A summary of the last decade of my life

Bram van Dijk (he/him)

- BSc thesis with Paulien (2011-2012)
- PhD thesis with Paulien (2015-2020)
- Postdoc with...

- From September 1st I’ll start my own research group here at Utrecht University (Theoretical Biology group)
Microbial evolution is different

“The walls that divide bacteria from one another are far from solid. Taken to extremes, the preponderance of HGT could even imply that microbiomes are better conceptualized as collections of locally adaptive genes, rather than communities of locally adapted species”

— James P.J. Hall, 2021
Is HGT really that extreme?

Well. Yes. But what type of evidence do we have?

Quantify mutations that entail whole genes:

- Duplication of a gene (**Expansion**)
- Deletion of a duplicated gene (**Reduction**)
- Deletion of the only copy of a gene (**Loss**)
- **Gain** (mostly HGT)

"Indeed, the estimated rates of gene family gain and loss in some groups of bacteria are such that multiple genes appear to come and go over the time required for a single nucleotide substitution to occur in an evolving gene."

**Note:** (most) Eukaryotes are not like this!
The walls that divide bacteria...

Implications for the bacterial species concept...
... but also requires rethinking of classical population genetics (i.e. alleles slowly changing in frequency in populations)
So... let’s take HGT to extremes, and see what happens!

Q1: How much diversity can we maintain?
Q2: How does gene mobility evolve?
First... what does evolution do?

Differential gene mobility!
Resistance is widespread, killing is rare

Rings depict closely related genes mapped onto the species tree (inner circle = resistance-genes, outer circle = toxin genes)
Why do toxin genes become mobile?

Only toxin genes that “outmobilise” their resistance factor are able to survive. If toxins are lost... resistance is futile \( (╯°□°)╯︵ ┻━┻ \)

So when that happens, diversity ought to collapse.
Low diversity without (evolution of) gene mobility

![Graph showing strain diversity over simulation time with different mobility parameters.](image)
Why don’t all genes become parasites?

- If we mix the gene pool, all genes DO become parasites
- Indicates a local feedback process
- Resistance is useful: it isn’t often lost!
- Higher gene mobility would just give cells more and more copies with no extra benefits
Core and accessory genomes

- "Core" is what they all share
- "Accessory" is what only one or few strains have
- Other terminology is core-, shell- and cloud genomes, where accessory genes are split into those that are very rare (cloud) and those that are abundant, but not core (shell)
Core and accessory genomes

1. Core-like resistomes
2. Patchy distribution of toxin genes
3. Accessory
4. Core/shell
5. Toxins genes are best conceptualised as "sometimes-useful parasites"
“Sparing” close kin without kin recognition

- In nature, bacteria often “spare” close kin from killing

- The model does not have any form of kin recognition: they don’t “know” who they are killing

- Instead: these “socially cohesive units of interaction” are the consequence of local reproduction and local gene transfer

- Low, instead of kin recognition, it is better conceptualised according to aforementioned “locally adaptive gene pools”
Conclusions project #1

On the cell-level:
• Despite HGT driving a lot of “genetic mixing”, that doesn’t mean everyone becomes the same.
• Concepts like “individuality” and “species” still persist for bacteria, but the gene- and group-level are equally (or more) important.

On the gene-level:
• Mobile genes Darwinian entities themselves, and evolve towards “parasitism” if they get the chance!
• However: a feedback with the local environment prevents this from getting out of control!
• Toxin genes can get away with being parasites in the short term, because they are occasionally beneficial.
New question based on project # 1

- Toxin genes persist by transferring and by being “occasionally useful” (but on average, only very little!)

- Resistance genes persist without transfer because they are always very beneficial, because toxins are bouncing around everywhere

- Let’s take a step back and forget about toxin/resistance for a second. Let’s think about one gene, in a single population.

- What about a gene with a given fitness effect, e.g. one that is always “slightly” beneficial? Does spatial structure matter here too?
A simple ODE model

A) ODE model (carriers and non-carriers)

- **We assume HGT has a cost** (uptake of DNA, transfer machinery, etc.)
- **How does costly HGT impact growth rates?**

\[
\begin{align*}
\frac{dC}{dt} &= (1 - ch + b)C - \frac{lC}{HGT} + \frac{hCN}{HGT} - \phi C \\
\frac{dN}{dt} &= (1 - ch)N + \frac{lC}{HGT} - \frac{hCN}{HGT} - \phi N \\
\phi &= \frac{(1 - ch + b)C}{HC} + \frac{(1 - ch)N}{NC} \\
C + N &= 1 \text{ (constant population size, ensured by chemostat assumption.)}
\end{align*}
\]

Q: How does HGT impact the population growth rate?
Maintenance of beneficial genes depends on parameters

A)
But does it benefit the microbial population?

\[
\frac{dC}{dt} = (1 - ch + b)C - \underbrace{IC}_{\text{reproduction of } C} - \underbrace{lC}_{\text{gene loss}} + \underbrace{hCN}_{\text{HGT}} - \underbrace{\phi C}_{\text{chemostat}}
\]

\[
\frac{dN}{dt} = (1 - ch)N + \underbrace{lC}_{\text{reproduction of } N} - \underbrace{hCN}_{\text{HGT}} - \underbrace{\phi N}_{\text{chemostat}}
\]

\[
\phi = \underbrace{(1 - ch + b)C}_{\text{total growth of } C} + \underbrace{(1 - ch)N}_{\text{total growth of } N}
\]

\[
C + N = 1 \quad \text{(constant population size, ensured by chemostat assumption.)}
\]

\[
\phi^*(h) = \begin{cases} 
1 - ch & \text{if } h \leq (l - b) \text{ (gene cannot persist)} \\
1 - ch + b - \frac{bl}{b + h} & \text{if } h > (l - b) \text{ (gene persists)}
\end{cases}
\]
Distinct gene "classes" exist

**Indispensable genes**
These genes are readily maintained. Transferring these genes with costly HGT always diminishes the growth rate of the population.

- **HGT is...**
  - REQUIRED: No
  - BENEFICIAL: No

- **$b > 1/c$**

**Enrichable genes**
These genes are maintained without HGT, but HGT can improve the population growth rate.

- **REQUIRED**
  - No
- **BENEFICIAL**
  - Yes

- **$b < 1/c$**

**Rescuable genes**
These genes are lost from the population without HGT. If the rate of HGT is high enough, the genes persist and the population growth rate is improved.

- **REQUIRED**
  - Yes
- **BENEFICIAL**
  - Yes

- **$b < 1$**

**Unrescuable genes**
These genes are lost from the population without HGT. Despite being maintained at intermediate rates of HGT, the population grows fastest without HGT.

- **REQUIRED**
  - Yes
- **BENEFICIAL**
  - No

- **$b < 4d/(1+c)^2$**

**Selfish genetic elements**
Selfish genetic elements (SGEs) can only persist at very high rates of HGT. When they do, the population growth rate is diminished.

- **REQUIRED**
  - Yes
- **BENEFICIAL**
  - No

- **$b < 0$**

Where would you put the **toxin genes** of the previous part?

What about the **resistance genes**?
Can cells evolve to invest energy into HGT to "rescue" genes?

B) ODE model (carriers and non-carriers with / without HGT)

\[
\begin{align*}
\text{Carrier} & \quad \varphi = 1 + b - c_h \\
\text{Non-carrier} & \quad \varphi = 1 - c_h
\end{align*}
\]

Proportional to total carriers \((C^+ + C^-)\)

\[
\begin{align*}
\text{Carrier} & \quad \varphi = 1 + b \\
\text{Non-carrier} & \quad \varphi = 1
\end{align*}
\]

Dilution by total growth \(\phi = (\varphi_{C^+} + \varphi_{N^+} + \varphi_{C^-} + \varphi_{N^-})\)
Can cells evolve to invest energy into HGT to "rescue" genes?

- Allee effect: HGT for rescuable genes is evolutionarily stable, but not "evolvable"
What about space though?

C) Individual-based, eco-evolutionary model

Q: What is the impact of spatial structure?

- Can cells invest into costly HGT of rescuable genes now?
A plot of what you just saw in the movie

C) Long-term coexistence of cells, beneficial genes, and strong SGEs

START INFLUX SGEs

STOP INFLUX SGEs

START INFLUX SGEs

STOP INFLUX SGEs

Bacteria infected by SGEs, $h$

Bacteria not infected by SGEs, $h$

Gene freq

HGT rate
From the gene-level
HGT can “rescue” genes from extinction. To put it differently, there may be genes (in nature) whose entire existence depends on repeatedly moving to new hosts through HGT

From the cell-level
This behaviour is evolutionarily stable, but only “evolvable” in a spatial system

Spatial system even opens the door to selfish genetic elements (SGEs), but the cells “take the good with the bad”
Are there genes that require HGT to persist?

What to use as a control?
Rescuable genes in Virtual Microbes?

A) Genome size over time [AUTC]

B) Mutational regime / replicate

C) DUP-populations

D) HGT-populations

E) i4-gene (by-product importer) is an "accessory gene" only preserved in the HGT-populations with smaller genomes.
What about the vehicles of HGT?

These are the “text-book” examples of HGT, but there are more mechanisms and vehicles:

Membrane vesicles, gene transfer agents, transposons, integrative and conjugative elements (ICEs), integrons, BORGs, starships and voyagers, even mobile chromosomes...

The list is endless...

“Mobile elements are entities that evolved to persist and replicate through adaptations that move DNA.”

— James P.J. Hall, 2021; “The secret lives of Mobile Genetic Elements”
Mobile elements are diverse and evolve on a parasitism–mutualism continuum.
Friends or foes?

- Transposons are the simplest “nested replicator”, they replicate inside chromosomes
- They can also jump from cell to cell, after uptake from the environment
- Occasionally, they carry useful genes in nature, such as antimicrobial resistance genes
- So what drives parasitism vs mutualism for these very simple entities?
Can TEs get “linked" to AMR genes?

Toy model of co-evolving genomes and transposons
TEs recognise flanking DNA

**Replication-rate of transposons** determined by $\phi$ of flanking DNA:

- Transposase flanked by nearby repeats?
  - No
  - Yes: every time step (rate $j$), replicate within host with chance $p = 0.1$. If
    - After uptake (rate $u$), integrate with chance $p = 0.1$. If

**Replication-rate of cells** (local competition) determined by fitness ($f$):

- All housekeeping genes present
  - Yes, $f = 1 - C_n \cdot \sum\text{ genes present}$
  - No, $f = 0.0$
TEs emerge after some time (no AMR selection)
System persists because of spatial structure

Host replicator (microbial cell)

Nested replicator (transposable element)

a. negative selection > replication rate: transposable element lost

b. negative selection < replication rate: exponential growth of transposable elements
Adding selection for AMR genes

- AMR did emerge, but it is not linked to a transposon at all...
Adding selection for AMR genes

- Pulsing for antibiotic resistance: still TEs and AMR do not get linked...
Adding selection for AMR genes

- Pulsing for antibiotic resistance: still TEs and AMR do not get linked...
Okay, so getting genes "linked" to beneficial genes is really hard...

But I started noticing something else!
What if jumping causes damage?

Transposon-induced mutations (TIMs)
What if jumping causes damage?

Transposons result in genome streamlining, which makes cells fitter!
How long does it take to go extinct?

But wait, isn’t it better to take very long to go extinct?!
Immediately dying pays off!

(a) TE-abundance in extinct lineages before and after streamlining

- Long extinction time, many TEs produced
- Shorter extinction time, fewer TEs produced
Immediately dying pays off: altruistic suicide!

(c) Streamlined genomes outcompete non-streamlined genomes by preventing TE proliferation (cartoon)
Setting yourself up for surprise

• As a modeller, it can be hard to decide what NOT to put into your model (especially if you enjoy programming)

• By adding complexity and degrees of freedom, you allow a model to surprise you!

• This can reduce your bias (we can’t nullify it of course, but it helps)
Conclusions project #3

On the gene-level (TEs)
- Non-coding DNA is their "resource"
- Not clear why carrying a beneficial gene pays off...

On the cell-level
- Better to encode AMR on the chromosome
- Reducing non-coding DNA can prevent TEs from taking over
- This is a group-level effect (it only works in space)
What about more “clever” TEs?

Every “pearl” carries an insertion-site parameter between 0 and 1
Every transposon additionally has a target-site between 0 and 1
Highly specific TEs care about this (has to match). Non-specific TEs don’t.

Transposons can evolve to be nasty (non-specific) or show some restraint (high specificity)
Dynamics with/without a beneficial gene

- Every TE copy reduces fitness with 0.02 (starting from 1)
- Optionally, having 1 or more copies give a 0.12 fitness benefit (cargo gene)
Dynamics with/without a beneficial gene

d. At low eDNA uptake (0.01), highly specific TEs evolve

Can anyone guess why this didn’t happen at high HGT?
Dynamics with/without a beneficial gene

e. Without eDNA uptake, the fate of TEs is no different than any other gene

Graph showing the time course of total host population, TE abundance, and average TE specificity with and without a beneficial cargo gene.
Side-by-side comparison

c. With high eDNA uptake (0.05), non-specific TEs evolve regardless of the beneficial cargo gene

Note: HGT rate is not only a parameters, but also depends on density of hosts

d. At low eDNA uptake (0.01), highly specific TEs evolve

e. Without eDNA uptake, the fate of TEs is no different than any other gene

[Graphs and images showing comparative analysis of TE evolution under different conditions]
Conclusions project #4

• Evolved behaviour of TES driven by frequency of horizontal transfer, and carrying a beneficial gene is not that impactful.

• At high HGT, parasitism. At low HGT, less parasitic ("commensal").

• Without HGT, the evolutionary fate of a TE is like any other gene. If you’re bad, you get lost, if you’re good, you replicate with the chromosome.

Why do so many TEs and other selfish genetic elements (e.g. phages) encode host-beneficial genes, if it doesn’t appear to pay off very well?
Maybe the answer to "why do SGEs carry beneficial traits" is:
they don't!
Perhaps SGEs are simply masters of manipulation!
Take home messages

• High HGT does not mean bacteria are not individuals anymore, but the lines are definitely more blurry than without HGT.

• HGT allows certain genes to be "rescued", but this includes good and bad genes!

• Mobile elements are their own Darwinian entities, that can live a life of their own (often in conflict with the bacteria)

• Spatial structure is important in preventing or stabilising interactions with parasitic mobile elements

• Impacts of mobile elements goes beyond microbial growth rates: it also shapes the stability and health of plants, animals, and perhaps entire ecosystems