes more computationally intensive, especially with more complex mixed microbial genomes. However, we will run PaliDIS with a lower minimum ITR length threshold on less complex genomes to discover ISs with smaller ITRs.

It is also important to note that all ISs in the catalogue contain a region that is flanked by ITRs within a 500 to 3000 bp proximity. Given the recursive mechanism of insertion events (i.e. ISs inserting within ISs), it is possible for a region to also contain another IS. Therefore, it is also possible for regions that have been lengthened by other insertion events to extend outside this proximity range and be missed by PaliDIS. Increasing the maximum IS length will account for this, and may be done for future iterations of the ISC.

In light of creating this tool and catalogue, we cannot turn a blind eye to the fact that disruptive sequencing technologies are advancing rapidly by becoming more accurate and generating longer reads. Very soon, it will be easy to apply tools for *de novo* discovery of ISs in genomic assemblies with resolved repeat regions, rather than relying on reference databases. However, reference catalogues, like ISC, generated from older data could be applied to monitor microbes that may acquire ISs that carry antimicrobial resistance genes,

which will be invaluable information for appropriate actions for tackling AMR. For instance, determining whether an IS carrying an ARG has already been in circulation that has to be controlled or is emerging that can be prevented from spreading early. Furthermore, having a catalogue of ISs will enable simple searches of genomic datasets, as well as comparisons with ISs from different species using less computationally intensive methods that are available to all in the community. We will continue to enrich the ISC towards a comprehensive catalogue by applying PaliDIS with different parameters to more mixed microbial genomes from a diverse range of sources, and encouraging submission of ISs from the scientific community.

Acknowledgements

The project was supported by the Centre for Host-Microbiome Interactions, King's College London, funded by the Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/M009513/1 awarded to D.L.M. S.S. was supported by Engineering and Physical Sciences Research Council (EPSRC), EP/S001301/1, Biotechnology Biological Sciences Research Council (BBSRC) BB/S016899/1 and Science for Life Laboratory (SciLifeLab).

Conflicts of interest

The authors declare no conflicting interests.

References

1. Roberts, A. P. & Mullany, P. Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. *FEMS Microbiol. Rev.* **35**, 856–871 (2011).
2. Frost, L. S., Leplae, R., Summers, A. O. & Toussaint, A. Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* **3**, 722–732 (2005).
3. Mahillon, J. & Chandler, M. Insertion Sequences. *Microbiol Mol Biol Rev* **62**, 725–774 (1998).
4. Aziz, R. K., Breitbart, M. & Edwards, R. A. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res*. **38**, 4207–4217 (2010).

5. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, (2018).
6. Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. & Chandler, M. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **34**, D32-36 (2006).
7. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet. TIG* **16**, 276–277 (2000).
8. Kamoun, C., Payen, T., Hua-Van, A. & Filée, J. Improving prokaryotic transposable elements identification using a combination of de novo and profile HMM methods. *BMC Genomics* **14**, 700 (2013).
9. Treangen, T. J. & Salzberg, S. L. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* **13**, 36–46 (2012).
10. Bingmann, T., Bradley, P., Gauger, F. & Iqbal, Z. COBS: A Compact Bit-Sliced Signature Index. in *String Processing and Information Retrieval* (eds. Brisaboa, N. R. & Puglisi, S. J.) 285–303 (Springer International Publishing, 2019). doi:10.1007/978-3-030-32686-9_21.
11. Blackwell, G. A. *et al.* Exploring bacterial diversity via a curated and searchable snapshot of archived DNA sequences. *PLOS Biol.* **19**, e3001421 (2021).
12. Khiste, N. & Ilie, L. E-MEM: efficient computation of maximal exact matches for very large genomes. *Bioinformatics* **31**, 509–514 (2015).
13. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
14. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

16. Gálvez-Merchán, Á., Min, K. H. (Joseph), Pachter, L. & Booeshaghi, A. S. Metadata retrieval from sequence databases with ffq. 2022.05.18.492548 (2022) doi:10.1101/2022.05.18.492548.
17. Turnbaugh, P. J. *et al.* The Human Microbiome Project. *Nature* **449**, 804–810 (2007).
18. Carr, V. R. *et al.* Abundance and diversity of resistomes differ between healthy human oral cavities and gut. *Nat. Commun.* **11**, 693 (2020).
19. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinforma. Oxf. Engl.* **30**, 2068–2069 (2014).