

# **MAPPING & GENE EXPRESSION AND GENOME EVOLUTION**

**13Nov15**

# ANNOUNCEMENTS

# MAPPING - BWT

Format specification: <http://samtools.github.io/hts-specs/SAMv1.pdf>

# MAPPING - BWT

## Alignment Fields

Col1

Col2

Col3

Col4

Col5

Col6

Col7

Col8

Col9

Col10

Col11

# MAPPING - BWT

<http://broadinstitute.github.io/picard/explain-flags.html>

# GENE EXPRESSION

# GENE EXPRESSION

<http://haroldpimentel.wordpress.com/2014/05/08/what-the-fpkm-a-review-rna-seq-expression-units/>

Inter- versus intra-sample comparison

# GENE EXPRESSION

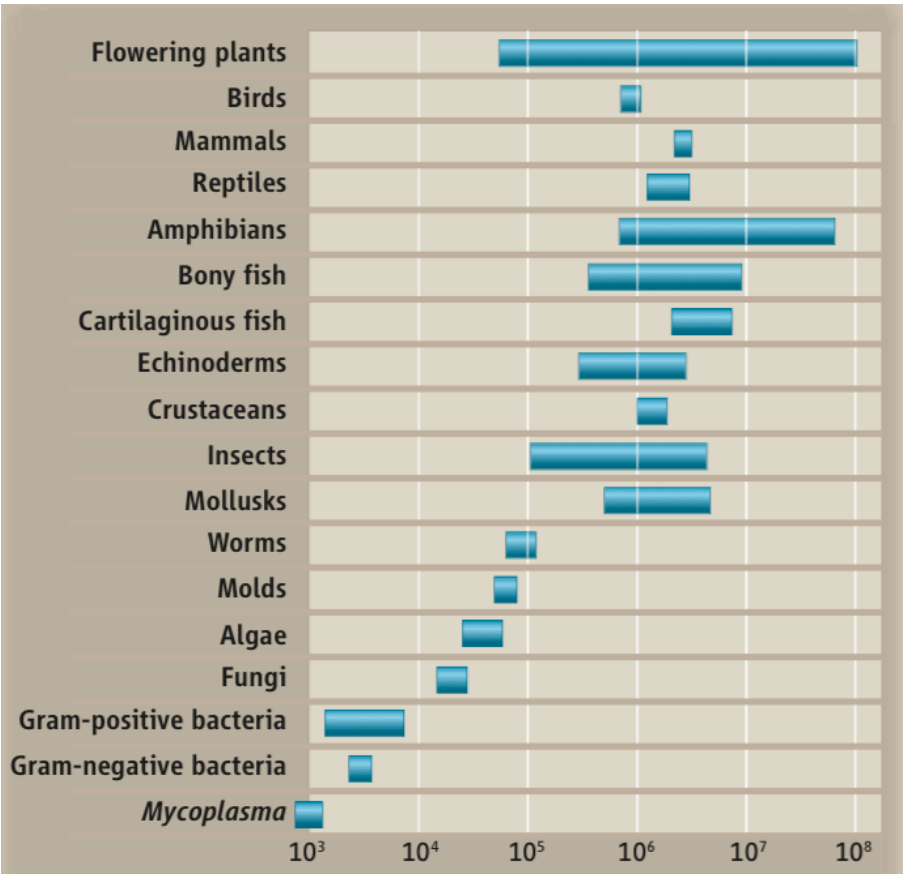
$$TPM_i = \frac{X_i}{\tilde{l}_i} * \left( \frac{1}{\sum_j \frac{X_j}{\tilde{l}_j}} \right) * 10^6$$



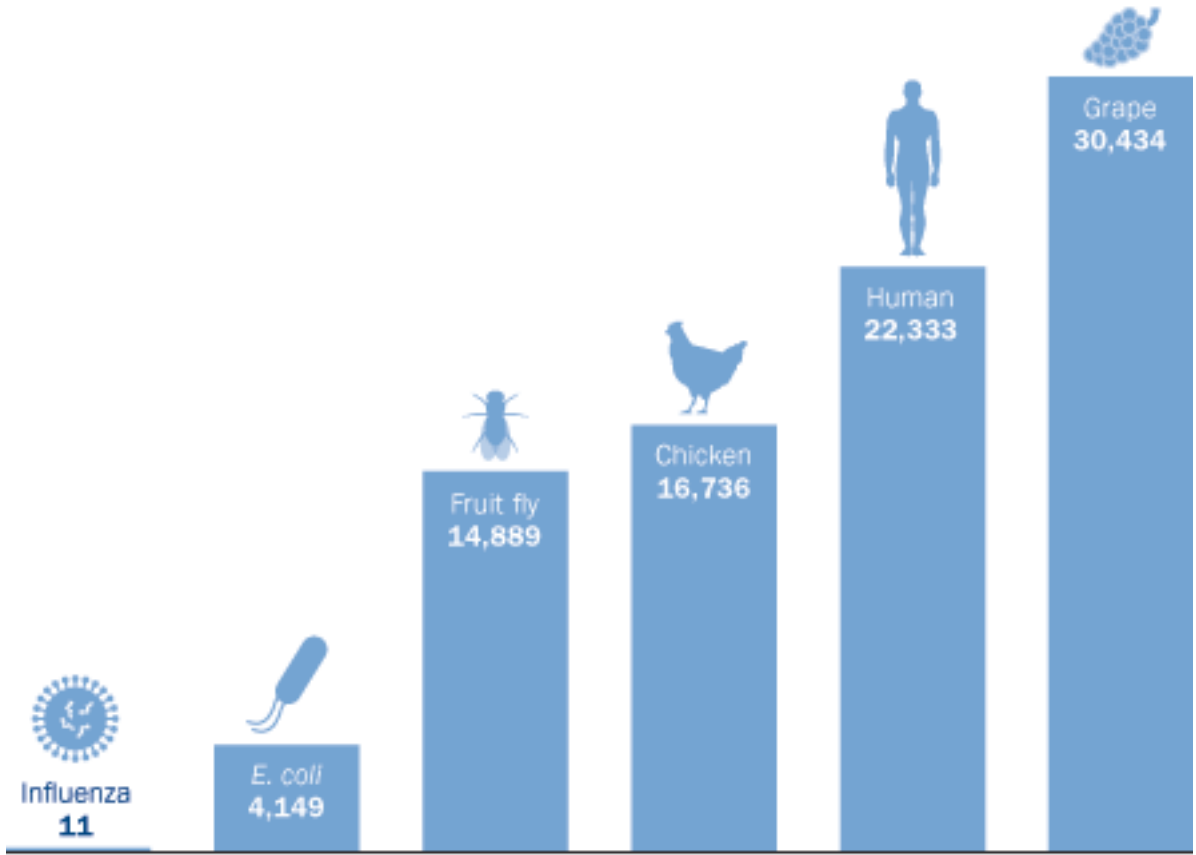
# GENOME DIVERSITY

How are genomes different?

# GENOME DIVERSITY: SIZE



# GENOME DIVERSITY: # GENES



# GENOME DIVERSITY:

Genome size unrelated to gene number

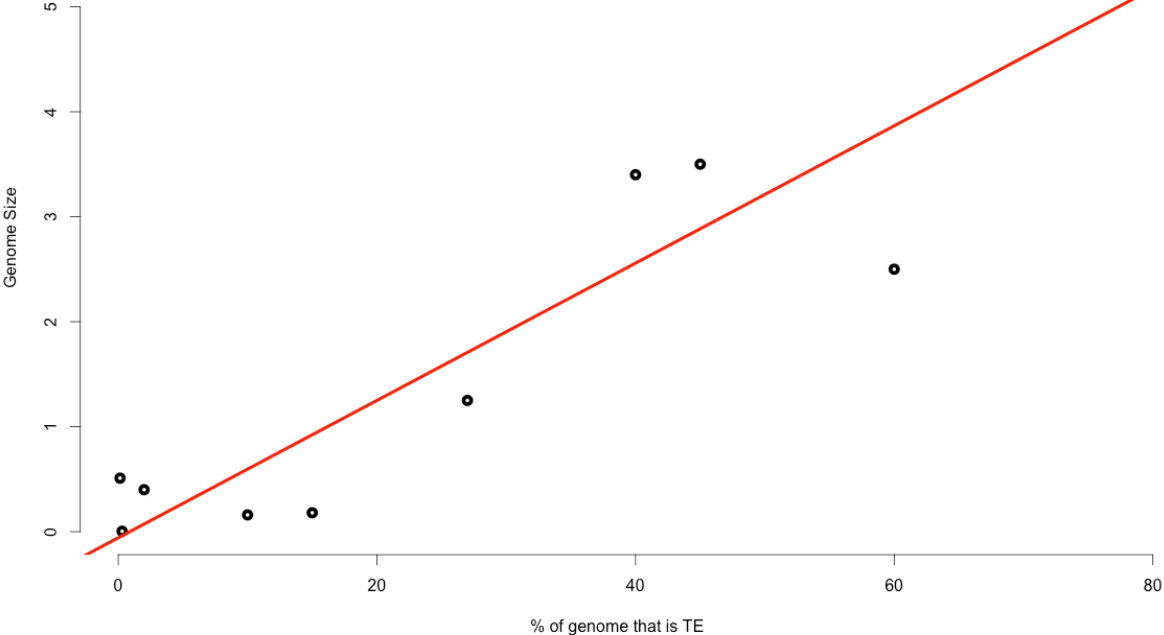
# GENOME DIVERSITY:

Genome size unrelated to gene number

Species	Common name	Genome size, pg	% TEs	Gene number
<i>Fritillaria assyriaca</i>	lily	127.4	95-99	
<i>Rana esculenta</i>	frog	5.6-8.0	77	
<i>Homo sapiens</i>	human	3.5	45	23,000
<i>Xenopus laevis</i>	frog	3.5	37	
<i>Mus musculus</i>	mouse	3.4	40	35,000
<i>Zea mays</i>	maize	2.5	60	
<i>Gallus domesticus</i>	hen	1.25	27	20,000
<i>Tetraodon nigroviridis</i>	fish	0.51	0.14	22,000
<i>Takifugu rubripes</i>	fish	0.4	2	31,000
<i>Anopheles gambiae</i>	malaria mosquito	0.28	16	14,000
<i>Drosophila melanogaster</i>	fruit fly	0.18	15-22	14,039
<i>Ciona intestinalis</i>	ascidian	0.16	10	15,500
<i>Arabidopsis thaliana</i>	arabidopsis	0.16	14	26,000
<i>Caenorhabditis elegans</i>	worm	0.1	12	20,060
<i>Saccharomyces cerevisiae</i>	yeasts	0.012	3-5	6,680
<i>Escherichia coli</i>	bacterium	0.0046	0.3	4,500

# GENOME DIVERSITY:

Genome size unrelated to gene number



# GENOME DIVERSITY:

Genome size unrelated to gene number