INTRODUCING

Genpax

Technical Overview 2023/2024

A new era of connected pathogen genomics

About Genpax

- Genpax is a research and development-based company building novel solutions for clinical pathogen genomics. We are focused on **the needs of infection prevention and control (IPC)**, providing optimal information to recognize and respond to the transmission of strains in the healthcare system and beyond.
- Since 2021, we have established a large team of specialist bioinformaticians focused upon bacterial pathogen genomics, building upon more than a century of prior collective experience to develop species-specific toolkits with new analytical capabilities. These are being made available through our **IDEM** platform, addressing over 30, healthcare-associated, public health, and food-associated pathogens.
- Considering pathogens can change at a rate of 0 to 5 SNPs per genome per year, the actionable information needed to proactively detect (and exclude) outbreaks, infer transmission, and effectively direct IPC is beyond what other analysis pipelines can reliably deliver.
- Compromises, such as Sequence Typing (e.g. cgMLST) achieve scalability and error tolerance at the expense of sensitivity and specificity. In contrast, SNP-solutions (e.g. wgSNP) cannot be scaled and have reference-dependent accuracy (and good references do not exist for many strains and species).
- Genpax exists to **eliminate these constraints**; to deliver a new generation of pathogen analysis capabilities which address the IPC challenges of emergent pathogens and AMR.
- This brochure highlights a selection of **key differentiating capabilities** in our quest to make the best possible genomic pathogen analysis accessible to everyone.

Genpax

Our automated cloud-based solution includes:

- **A SNP-throughout analysis** that makes maximal use of the safely interpretable genome sequencing information.
- An analysis solution that **does not require typing or other steps** to select a reference genome.
- **Performance equivalent to wgSNP under its most optimal conditions**, delivering equally optimal and comparable results for all species and strains.
- Entirely c**onsistent findings from clinical replicates**, identical strains from common sources, and samples in clinical ring-trials.
- A **near-zero error rate** meaning results from the multiple labs can be reliably combined and compared.
- **Unprecedented accuracy** combined with **addressing more of the genome sequence** information than either Sequence Typing or previous whole genome SNP comparisons.
- Unmatched capabilities to identify stains and **infer their likely membership of outbreaks (or not) and order of transmission**, even with only two isolates within a transmission cluster.
- Effective analysis that can be used as **part of a clinical solution** to help optimize infection prevention and control workflows for improved patient care and safety at the same time as reducing the costs of healthcare.
- **Open scalability**, so that each newly analyzed strain can be compared with all strains previously processed, within and between sites that choose to openly display their results.
- A **user-friendly platform** with interactive and continuously updated communication of findings.
- A **rapid turnaround time**, whether analyzing one or hundreds of isolates from a sequencing run, while comparing them with hundreds or tens of thousands (or more) previously analyzed strains.

Genpax

The Genpax Posters

As a commercial entity, Genpax cannot share its code and solutions. However, we are a team, comprised largely of former academics with hundreds of publications in the field between us, who want to share our system and its capabilities as openly and clearly as possible. To do this, we have performed a set of validation and demonstration analyses using information from the highest quality publications and studies that we could identify, selecting those with the best evidence for 'ground truth' against which to be measured.

- In tests of **reproducibility** and near-zero error, genuine biological replicates from clinical ring trials (ECCMID, *Staph. aureus*) and large well-documented outbreaks and re-isolation studies have been addressed (*E. coli* and *K. pneumoniae*).
- In tests of **accuracy** (*E. coli*, *Campylobacter jejuni*, and *Ps. aeruginosa*), exceptional situations in which published or specifically generated almost identical high-quality reference genomes were used in the original studies were selected and thus represent the most stringent findings to measure performance against that we could identify.
- In tests of **reference-free** performance and transmission-chain re-structuring, species were selected that represent extremes of highly recombining panmictic (*Campylobacter jejuni*) and deeply rooted, highly-clonally diverse (*Ps. aeruginosa*) population structures.
- In the test of **scalable** comparisons, using *Listeria monocytogenes,* we processed data generated in large studies from different European laboratories.
- The test of MRSA **gene finding** (ASM, *Staph. aureus*) used sequencing and MRSA/MSSA data from two published studies from an EU reference laboratory.
- Likewise, our **economic modeling** adopts conservative assumptions, taking a cautious approach towards outbreak sizes and containment speed, in contrast to the assumptions of the published models it is built upon. Additionally, our analysis includes up-to-date costs of sequencing and analysis ensuring accurate financial impacts.

Each poster primarily addresses one or two aspects of our platform's performance for IPC applications: accuracy, low noise, high resolution, comparability, species applicability, and scalability. In combination they represent a real enhancement in what sequencing analysis can offer infection prevention and control.

Genpax

The Genpax Posters

Novel genome comparison tool reveals both false-positive and false-negative MRSA and MSSA strain identification and a failure to detect transmission-linked strains using phenotypic, PCR, and previous genomic strategies – **ASM 2023**

Openly comparable and scalable SNP-resolution analysis for *Listeria monocytogenes* using a novel genome comparison tool – **ASM 2023**

Economic and health impact modelling of a whole genome sequencing-led intervention strategy for bacterial healthcare-associated infections for England and for the USA – **ASM 2023**

Near zero error using large-scale hospital outbreak whole genome sequence data for *Klebsiella pneumoniae* – **ASM 2023**

A novel genome comparison tool producing near-zero error for same-patient isolates of *E. coli* ST131 – **ASM 2023**

Reference-free WGS SNP-resolution analysis of *Campylobacter jejuni* – **ASM 2023**

Reference-free whole genome SNP analysis of *Pseudomonas aeruginos*a, with the restructuring of outbreaks analyzed with established methods – **ASM 2023**

Calling Zero: A new foundation for diagnostic bacterial genomics – **ECCMID 2023**

Genpax

Introduction to the Following Health Economics Paper

A clinical genomics solution must provide an increase in patient and public safety with improved patient care and improved management of healthcare resources, while also being economically viable. Several previous publications indicate that this is the case for proactive clinical genomics for IPC. (These are cited in the following paper.)

However, these papers do not always include all attributable costs of sequencing, analysis, staffing and other infrastructure costs, and others are dated or don't work with the most current epidemiological information or a more limited set of species. This new analysis draws upon the best of this previous work, updates it, addresses a core set of healthcare-associated hospital-acquired AMR priority pathogens, using inclusive current real world costs, and a best-available set of epidemiological information. The publication also makes its model available in an Excel format to enable local hospitals or others to modify it to generate locally informed versions and ongoing updates. Key findings of this modelling include:

- Addressing NHS England as an example:
  - it would be possible to **save over 70,000 bed days per year**, the equivalent of building and fully equipping and staffing a new 200-bed hospital with full occupancy
  - it is possible to **prevent more than 1,200 avoidable hospital care associated deaths**; representing 10-20% of estimated avoidable hospital deaths per year
  - it is possible to **save at least £480 million per year** in avoidable costs
- There is **no economic obstacle to adoption**, because the savings to hospitals and healthcare delivery systems are considerably greater than the costs of adopting and delivering proactive bacterial genomics surveillance for IPC.
- **Improved patient safety and actions to contain and prevent the spread of AMR** within the hospital can be achieved at negative costs.
- The **hospital-level costs savings** are dominated by improved use of healthcare resources, such that large savings remain with wide variations in the costs of sequencing and analysis.
- Larger savings and proportionate returns on investment are available in the US than the UK

The only remaining obstacles to adoption are sequencing, which is now available in-house to any modern laboratory capable of typical microbiology and pathology services or through external services, and the expertise and resources that are required for analysis and interpretation that are now openly available from Genpax.

Finally, it should be noted that **these models are intentionally conservative**. They do not include savings from other activities such as combined environmental, healthcare worker, and pre-admission screening; the additional benefits of addressing non-AMR/antibiotic sensitive strains with similar transmission mechanisms and clinical consequences (e.g. MSSA which has a 20 to 30% mortality); nor additional species. They also do not include costs associated with exceptional responses such as ward closure, rebuild and refits, equipment replacement, insurance company non-payment or claw-backs, or legal liabilities for hospital transmitted infections. Nor the savings from avoidable responses to ‘non-outbreaks’ that suspected on epidemiological grounds are caused by strains are unrelated and not connected, or being able to demonstrate that infections were not caused by hospital-associated strains.

MICROBIAL GENOMICS

RESEARCH ARTICLE

Fox *et al.*, *Microbial Genomics* 2023;9:001087

DOI 10.1099/mgen.0.001087

MICROBIOLOGY SOCIETY

OPEN DATA OPEN ACCESS

Economic and health impact modelling of a whole genome sequencing-led intervention strategy for bacterial healthcare-associated infections for England and for the USA

John M. Fox, Nigel J. Saunders and Susie H. Jerwood*

Abstract

Bacterial healthcare-associated infections (HAIs) are a substantial source of global morbidity and mortality. The estimated cost associated with HAIs ranges from $35 to $45 billion in the USA alone. The costs and accessibility of whole genome sequencing (WGS) of bacteria and the lack of sufficiently accurate, high-resolution, scalable and accessible analysis for strain identification are being addressed. Thus, it is timely to determine the economic viability and impact of routine diagnostic bacterial genomics. The aim of this study was to model the economic impact of a WGS surveillance system that proactively detects and directs interventions for nosocomial infections and outbreaks compared to the current standard of care, without WGS. Using a synthesis of published models, inputs from national statistics, and peer-reviewed articles, the economic impacts of conducting a WGS-led surveillance system addressing the 11 most common nosocomial pathogen groups in England and the USA were modelled. This was followed by a series of sensitivity analyses. England was used to establish the baseline model because of the greater availability of underpinning data, and this was then modified using USA-specific parameters where available. The model for the NHS in England shows bacterial HAIs currently cost the NHS around £3 billion. WGS-based surveillance delivery is predicted to cost £61.1 million associated with the prevention of 74 408 HAIs and 1257 deaths. The net cost saving was £478.3 million, of which £65.8 million were from directly incurred savings (antibiotics, consumables, etc.) and £412.5 million from opportunity cost savings due to re-allocation of hospital beds and healthcare professionals. The USA model indicates that the bacterial HAI care baseline costs are around $18.3 billion. WGS surveillance costs $169.2 million, and resulted in a net saving of ca.$3.2 billion, while preventing 169 260 HAIs and 4862 deaths. From a 'return on investment' perspective, the model predicts a return to the hospitals of £7.83 per £1 invested in diagnostic WGS in the UK, and US$18.74 per $1 in the USA. Sensitivity analyses show that substantial savings are retained when inputs to the model are varied within a wide range of upper and lower limits. Modelling a proactive WGS system addressing HAI pathogens shows significant improvement in morbidity and mortality while simultaneously achieving substantial savings to healthcare facilities that more than offset the cost of implementing diagnostic genomics surveillance.

Impact Statement

This article estimates the impact of effective whole genome sequencing-based surveillance for tracking and intervening in bacterial nosocomial outbreaks of the 11 most common healtcare-associated infection (HAI) species in both England and the USA. The projected outcome would be to reduce the bacterial morbidity and mortality of HAI in hospitals while simultaneously reducing the cost of patient care and increasing the wider cost savings of England and the USA by £478.3 million and $3.2 billion respectively, with more efficient use of hospital resources.

Genpax

www.genpax.co
+44 20 3603 6869
research@genpax.co