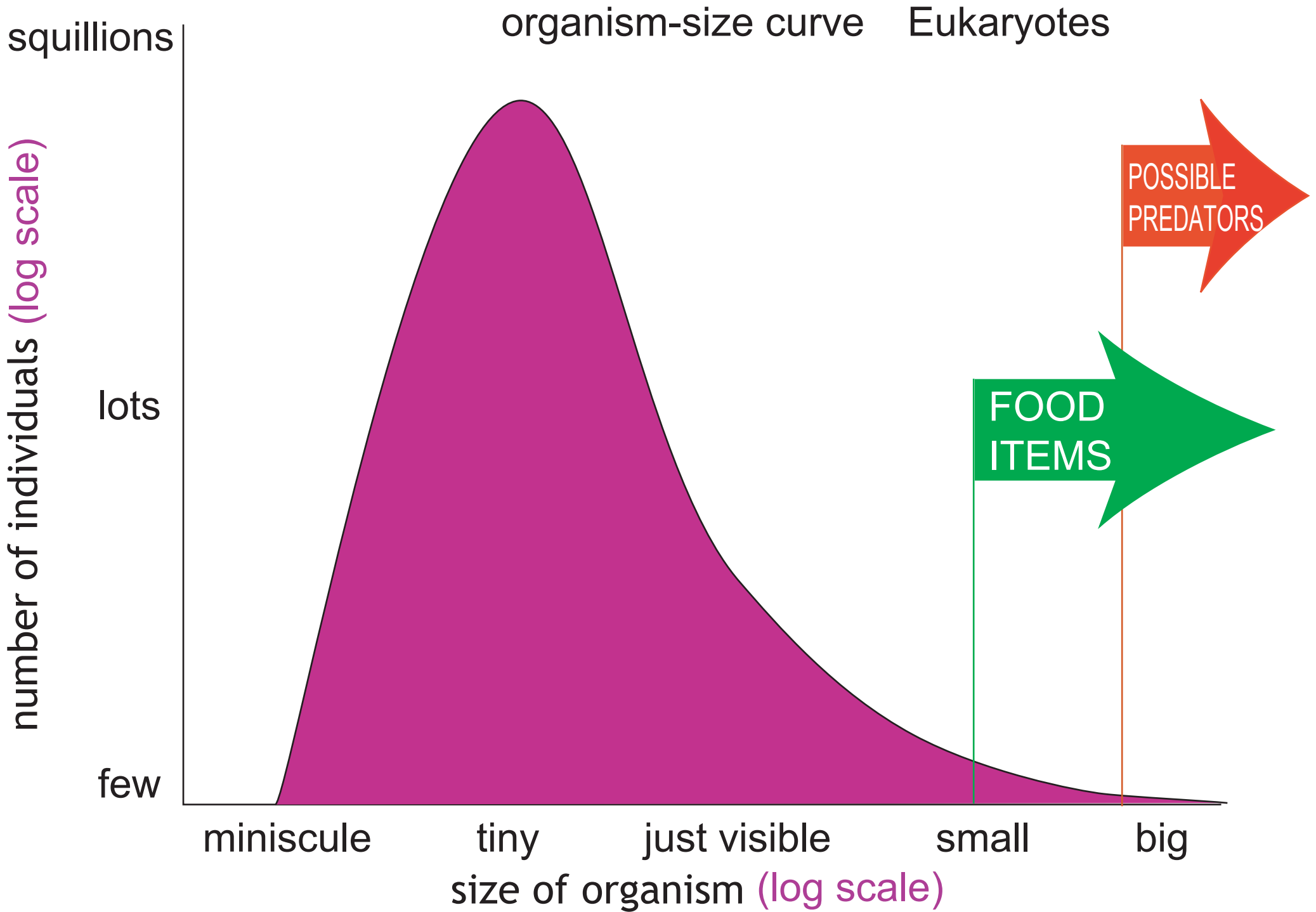




"... endless forms
most beautiful
and most wonderful
have been,
and are being,
evolved"

(Darwin 1859)

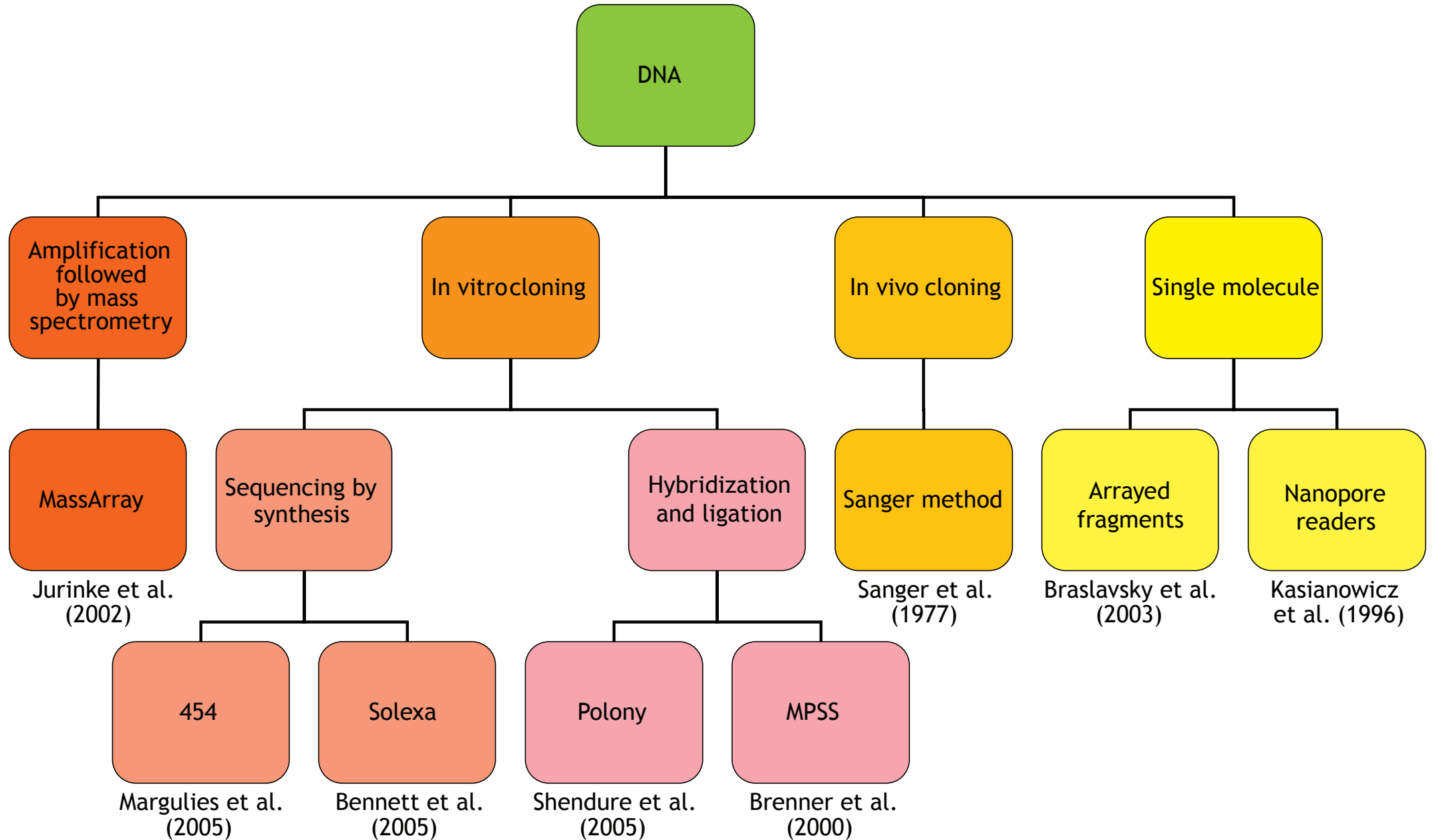
16S sequences	500000
18S sequences	125000
CO1 sequences <i>CBoL</i>	300000
rbcL sequences	35000
genomes (cellular)	500++
NCBI 'species' (rel158)	165000
<i>named 'species'</i>	1.5×10^6
<i>probable 'species'</i>	$1 \times 10^7+++$



Upcoming sequencing tech

**gigabases per hour =
genomes per day =
ecosystems per week**

Sequencing Genomes



From Neil Hall (2007) J. Exp. Biol.

Next generation now

SOLEXA

sequencing-by-synthesis

1 Gb / 3 days

35 bases x 30 10e6 reads

Next generation now

SOLEXA *s-b-s*

coming next year

50 bases

60-100 10e6 reads

4 Gb plus

Next generation now

454

pyrosequencing

100 Mb / 9 hours

250 bases x 400,000 reads

Next generation now

454 *pyrosequencing*

coming next year

500 bases

1,000,000 reads

0.5 Gb

Next generation [emerging]

SOLiD *MPSS short*

Polony *MPSS cheap*

nanopore sequencing *long reads*

Costs (now-ish)

mode	instrument	per base	per run
<i>MLB, PhD</i>	21k	10	0.0000036 Gb
Sanger	200k	1	0.00005 Gb
454	200k	0.01	0.1 Gb
SOLEXA	200k	0.001	1 Gb
SOLiD	200k	0.001	1 Gb

Costs (5 years-ish)

mode	instrument	per base
Sanger	200k	1
454	200k	0.01
SOLEXA	200k	0.001
SOLiD	200k	0.001
nanopore	10k	0.001
chip-seq	10k	0.0001
personal pyrosequencer	10k ... 1k	0.00001

To the Unseeable Animal

My Daughter:

"I hope there's an animal
Somewhere that nobody has ever seen.
And I hope nobody ever sees it."

Wendell Berry

