



Hacking Life – Decoding and manipulating cells with AI

Moritz Schaefer, PhD

38c3 (2024)



Hacking cells: What's possible?



Hacking cells: What's possible?

Glowing animals (under UV)





Hacking cells: What's possible?

Glowing animals (under UV)



"Easy"



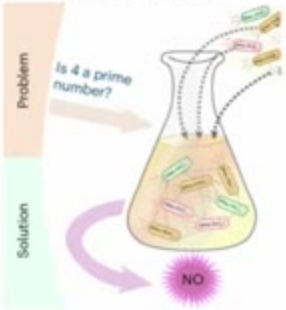
Hacking cells: What's possible?

Glowing animals (under UV)



"Easy"

Cells as biocomputers





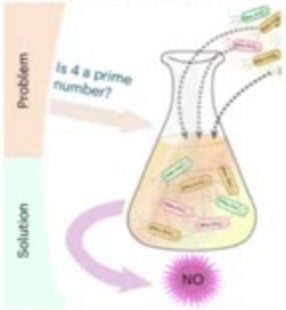
Hacking cells: What's possible?

Glowing animals (under UV)



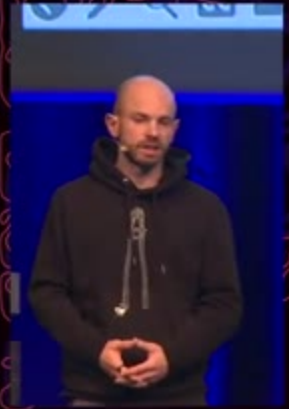
"Easy"

Cells as biocomputers



"Hard"

Bonnerjee et al., 2024, nature chemical biology



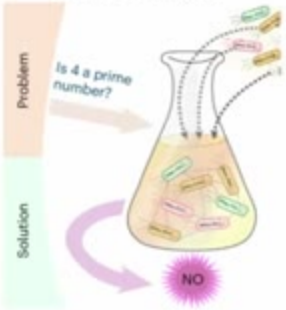
Hacking cells: What's possible?

Glowing animals (under UV)



"Easy"

Cells as biocomputers



"Hard"

Gene effect prediction



Bonnerjee et al., 2024, nature chemical biology



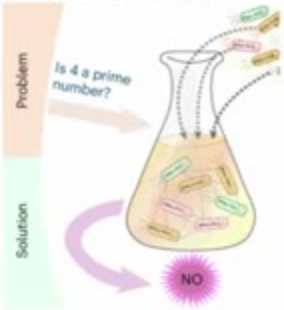
Hacking cells: What's possible?

Glowing animals (under UV)



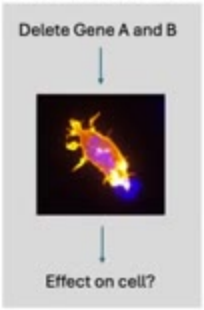
"Easy"

Cells as biocomputers



"Hard"

Gene effect prediction



"Unsolved"

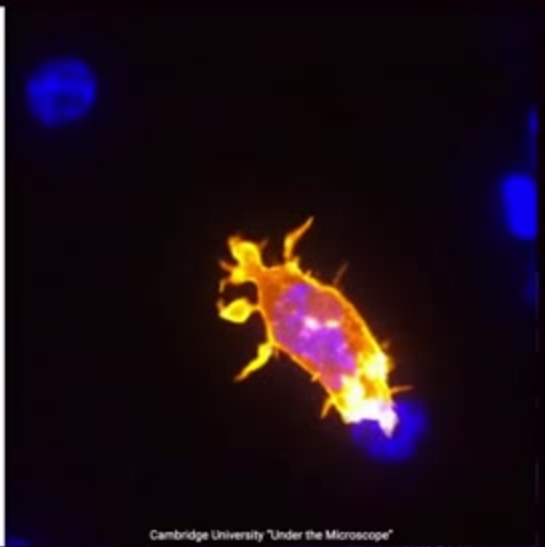
Bonnerjee et al., 2024, nature chemical biology





Outline

- What are cells?



Cambridge University "Under the Microscope"

Cells are complex



Cellscape, Tyler DeWitt (YouTube)

Cells are complex



Cellscape, Tyler DeWitt (YouTube)

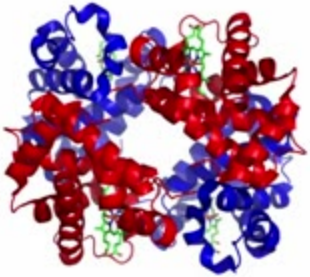
Cells are complex



Cellscape, Tyler DeWitt (YouTube)

Cells are driven by genes and proteins

Hemoglobin (~10,000 atoms)



Proteins

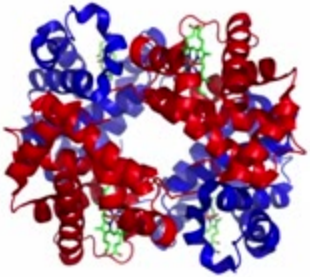
Genes/DNA (n=20,000)

Wikimedia CC BY-SA 3.0



Cells are driven by genes and proteins

Hemoglobin (~10,000 atoms)



Proteins

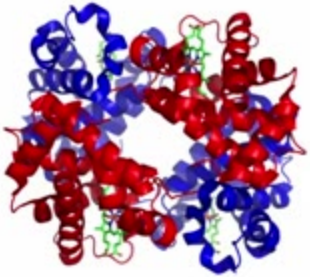
Genes/DNA (n=20,000)

Wikimedia CC BY-SA 3.0



Cells are driven by genes and proteins

Hemoglobin (~10,000 atoms)



Proteins



Genes/DNA (n=20,000)

Wikimedia CC BY-SA 3.0



Cells are driven by genes and proteins



Proteins



mRNA



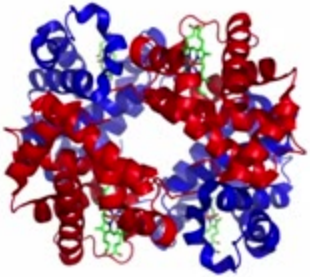
Genes/DNA (n=20,000)

Wikimedia CC BY-SA 3.0



Cells are driven by genes and proteins

Hemoglobin (~10,000 atoms)



Proteins



mRNA



Genes/DNA (n=20,000)

Wikimedia CC BY-SA 3.0



mRNA provides a window into the state of a cell







mRNA provides a window into the state of a cell





Gene A: on
Gene B: off
...
20,000



mRNA provides a window into the state of a cell





Gene A: on
Gene B: off
...
20,000



mRNA provides a window into the state of a cell





"Sequencer"

Gene A: 73
Gene B: 2
...
20,000



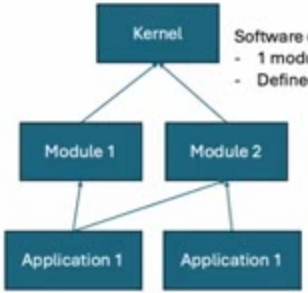
Research studies





“Nothing in Biology Makes Sense Except in the Light of Evolution”

Theodosius Dobzhansky

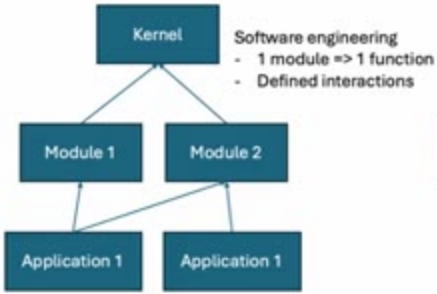


Software engineering
- 1 module => 1 function
- Defined interactions

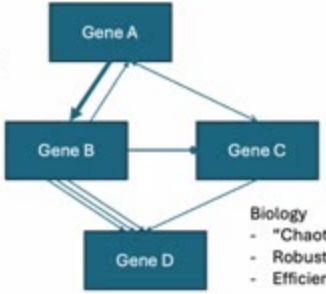


“Nothing in Biology Makes Sense Except in the Light of Evolution”

Theodosius Dobzhansky



Software engineering
- 1 module => 1 function
- Defined interactions



Biology
- “Chaotic”
- Robust
- Efficient



Intro Recap





Intro Recap





Multimodal LLM decodes images...



What does this image show?



Multimodal LLM decodes images...



What does this image show?

This image shows a logo featuring a stylized orange human figure leaping forward, with a bright yellow circle resembling a sun. The design conveys a sense of motion,

Multimodal LLM decodes images...



What does this image show?

This image shows the logo for the Rajiv Gandhi International Cricket Stadium, an emblem used by the Hyderabad Cricket Association in India.

The image shows a stylized rocket ship in flight.

Multimodal LLM decodes images… and cells



What does this image show?



This image shows the logo for the Rajiv Gandhi International Cricket Stadium, an emblem used by the Hyderabad Cricket Association in India.

The image shows a stylized rocket ship in flight.

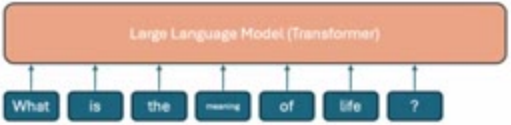


What is the biological function of this cell?

This is an immune cell in an active killing state

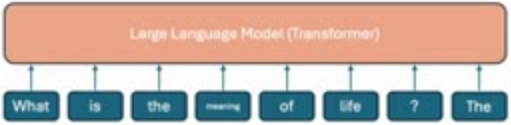
- Multimodal LLM architecture?
- Training data?

Multimodal LLM architecture...



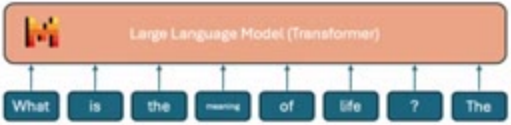


Multimodal LLM architecture...



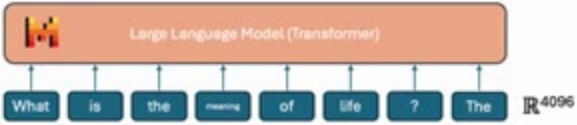


Multimodal LLM architecture...



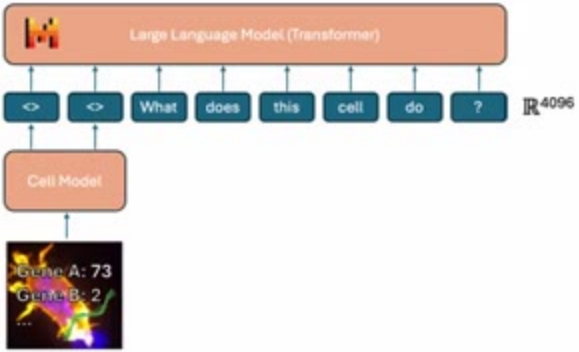


Multimodal LLM architecture...



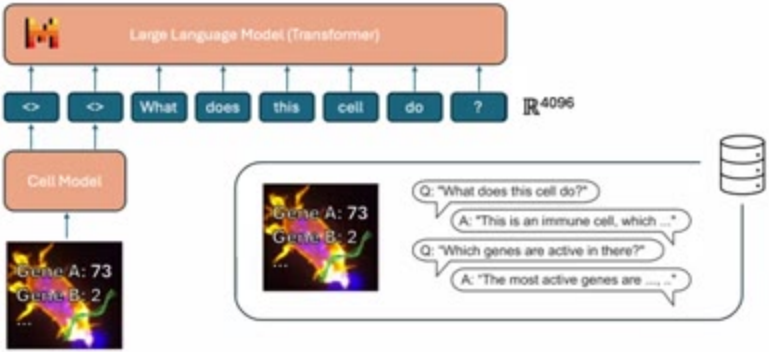


Multimodal LLM architecture...



Theodoris et al., 2023 (Nature, "Geneformer")

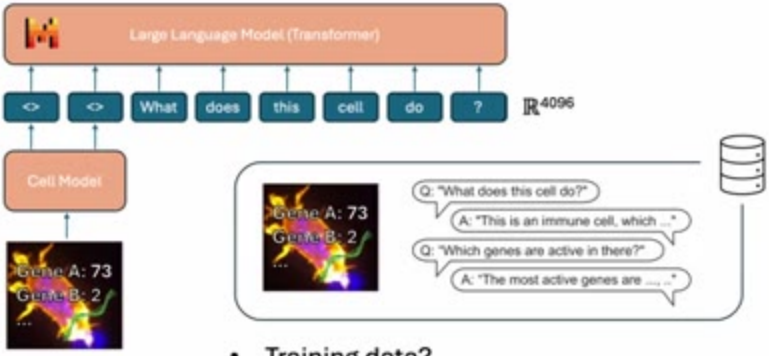
Multimodal LLM architecture… and training



Theodoris et al., 2023 (Nature, "Geneformer")

Liu et al., 2023 (NeurIPS, "Visual instruction tuning")

Multimodal LLM architecture… and training



- Training data?

Theodoris et al., 2023 (Nature, "Geneformer")

Liu et al., 2023 (NeurIPS, "Visual instruction tuning")



Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata





Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata





Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata



ChatGPT

Prompt: *Filter biological information and write a short abstract*

```
Examination of
molecular differences
between metatstatic
versus non-metatstatic
in stage I and stage II
```

- `growth_protocol: None`



Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata



ChatGPT

Prompt: *Filter biological information and write a short abstract*

```
Examination of
molecular differences
between metatstatic
versus non-metatstatic
in stage I and stage II
```

- `growth_protocol: None`



Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata



Examination of molecular differences between metatstatic versus non-metatstatic in stage I and stage II

- growth_protocol: None

ChatGPT

Prompt: *Filter biological information and write a short abstract*

Renal cell carcinoma tissue sample taken from a male patient at stage 2, with no metastasis, preserved in FFPE blocks

Q: "What does this cell do?"

A: "This is an immune cell, which ..."

Q: "Which genes are active in there?"

A: "The most active genes are ..... ."

100,000x



Semantic data curation with LLMs



- 20,000 studies
- 1 mio. samples
- metadata

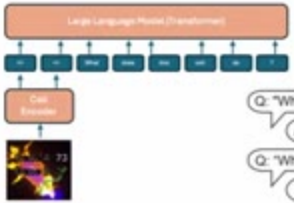


Examination of molecular differences between metatstatic versus non-metatstatic in stage I and stage II

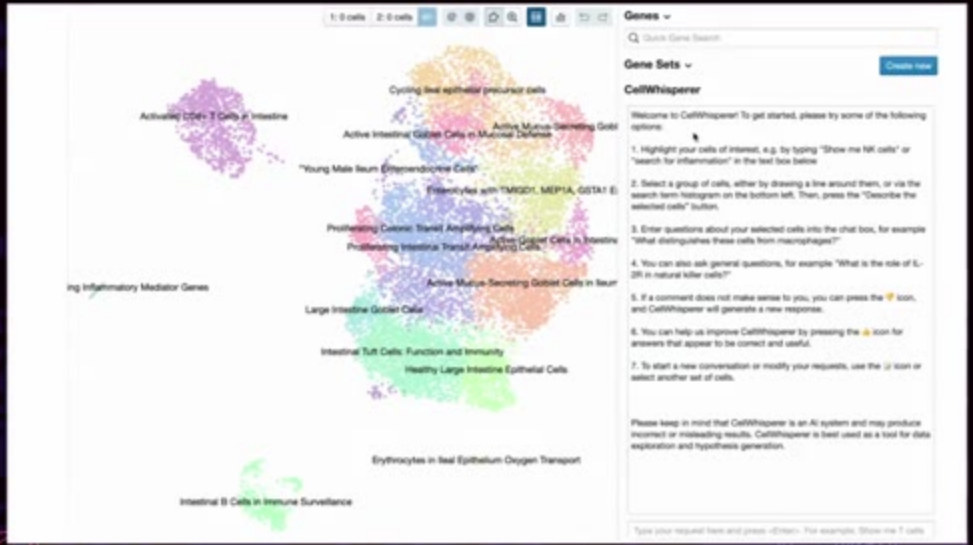
- growth_protocol: None



Renal cell carcinoma tissue sample taken from a male patient at stage 2, with no metastasis, preserved in FFPE blocks







Genes

Quick Gene Search

Gene Sets

Create new

CellWhisperer

Welcome to CellWhisperer! To get started, please try some of the following options:

1. Highlight your cells of interest, e.g. by typing "Show me NK cells" or "search for inflammation" in the text box below.

2. Select a group of cells, either by drawing a line around them, or via the search term histogram on the bottom left. Then, press the "Describe the selected cells" button.

3. Enter questions about your selected cells into the chat box, for example "What distinguishes these cells from macrophages?"

4. You can also ask general questions, for example "What is the role of IL-2R in natural killer cells?"

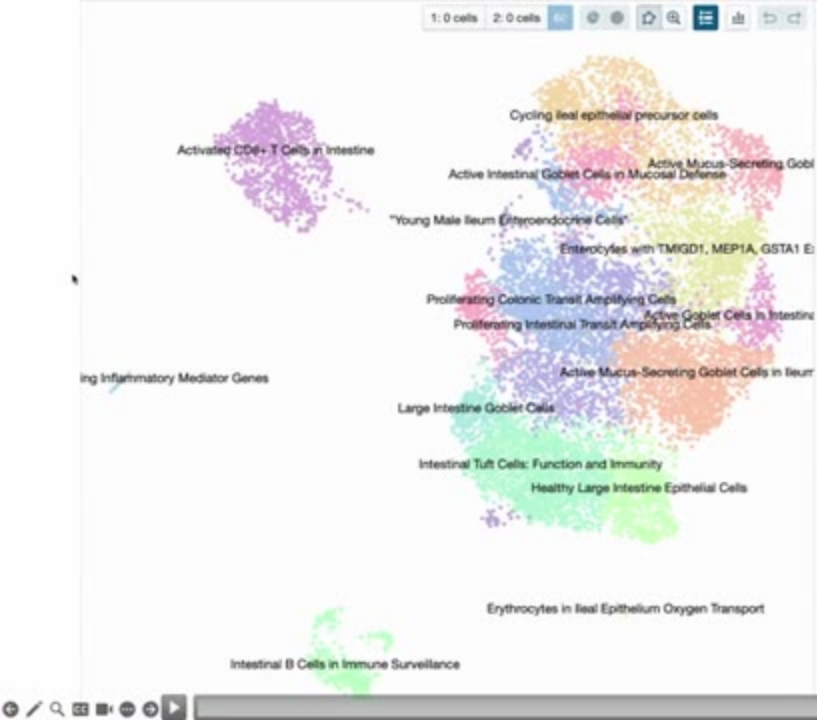
5. If a comment does not make sense to you, you can press the 👎 icon, and CellWhisperer will generate a new response.

6. You can help us improve CellWhisperer by pressing the 👍 icon for answers that appear to be correct and useful.

7. To start a new conversation or modify your requests, use the ↺ icon or select another set of cells.

Please keep in mind that CellWhisperer is an AI system and may produce incorrect or misleading results. CellWhisperer is best used as a tool for data exploration and hypothesis generation.

Type your request here and press <Enter>. For example, Show me T cells.



Genes

Quick Gene Search

Gene Sets

Create new

CellWhisperer

Welcome to CellWhisperer! To get started, please try some of the following options:

1. Highlight your cells of interest, e.g. by typing "Show me NK cells" or "search for inflammation" in the text box below

2. Select a group of cells, either by drawing a line around them, or via the search term histogram on the bottom left. Then, press the "Describe the selected cells" button.

3. Enter questions about your selected cells into the chat box, for example "What distinguishes these cells from macrophages?"

4. You can also ask general questions, for example "What is the role of IL-2R in natural killer cells?"

5. If a comment does not make sense to you, you can press the 👎 icon, and CellWhisperer will generate a new response.

6. You can help us improve CellWhisperer by pressing the 👍 icon for answers that appear to be correct and useful.

7. To start a new conversation or modify your requests, use the 📝 icon or select another set of cells.

Please keep in mind that CellWhisperer is an AI system and may produce incorrect or misleading results. CellWhisperer is best used as a tool for data exploration and hypothesis generation.

Example: Show me T cells



Genes

Quick Gene Search

Gene Sets

Create new

CellWhisperer

Welcome to CellWhisperer! To get started, please try some of the following options:

1. Highlight your cells of interest, e.g. by typing "Show me NK cells" or "search for inflammation" in the text box below

2. Select a group of cells, either by drawing a line around them, or via the search term histogram on the bottom left. Then, press the "Describe the selected cells" button.

3. Enter questions about your selected cells into the chat box, for example "What distinguishes these cells from macrophages?"

4. You can also ask general questions, for example "What is the role of IL-2R in natural killer cells?"

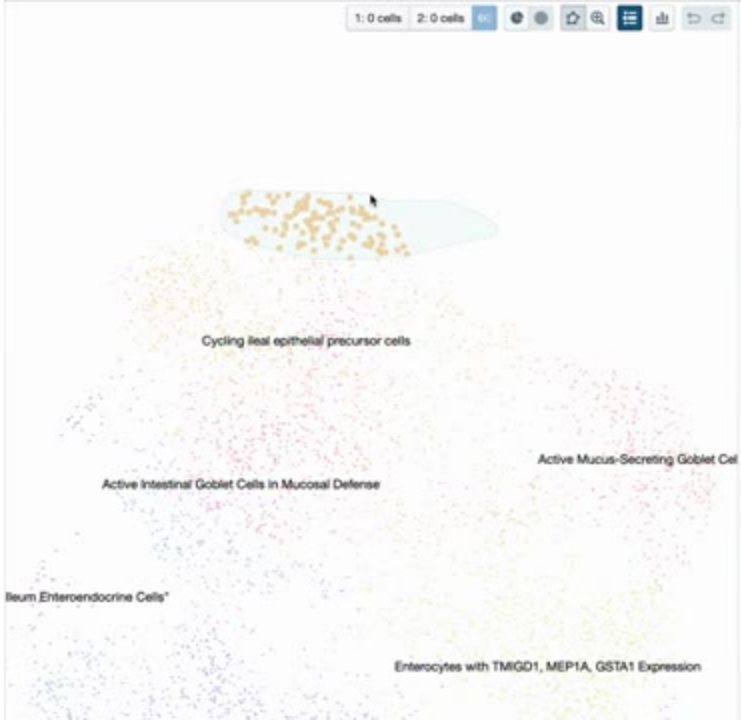
5. If a comment does not make sense to you, you can press the 👎 icon, and CellWhisperer will generate a new response.

6. You can help us improve CellWhisperer by pressing the 👍 icon for answers that appear to be correct and useful.

7. To start a new conversation or modify your requests, use the 📝 icon or select another set of cells.

Please keep in mind that CellWhisperer is an AI system and may produce incorrect or misleading results. CellWhisperer is best used as a tool for data exploration and hypothesis generation.

Type your request here and press <Enter>. For example: Show me T cells



Genes

Quick Gene Search

Gene Sets

Create new

CellWhisperer

Welcome to CellWhisperer! To get started, please try some of the following options:

1. Highlight your cells of interest, e.g. by typing "Show me NK cells" or "search for inflammation" in the text box below

2. Select a group of cells, either by drawing a line around them, or via the search term histogram on the bottom left. Then, press the "Describe the selected cells" button.

3. Enter questions about your selected cells into the chat box, for example "What distinguishes these cells from macrophages?"

4. You can also ask general questions, for example "What is the role of IL-2R in natural killer cells?"

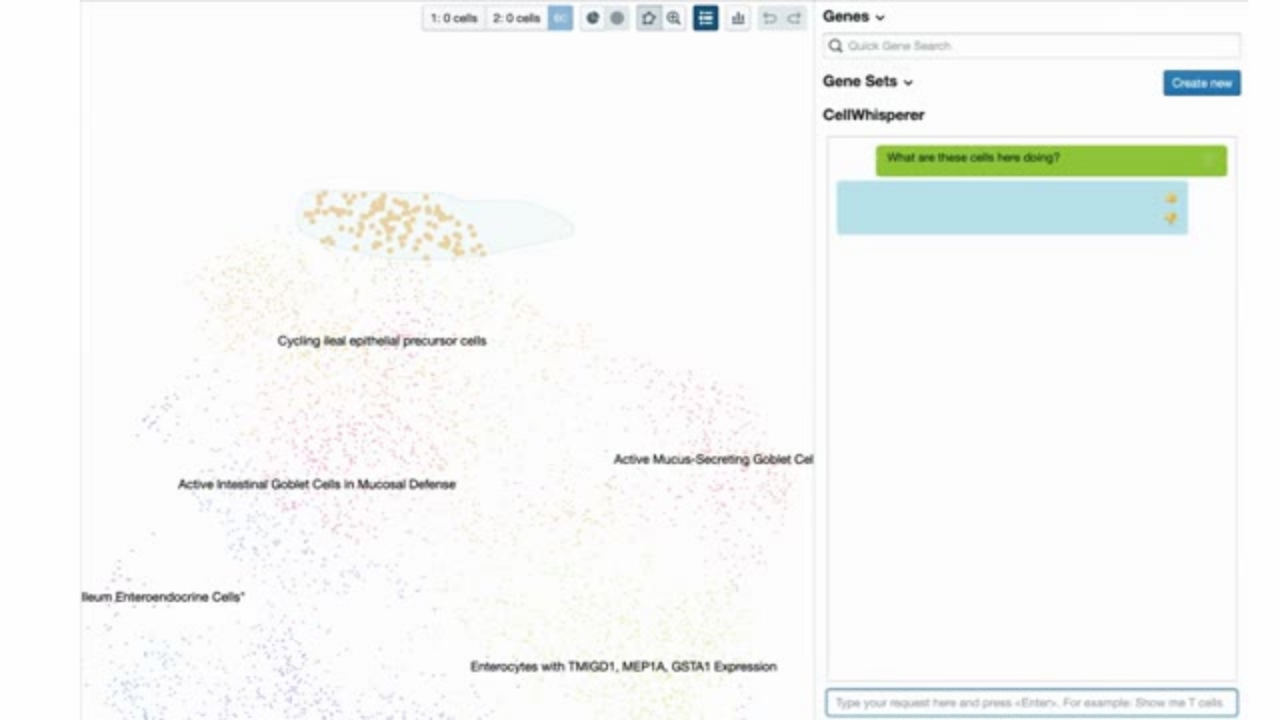
5. If a comment does not make sense to you, you can press the 👎 icon, and CellWhisperer will generate a new response.

6. You can help us improve CellWhisperer by pressing the 👍 icon for answers that appear to be correct and useful.

7. To start a new conversation or modify your requests, use the 📝 icon or select another set of cells.

Please keep in mind that CellWhisperer is an AI system and may produce incorrect or misleading results. CellWhisperer is best used as a tool for data exploration and hypothesis generation.

Type your request here and press <Enter>. For example: Show me T cells


1: 0 cells
2: 0 cells
Genes
Quick Gene Search
Gene Sets
Create new
CellWhisperer
What are these cells here doing?
Cycling ileal epithelial precursor cells
Active Mucus-Secreting Goblet Cel
Active Intestinal Goblet Cells in Mucosal Defense
Ileum Enteroendocrine Cells*
Enterocytes with TMIGD1, MEP1A, GSTA1 Expression
Type your request here and press <Enter>. For example: Show me T cells.

Hacking cells with AI. Why?

Hacking cells with AI. Why?



Hacking cells with AI. Why?



Hacking cells with AI. Why?



Blåhaj





Hacking cells with AI. Why?



Blåhaj





Hacking cells with AI. Why?



Blåhaj





Hacking cells with AI. Why?



Blåhaj







Hacking cells with AI-generated genes





Hacking cells with AI-generated genes



David Baker

The Nobel Prize in Chemistry 2024

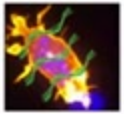
Prize motivation: "for computational protein design"



www.nobelprize.org/prizes/chemistry/2024/press-release/



Hacking cells with language AI

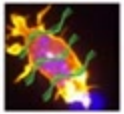




Gene A: 73
Gene B: 2
...



Hacking cells with language AI



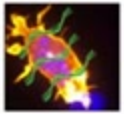


Gene A: 73
Gene B: 2
...

User: "How to make this immune cell more active?"



Hacking cells with language AI





Gene A: 73
Gene B: 2
...

User: "How to make this immune cell more active?"





Hacking cells with language AI

User: “How to make this immune cell more active?”



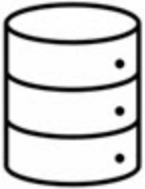
Hacking cells with language AI

User: “How to make this immune cell more active?”

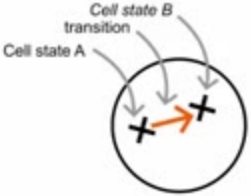
AI



Hacking cells with language AI

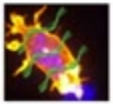


User: “How to make this immune cell more active?”



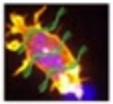


Recap





Recap



AI

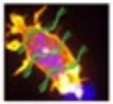


Risks

- Bioweapons?
- False predictions?
- ...



Recap



AI



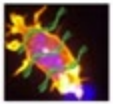
Risks

- Bioweapons?
- False predictions?
- ...

⇒ Don't worry! (2024)



Recap



AI



Risks

- Bioweapons?
- False predictions?
- ...

⇒ Don't worry! (2024)

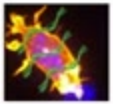
Vision

Hacking/Makers

- 1995: Software



Recap



AI



Risks

- Bioweapons?
- False predictions?
- ...

⇒ Don't worry! (2024)

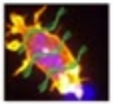
Vision

Hacking/Makers

- 1995: Software
- 2005: Hardware
- 2015: 3D printing
- 2025: Biology?



Recap



AI



Risks

- Bioweapons?
- False predictions?
- ...

⇒ Don't worry! (2024)

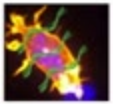
Vision

Hacking/Makers

- 1995: Software
- 2005: Hardware
- 2015: 3D printing
- 2025: Biology?



Recap



AI



Risks

- Bioweapons?
- False predictions?
- ...

⇒ Don't worry! (2024)

Vision

Hacking/Makers

- 1995: Software
- 2005: Hardware
- 2015: 3D printing
- 2025: Biology?





Acknowledgements

CENTER FOR MEDICAL DATA SCIENCE
MEDICAL UNIVERSITY OF VIENNA
Institute of Artificial Intelligence

CeMM



Peter Peneder* (Tomazou group)



Daniel Malzl (Menche group)



Mihaela Peyecheva (Bock group)



Jake Burton (Bock group)



Varun Sharma (Bock & Menche)



Anna Hakobyan (Cresswell group)



Thomas Krausgruber (Bock group)

Labs/PIs: **Eleni Tomazou** (CCRI)
Jörg Menche (MPL/CeMM/LBI-NetMed)
Christoph Bock (MedUni/CeMM)



Acknowledgements

CENTER FOR MEDICAL DATA SCIENCE
MEDICAL UNIVERSITY OF VIENNA
Institute of Artificial Intelligence

CeMM



Peter Peneder* (Tomazou group)



Daniel Malzl (Menche group)



Mihaela Peyecheva (Bock group)



Jake Burton (Bock group)



Varun Sharma (Bock & Menche)



Anna Hakobyan (Cresswell group)



Thomas Krausgruber (Bock group)

Labs/PIs: **Eleni Tomazou** (CCRI)
Jörg Menche (MPL/CeMM/LBI-NetMed)
Christoph Bock (MedUni/CeMM)



Acknowledgements

CENTER FOR MEDICAL DATA SCIENCE
MEDICAL UNIVERSITY OF VIENNA
Institute of Artificial Intelligence

CeMM



Peter Peneder* (Tomazou group)



Daniel Malzl (Menche group)



Mihaela Peyecheva (Bock group)



Jake Burton (Bock group)



Varun Sharma (Bock & Menche)



Anna Hakobyan (Cresswell group)



Thomas Krausgruber (Bock group)

Labs/PIs: **Eleni Tomazou** (CCRI)
Jörg Menche (MPL/CeMM/LBI-NetMed)
Christoph Bock (MedUni/CeMM)



Acknowledgements

CENTER FOR MEDICAL DATA SCIENCE
MEDICAL UNIVERSITY OF VIENNA
Institute of Artificial Intelligence

CeMM



Peter Peneder* (Tomazou group)



Daniel Malzl (Menche group)



Mihaela Peyecheva (Bock group)



Jake Burton (Bock group)



Varun Sharma (Bock & Menche)



Anna Hakobyan (Cresswell group)



Thomas Krausgruber (Bock group)

Labs/PIs: **Eleni Tomazou** (CCRI)
Jörg Menche (MPL/CeMM/LBI-NetMed)
Christoph Bock (MedUni/CeMM)

https://cellwhisperer.bocklab.org

https://moritzs.de




LLEGAL


LLEGAL
803


EGAL


EGAL


EGAL


EGAL


ILLEGAL


















GAL


EGAL




EGAL



Recorded by VOC
together with:

ags

Wissenschaftliche Arbeitsgemeinschaft
für Studio- und Senderfragen an der
TU Braunschweig e.V.

FEM

Forschungsgemeinschaft
elektronische Medien e.V.

creative commons

Attribution 4.0 International

https://creativecommons.org/licenses/by/4.0



For more talks and
conference recordings visit:

media.ccc.de