CASE STUDY:

New STAG Strain-Typing Method Could Unlock the Mystery of Pathogen Evolution

Palmona Pathogenomics developed a new strain-typing algorithm called STAG (sequence typing by accessory genome)—a breakthrough that takes strain typing to the next level. Unlike other methods that solely rely on typing the core genome, STAG categorizes the greater pangenome to gain further insights into accessory genes.

The result: The ability to predict pathogen behavior, demonstrated in one of the world's most destructive pathogens.

THE APPROACH

The STAG algorithm was applied to classify, model, and predict the behavior of different strains of *C. difficile*.

MODEL

451

STRAIN-SPECIFIC GENOME-SCALE MODELS OF METABOLISM (GEMs)

CLASSIFY

176

GENETICALLY DISTINCT GROUPS OF STRAINS

PREDICT

76%

OF CASES HAD CORRECTLY PREDICTED GROWTH, LINKING PHENOTYPES AND STRAIN-SPECIFIC GENETIC DIFFERENCES

THE DISCOVERY

STAG's innovative algorithm opened the door to a deeper understanding of strain characteristics and traits through prediction.

THE OPPORTUNITY

A way to identify genetic features that could contribute to the emergence of new problematic strains.

PALMONA PATHOGENOMICS

This case study was adapted from a paper published in the *PNAS* journal—**scan here to view.**

"Through STAG, there is a route where these algorithms provide an expedited path to developing a more complete view of identifying AMR [antimicrobial resistance] and associated strategies during infection to improve diagnostics and treatment courses."

—Charles J. Norsigian, PhD, study author

AN OPPORTUNISTIC PATHOGEN, LIMITED TYPING METHODS

THE CHALLENGE: Antibiotic use eliminates normal microbiota—or good bacteria—allowing *C. difficile* bacterium to establish itself. This can cause *C. difficile* infection (CDI), a potentially fatal condition that recurs in ~35% of infected patients. *C. difficile* is the most common cause of healthcare-acquired infections (HAIs), due to its complex and ever-evolving genetic makeup.

The Centers for Disease Control and Prevention (CDC) recognizes *C. difficile* as an urgent threat.

Estimated annual US statistics due to *C. difficile*

220,000 cases in hospitalized patients | 13,000 deaths

While polymerase chain reaction (PCR) ribotyping and multilocus sequence typing (MLST) have proven successful in linking *C. difficile* to CDI, they are unable to distinguish more closely related strains.

ENTER STAG: DEVELOPING THE METHOD

The STAG study presents a systems biology analysis that expands the capabilities of strain-typing methods by connecting the relationship between genotypes (genetic traits) and phenotypes (behavior), which unlocks predictive power into understanding how the bacterium spreads and why.

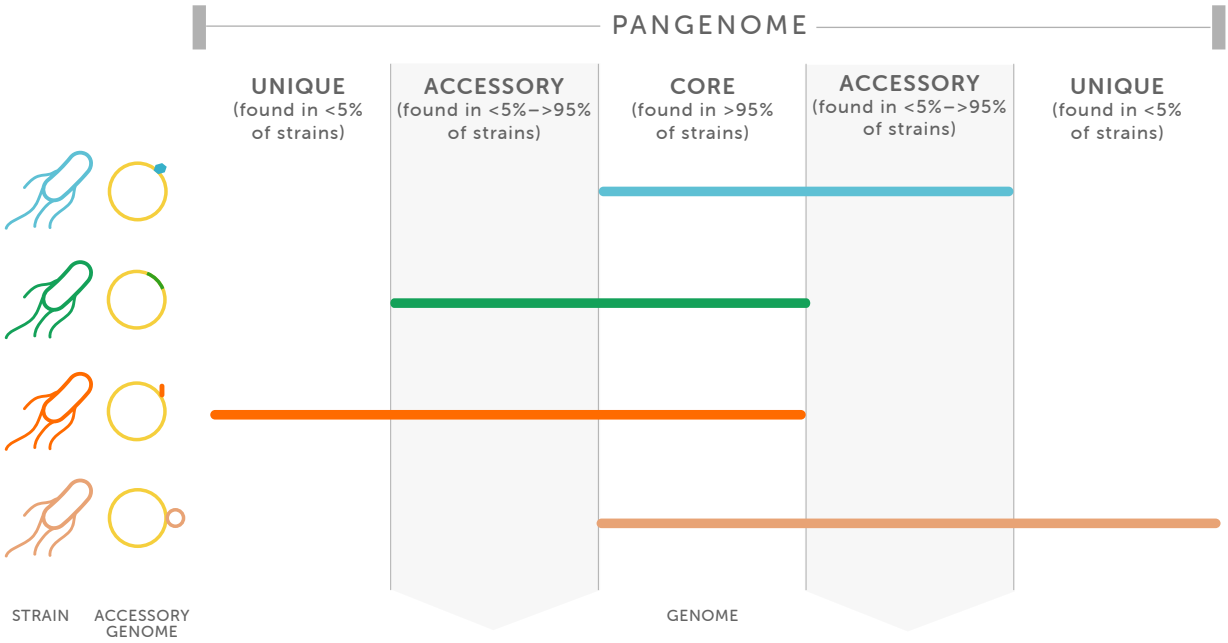
The following charts explain the components in developing the STAG algorithm, which uses the *C. difficile* pangenome to identify the core, unique, and **accessory gene** clusters.

COMPONENTS OF THE STAG STUDY

Whole-genome sequencing (WGS)	Genome-scale models of metabolism (GEMs)	Construction of **pangenome**
Genomes of the 451 clinical isolates were compared to a reference strain	• System modeling of metabolic capabilities • Links genotype to phenotype	*C. difficile* pangenome (based on 451 strains)

THE PANGENOME UP CLOSE

The strains below share a core genome (found in >95% of strains), while accessory genes and unique genes display increasingly greater variance.



Evaluating accessory genes provides a genetic bank rich in discriminatory power.

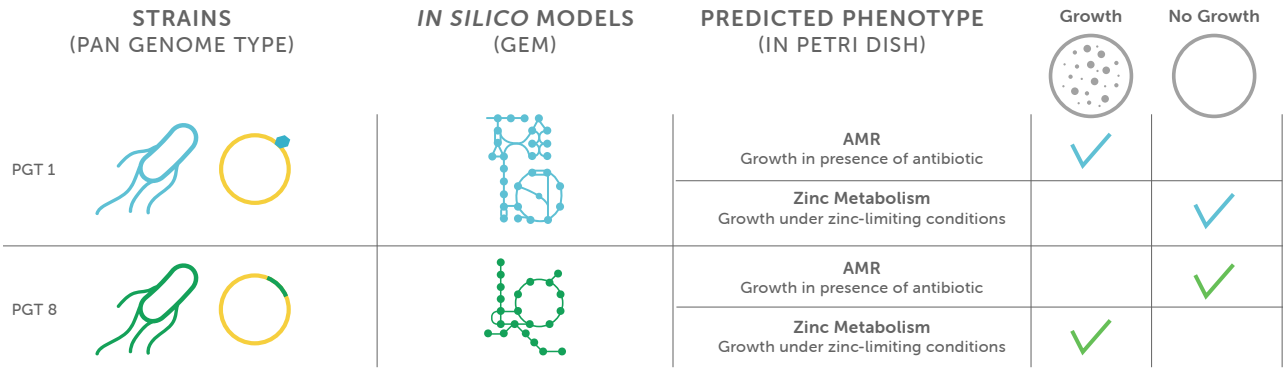
Accessory genes provide genetic flexibility that allows adaptation in various environments.

The **pangenome** is the entire gene set of all strains of a species. It includes genes present in all strains (core genome) and genes present only in some strains of a species (accessory genome).

Whole-genome sequencing is the analysis of the entire genomic DNA sequence of a cell or bacterium, which allows for the most comprehensive characterization of the genome.

THE POWER OF PREDICTION

The STAG algorithm was used to build models of strain-specific metabolic capabilities to predict growth/no growth of those strains across different environmental conditions. This modeling can be applied to various factors to predict pathogen behavior, such as antibiotic resistance and metal homeostasis.



"An advantage of STAG modeling is that we were able to see differences that were previously not appreciated. For example, zinc metabolism is not something I ever considered as a virulence factor/fitness advantage present in these strains. A standard workflow requires experimentalists to know the exact question they want answered, in order to adjust parameters. By using STAG, we identified sub-groupings that led to experimentation without actually asking about zinc metabolism."

—Dr Heather Danhoff, study author

MORE TO DISCOVER

In this case study, we demonstrate how the STAG approach was used to type strains of *C. difficile*, and to subsequently make important predictions on behavior (eg, AMR) that were experimentally validated. But the promise of STAG is not limited to *C. difficile,* as it can be applied to other types of bacterial species, bringing an understanding of genetic diversity to the forefront and allowing deeper insight to a broad range of clinically relevant processes.

Palmona Pathogenomics has used similar techniques to identify:

- A gene cluster in *Escherichia coli* that predicts the risk of urinary tract infections (UTIs) progressing to bloodstream infections
- A gene cluster in *Leptospira* that predicts pathogenic behavior
- Polymorphisms in *Mycobacterium tuberculosis* and *Staphylococcus aureus* that predict resistance to antibiotics

These represent just some examples of Palmona Pathogenomics' ability to categorize and predict bacterial behavior, unlocking invaluable insights into the mysteries of pathogen evolution.

READY TO TAKE ADVANTAGE OF THE POWER OF PREDICTION?

Scan code to visit palmona.com

PALMONA
PATHOGENOMICS
an Avellino Company

(650) 396-3741 **www.palmona.com** PAL 12-08-0009