1 - Why are we here? (*ludic rhetorics*)

Tuesday afternoon

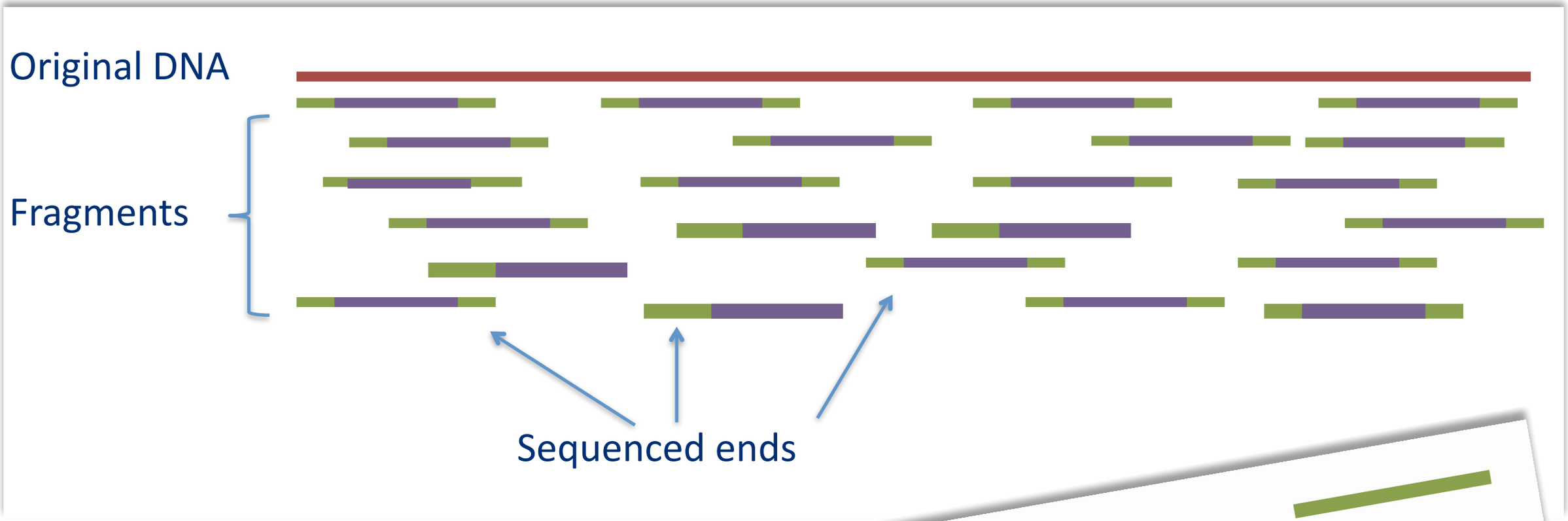
Bernardo J. Clavijo
Richard Smith-Unna
Gonzalo Garcia

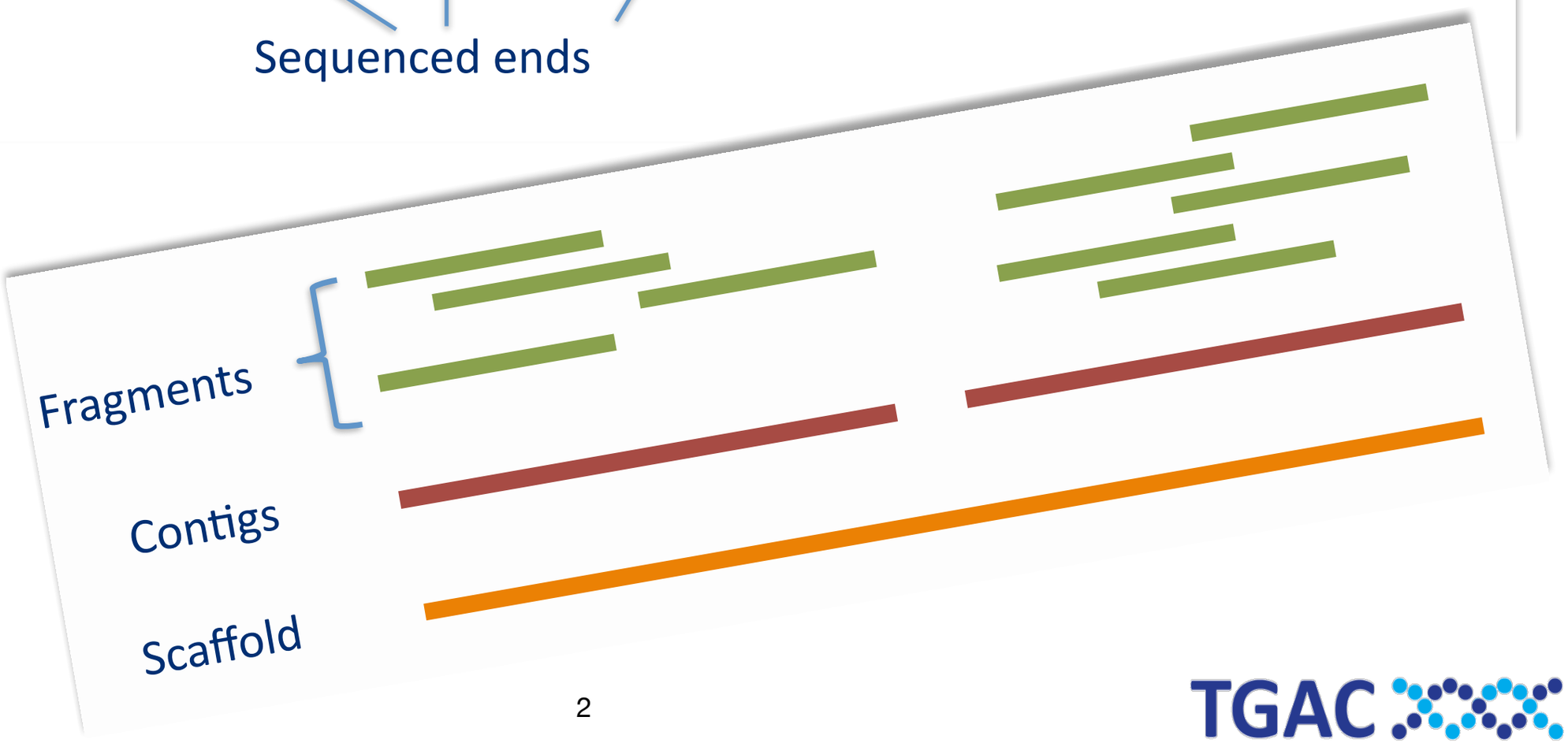




TGAC

The genome assembly problem (WGS)





TGAC

A correct assembly has:

The right *motifs,*

the correct number of times,

in correct order and position.

TGAC

Assembly Games

Black box processing

TGAC

Heuristic processing: using shortcuts

heuristic | ˌhjʊ(ə)ˈrɪstɪk |

adjective
enabling a person to discover or learn something for themselves. *a 'hands-on' or interactive heuristic approach to learning.*
• Computing proceeding to a solution by trial and error or by rules that are only loosely defined.

noun
a heuristic process or method.
• (**heuristics**) [usu. treated as sing.] the study and use of heuristic techniques.

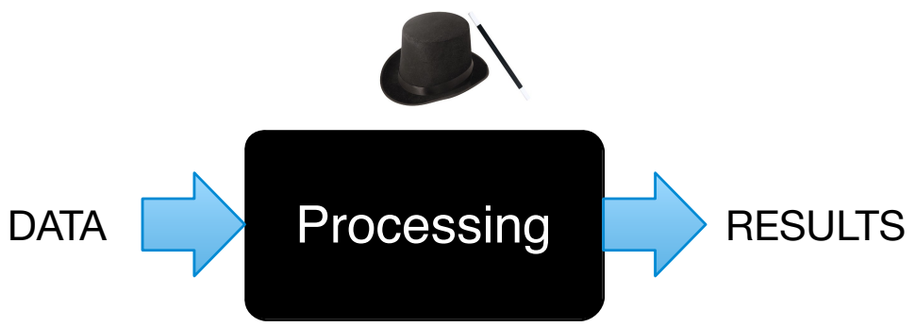
DERIVATIVES
heuristically adverb

ORIGIN early 19th cent.: formed irregularly from Greek ***heuriskein 'find'***.

exhaustive | ɪgˈzɔːstɪv, ɛg- |

adjective
including or considering all elements or aspects; fully comprehensive: *the guide outlines every bus route in exhaustive detail.*

DERIVATIVES
exhaustively adverb ,
exhaustiveness noun

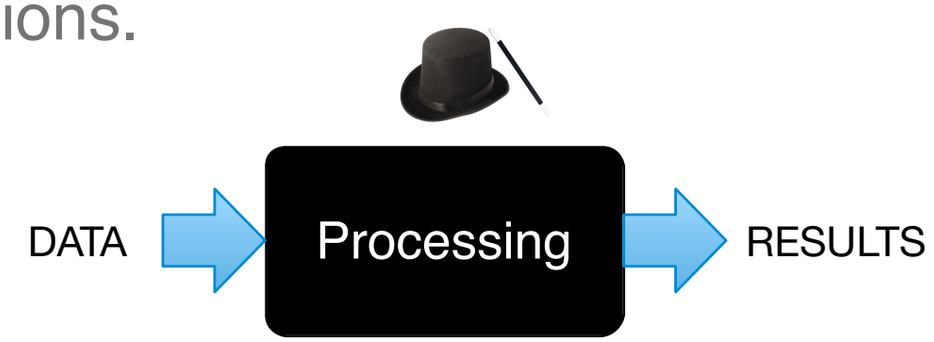


TGAC

Why use heuristics for genome assembly?

- The problem is not completely defined.

- Exhaustive methods are:
  - Too limited, thus producing simple partial solutions.
  - Too slow, not scaling well.

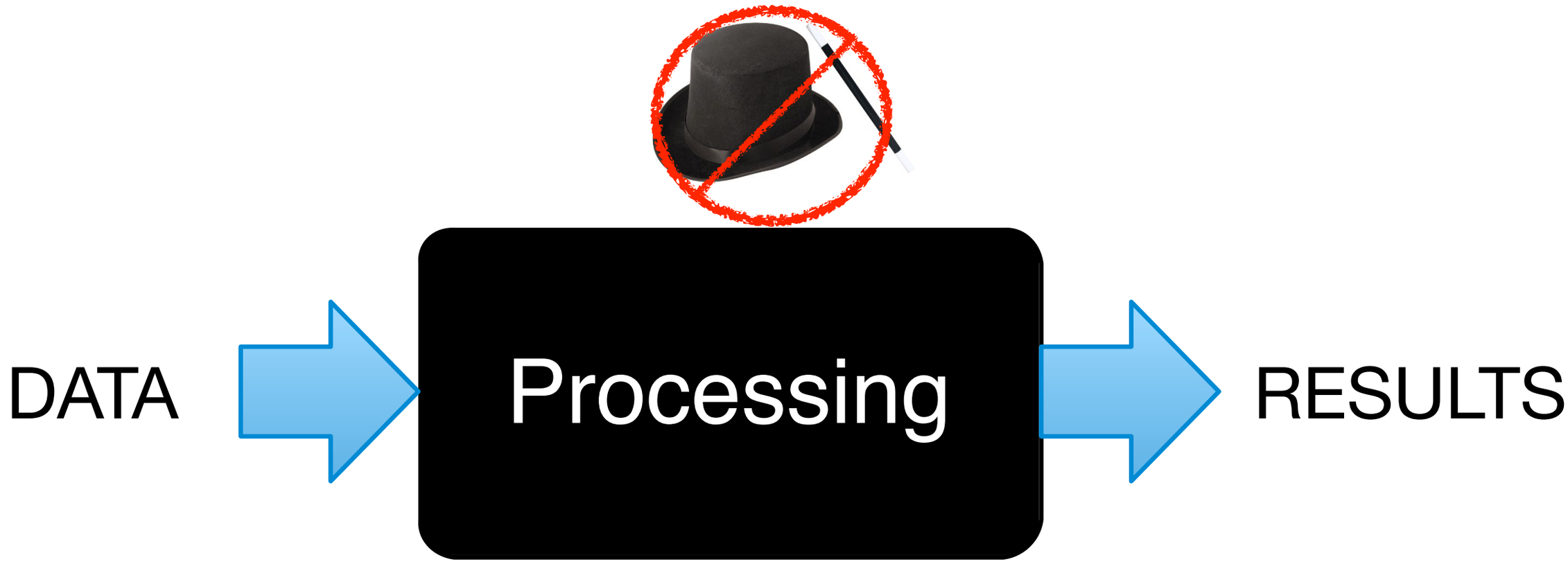


- Data varies too much and no good models are available.

- It is so much faster and easier… and it works! (sometimes, anyway)

TGAC

Black box processing done right



- Use good data, check its pre-conditions to be well processed.
- Know (roughly) how the processing works.
- Check soundness and sanity of results.

TGAC

Questions?