BACTERIOPHAGE GENOME ANNOTATION: COMPARING AUTOMATIC AND MANUAL APPROACHES

A. Culot, G. Abriat, K. Furlong

uOttawa

GENOMICS FOR PHAGE THERAPY

Importance and challenges

PHAGE THERAPY

Bacteria + Phage = Dead bacteria

GENOMICS FOR PHAGE THERAPY

Importance and challenges

PHAGE THERAPY

Bacteria + Phage = Dead bacteria

PHAGE MESSING UP ECOSYSTEMS & PUBLIC HEALTH

Bacteria + Phage = New bacterial strain

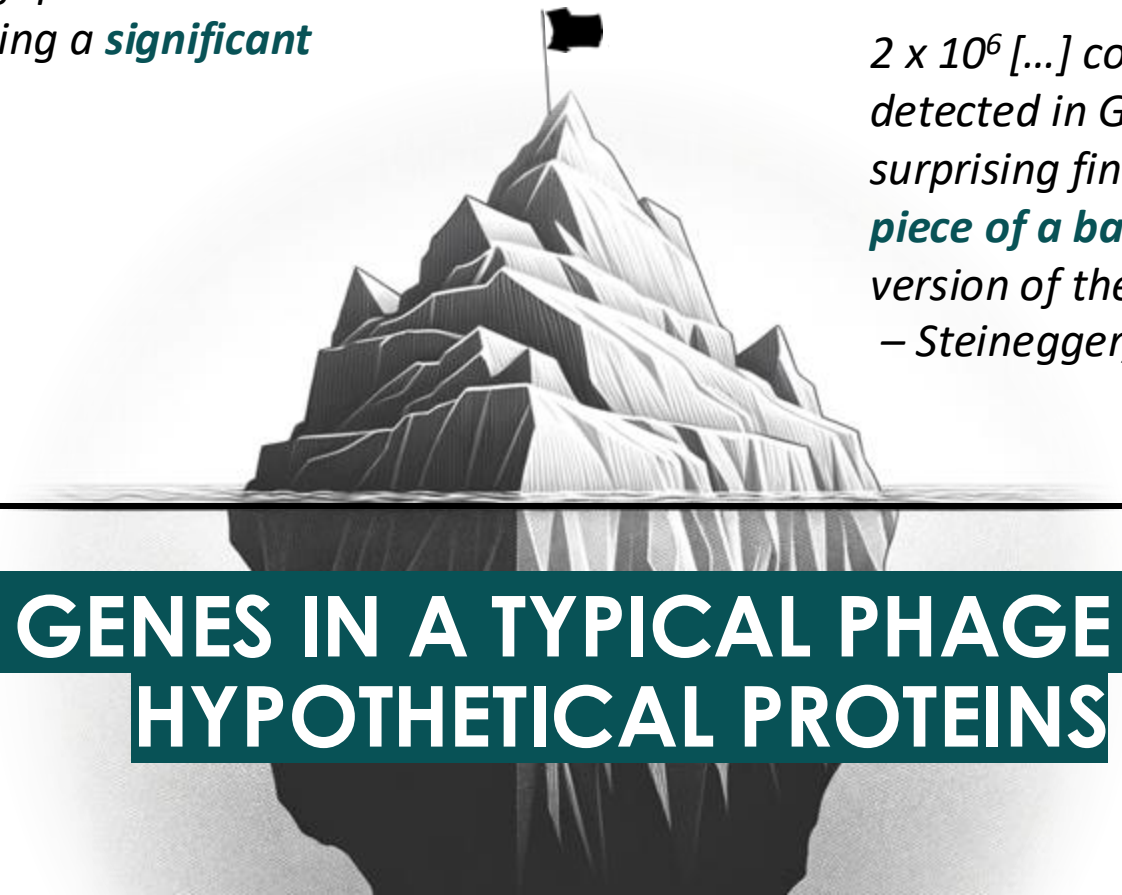
→ We need to understand phage gene functions to avoid this

BIOINFORMATICS IN MICROBIOLOGY

The taboo of bad bioinformatics

*"**High-quality gene finding** and **functional annotation** are **vital** for successful discovery of new phage-based therapeutics."*
– Kutter et al, 2015

*"**Lamentably**, […] as phage sequencing becomes more high-throughput and automated, we are observing a **significant increase in problems**"*
– Turner et al 2021



*2 x 10^6 […] contaminated sequences were detected in GenBank. This single most surprising finding was the **presence of a piece of a bacterium**, in […] the current version of the **human reference genome.**"*
– Steinegger, 2021

70-80% OF GENES IN A TYPICAL PHAGE GENOME ARE HYPOTHETICAL PROTEINS

WHY ARE WE FACING THIS CHALLENGES ?

The taboo of bad bioinformatics

POOR PEER REVIEW

LACK OF TRAINED BIOINFORMATICIANS

UNFRIENDLY USER INTEFACES

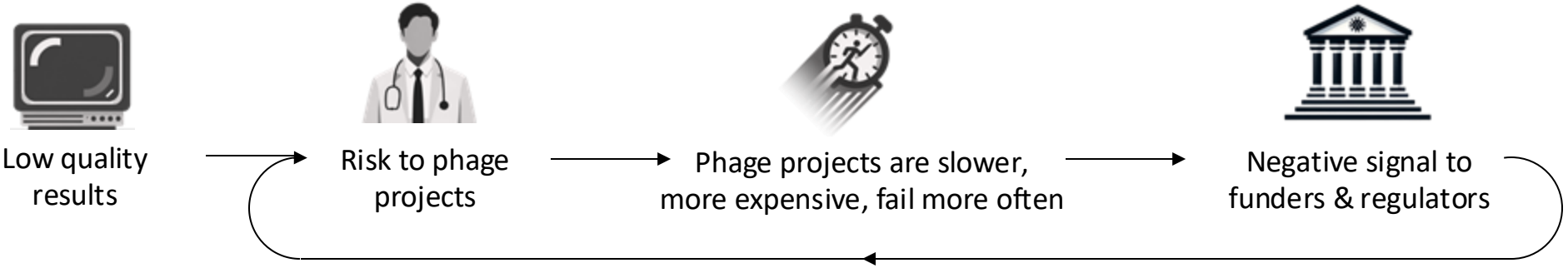
PRICE AS THE HISTORICAL DRIVER OF SEQUENCING MARKET

BIOINFORMATICS IN MICROBIOLOGY

Why working properly matters

LOW QUALITY BIOINFORMATICS DRIVE SCIENCE DOWN

LOW QUALITY BIOINFORMATICS DRIVE THE FIELD DOWN



BIOINFORMATICS PRACTICES IMPACTS SCIENCE, INDUSTRY AND PATIENTS

OVERCOMING THE CHALLENGES OF PHAGE GENOMICS

What is already being done

PHAROKKA

GALAXY

Galaxy PROJECT

COLABFOLD

SEA-PHAGES

SEA PHAGES

SEA-PHAGES
The project

UNDERGRADUATE EDUCATION

- Promote student interest and retention through discovering, characterizing and naming their own bacteriophages

RESSOURCES

- The SEA-PHAGES program involves (but is not limited to) DNA sequencing and the subsequent data analysis
- Participants in the program are provided with protocols, and custom tools to assist them in the analysis of the data

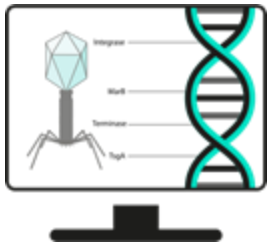
SCIENTIFIC PUBLICATION

- Synergistic opportunity : publishing high quality manually curated, phage genomes
- More than 5 000 high-quality genomes produced
- But, it takes a long time and several students to work on a single phage genome

HIGH PERFORMANCE BIOINFORMATICS TO PROMOTE A SAFE PHAGE THERAPY FOR PATIENTS AND THE ENVIRONMENT

The RIME Bioinformatics project

Software development



- rTOOLS2

Methods, Repeatability, Quality

- ISO 9001
- GxP

Expertise

- Phage therapy
- Genomics
- Regulation

Engagement

- Safe phage therapy
- Help academic & private projects
- Phage ecosystem construction

BENCHMARKING RimeTOOLS2 & SEA-PHAGES

AIM

« Can automatic genome annotation compare with manual approaches ? »

- Kieran Furlong, University of Ottawa at the 2023 Evergreen Conference

BENCHMARKING RimeTOOLS2 & SEA-PHAGES

M&M

« Can automatic genome annotation compare with manual approaches ? »

- Kieran Furlong, University of Ottawa at the 2023 Evergreen Conference

27 Phage genomes from uOttawa's SEA-PHAGES program

Novel annotation with RimeTOOLS 2 – automatic version

Compare structural and functional annotation results

BENCHMARKING RimeTOOLS2 & SEA-PHAGES

Results & discussion

SEA PHAGES

With rTOOLS2 as a reference

RIME BIO INFORMATICS

With SEA PHAGES as a reference

STRUCTURAL ANNOTATION

+1.5
genes detected / genome

-1.5
genes detected / genome

FUNCTIONAL ANNOTATION

-4.2
better annotated genes / genome

+4.2
better annotated genes / genome

Average difference per genome

BENCHMARKING RimeTOOLS2 & SEA-PHAGES
Conclusion

« Can automatic genome annotation compare with manual approaches ? »

- Kieran Furlong, University of Ottawa, 2023 Evergreen Conference

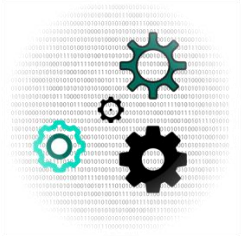
→ Yes, but manual curation is still necessary

BENCHMARKING RimeTOOLS2 & SEA-PHAGES
Conclusion

« Can automatic genome annotation compare with manual approaches ? »

- Kieran Furlong, University of Ottawa, 2023 Evergreen Conference

→ Yes, but manual curation is still necessary



We are already working on it !

→ **RimeTOOLS3 :**

- Phage genomes :
  - Improved structural annotation
  - Improved functional annotation
- Bacterial genomes :
  - Improved prophage detection
  - Tailocin naïve discovery

- **Full Release : Q1 2025**

Regulation compliant

CERTIFIED ISO 9001:2015 COMPANY

GxP

RIME BIO INFORMATICS

THANK YOU !

uOttawa

Antoine.Culot

@rime-bioinformatics.com

Guillaume.Abriat

@rime-bioinformatics.com

Kieran Furlong