



① Few-Shot Bitstring Program Synthesis

(100+ programs)

Identity:

```
10100000 -> 10100000 SEP
11100011 -> 11100011 SEP
11001010 -> 11001010 SEP
11000000 -> ?
```

Bitwise NOT:

```
10100000 -> 01011111 SEP
11100011 -> 00011100 SEP
11001010 -> 11001010 SEP
11000000 -> ?
```

Majority 1:

```
10100000 -> 00000000 SEP
11100011 -> 11111111 SEP
11001110 -> 11111111 SEP
11000000 -> ?
```

Majority 1 + NOT:

```
10100000 -> 11111111 SEP
11100011 -> 00000000 SEP
11001110 -> 00000000 SEP
11000000 -> ?
```

Reverse:

```
10100000 -> 00000101 SEP
11100011 -> 11000111 SEP
11001110 -> 01110011 SEP
11000000 -> ?
```

② Modality-Specific Encoding:

Genomic Encoding

(a) Random 3 Nucleotides:

A T C G

(b) Map the bit strings:

“0” ⇒ “A”, “1” ⇒ “T”,
“SEP” ⇒ “G”, “-” ⇒ removed

Example:

```
1010 -> 1010SEP
TATATATAG
```

Linguistic Encoding

(a) Random 3 Digits:

0 1 2 3 4 5 6 7 8 9

(b) Map the bit strings:

“0” ⇒ “1”, “1” ⇒ “3”,
“SEP” ⇒ “6”, “-” ⇒ removed

Example:

```
1010 -> 1010SEP
313131316
```

③ Inference Model and Compute Accuracy

Genomic Encoding

k-shot demonstrations

```
AAAT TAAA G
ATAA AATA G
TTTA ATTT G
ATAT TATA G
AATT ?
```

Evo2 Greedy decode & evaluate

Linguistic Encoding

k-shot demonstrations

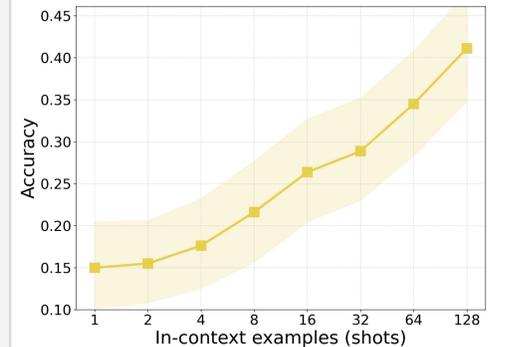
```
1113 3111 6
1311 1131 6
3331 1333 6
1313 3131 6
1133 ?
```

Qwen Greedy decode & evaluate

④ In-Context Learning Trends & Comparison

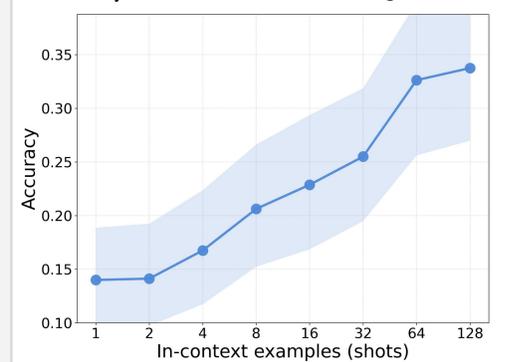
Genomic Models

Accuracy vs number of shots (Evo2-40B)



Linguistic Models

Accuracy vs number of shots (Qwen3-14B)



Models trained on the **genome** develop ICL in similar ways to models trained on **human language** - indicating ICL is a consequence of **next-token training on complex data**.

Our Contributions:

- First demonstration of **Emergent ICL** in **genomics models** (prior demonstrations are **Meta ICL**)
- Controlled comparison between **linguistic models** and **genomic models**
- Comparison across **model scale** and **shot count** - performance improves with both in linguistic and genomic models
- These findings challenge the notation that ICL is **language-specific** and explanations of ICL that rely on **properties of human language alone**. They instead suggest a **compression/prediction** view of ICL.

But what is emergent ICL?

Meta ICL	Emergent ICL
Explicitly trained for - model is taught to infer function from few-shot examples.	Emerges naturally as a byproduct of training on a complex modality (language, genome).
Tells us little about the content of the trained modality, more about prompt format.	Tells us whether some form of intelligence can emerge from predicting a given modality.