

# How many bacterial genes are there in the world?

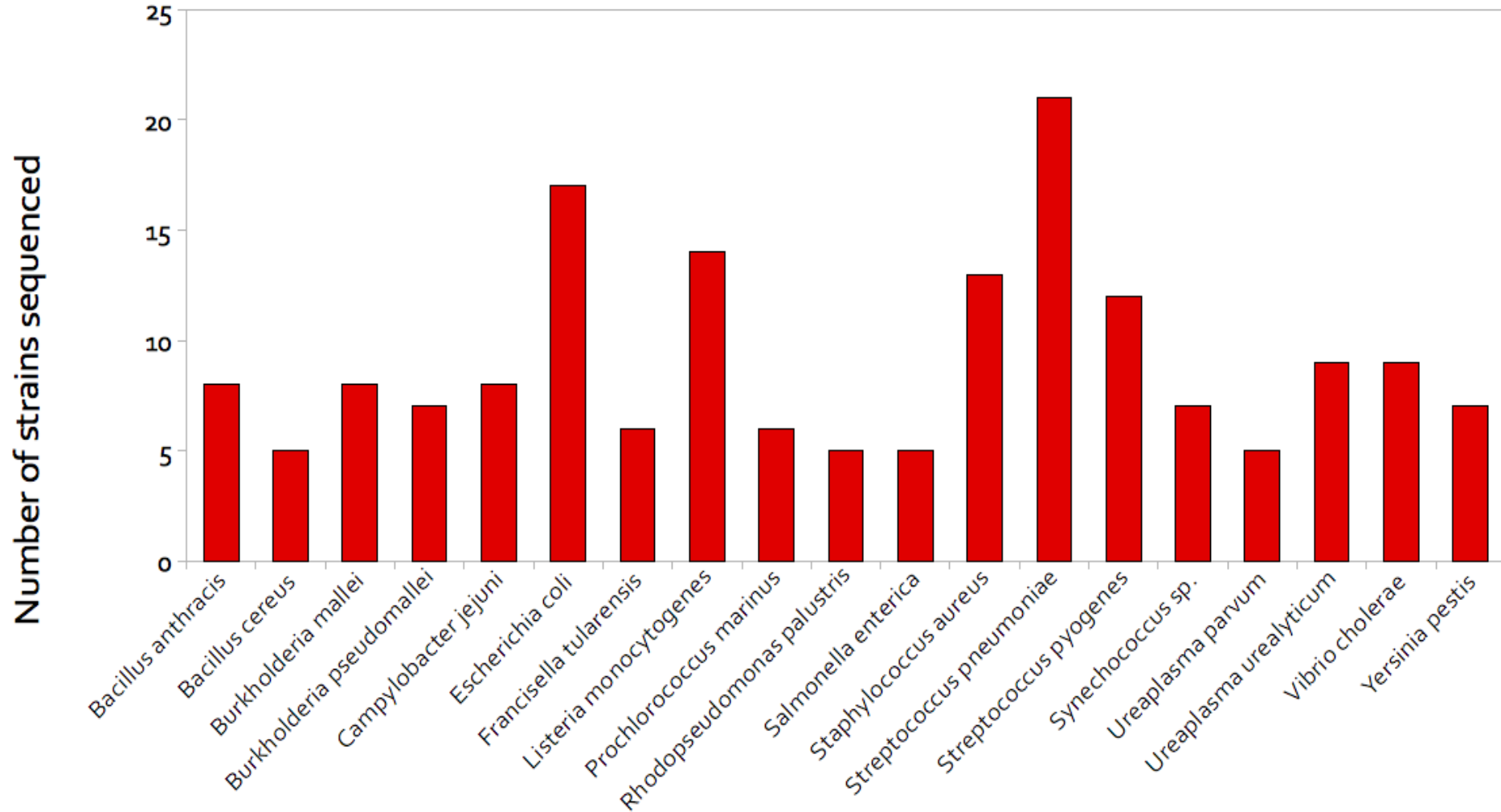
## Math guys:

Jim Nulton, Peter Salamon, Ben Felts

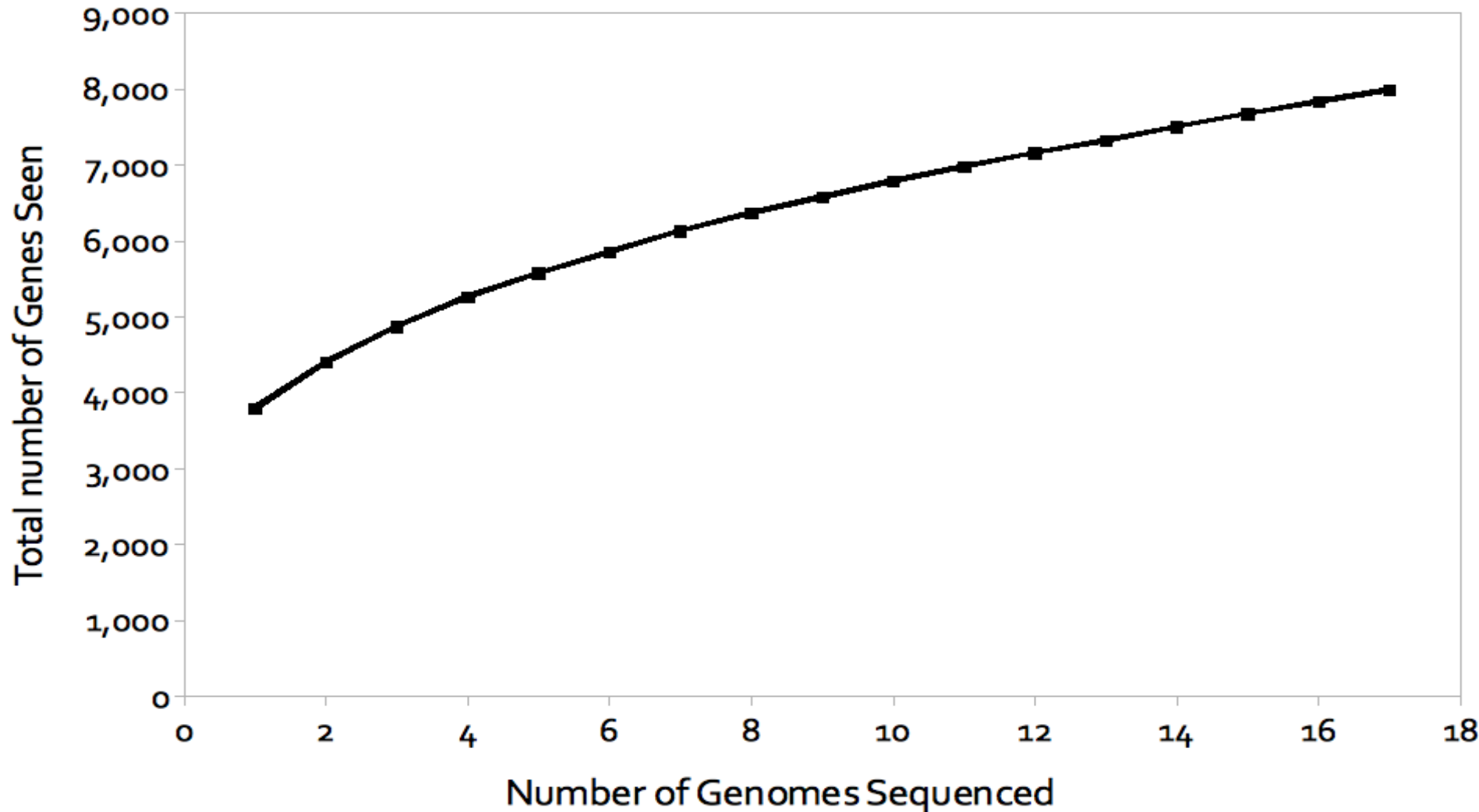
## Biology guys:

Liz Dinsdale, Forest Rohwer, Florent Angly

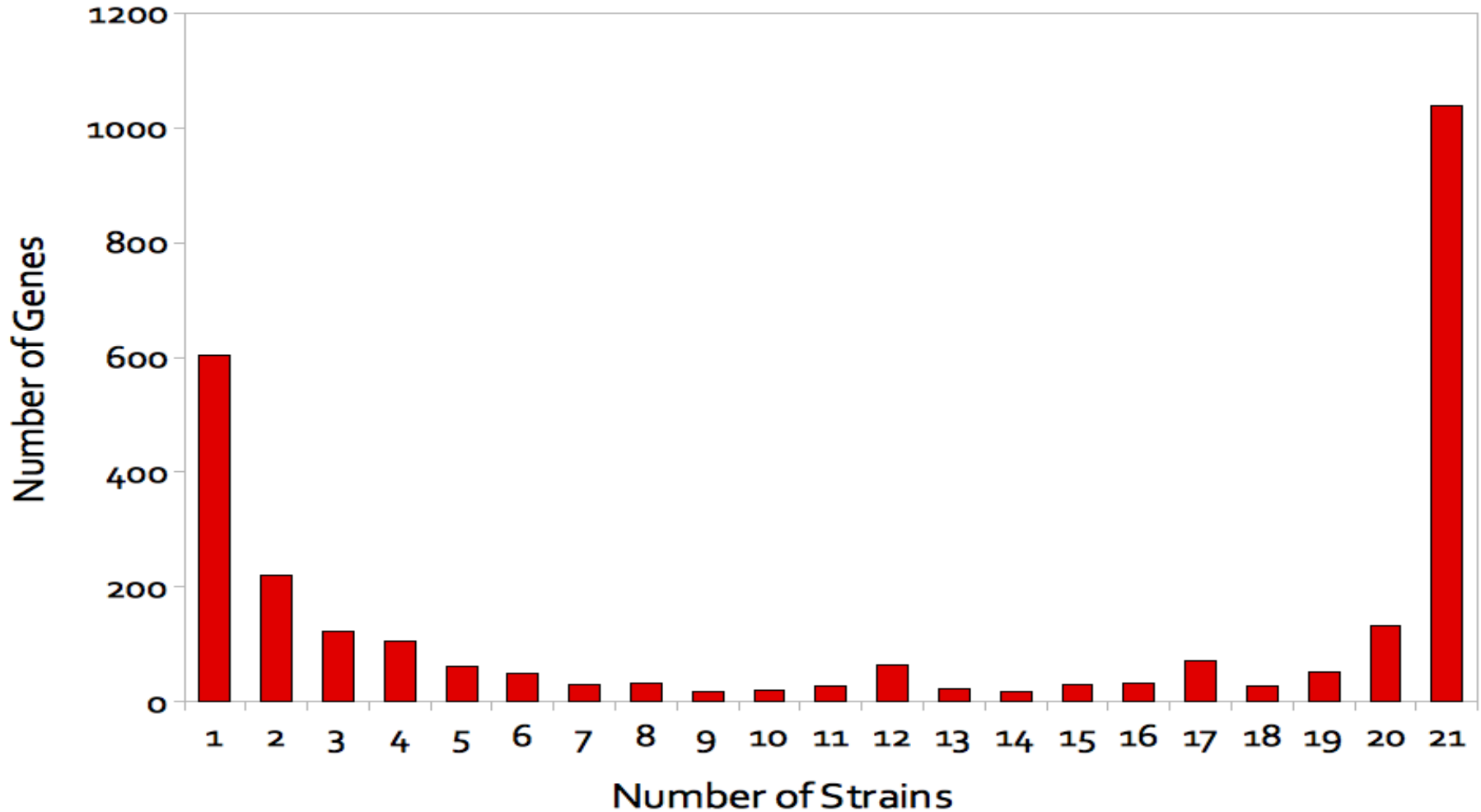
# How many genomes have we sequenced?



# Finding new genes in *E. coli*



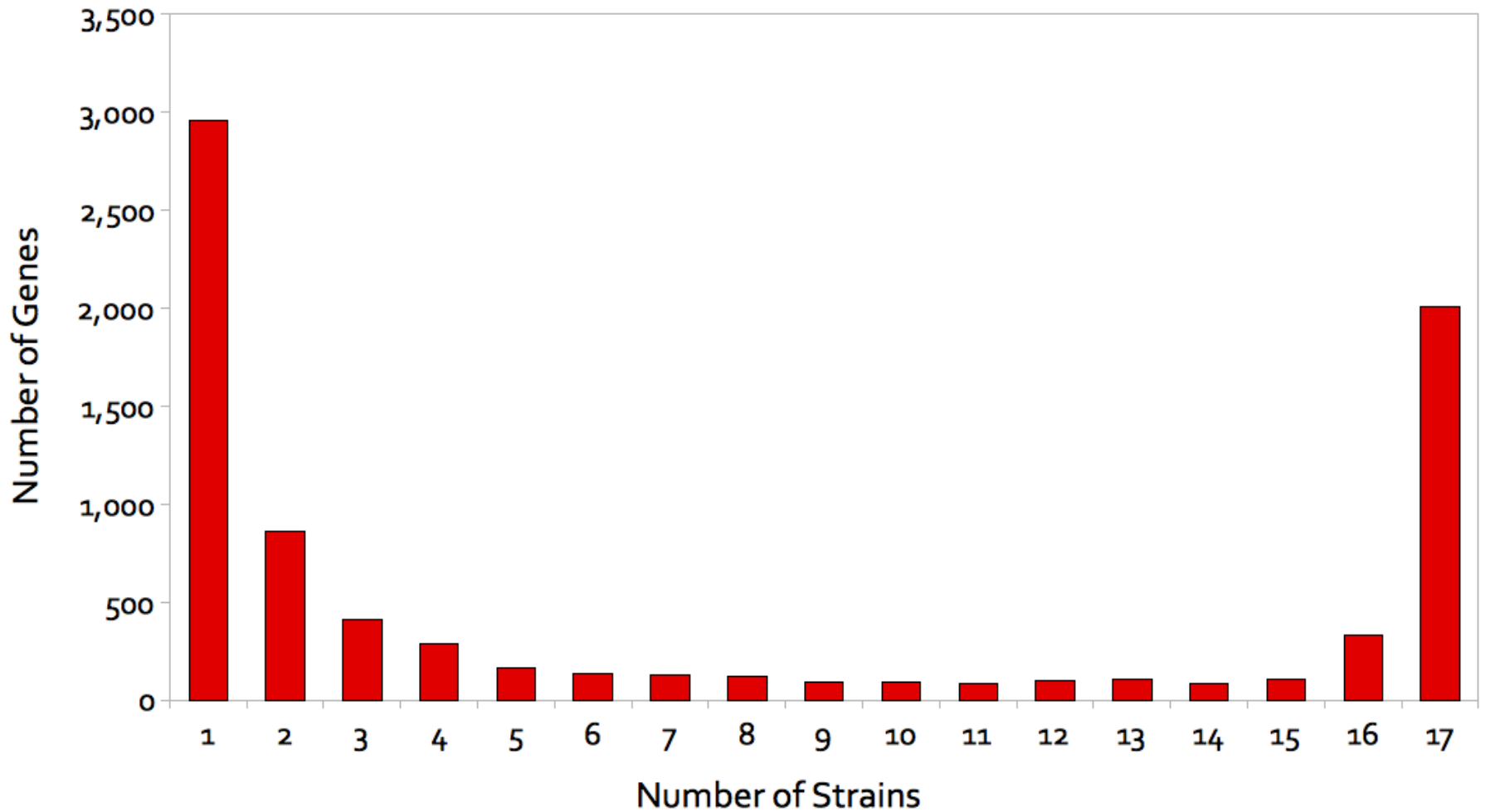
# Genes and Strains: *Streptococcus pneumoniae*



*Streptococcus pneumoniae* has 3,807 proteins

Pool Size	Probability of choosing a gene
1,745	0.027397
441	0.18988
229	0.54277
236	0.82951
1,156	0.99471

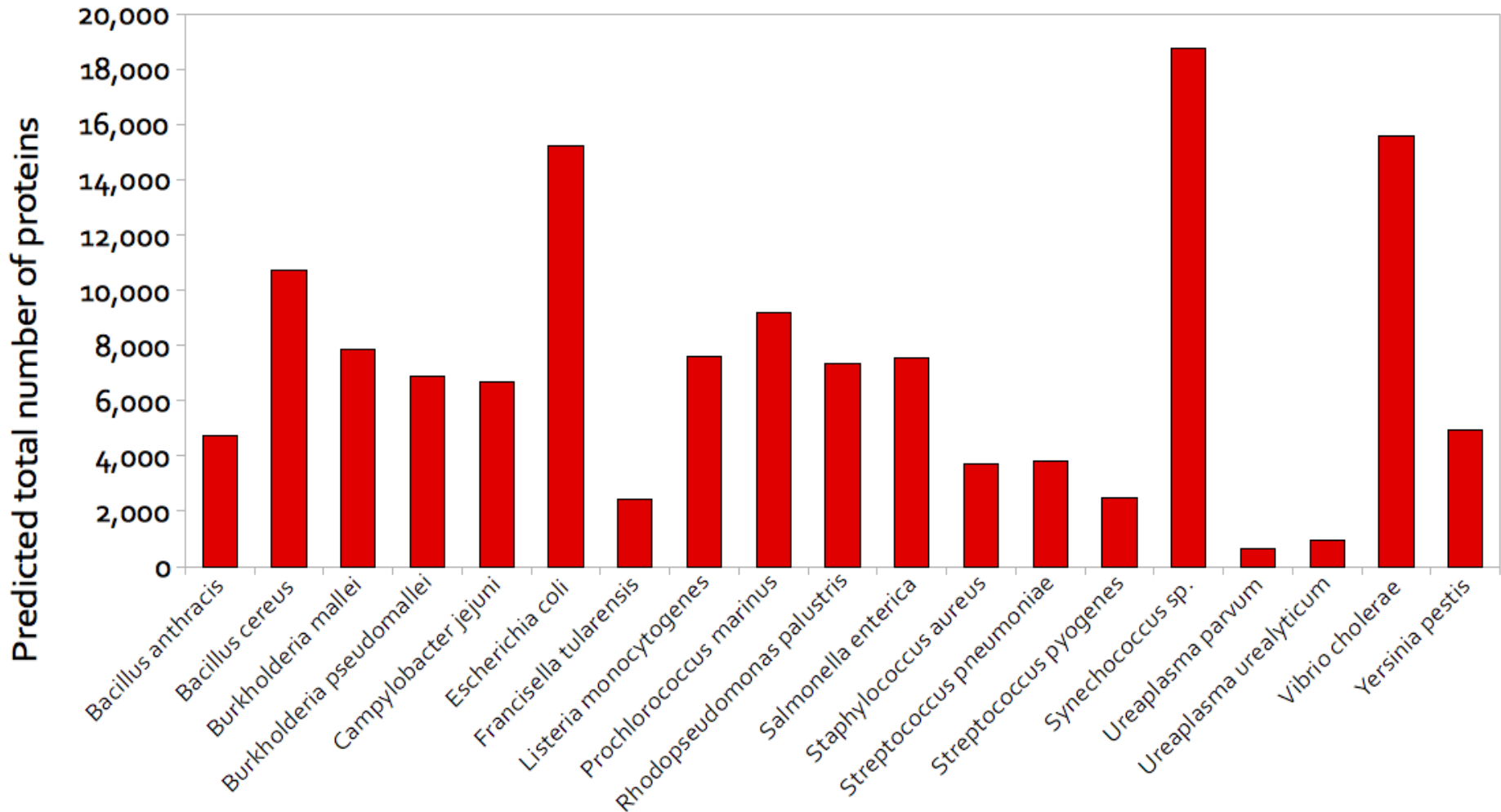
# Genes and Strains: *Escherichia coli*



# *Escherichia coli* has 15,242 proteins

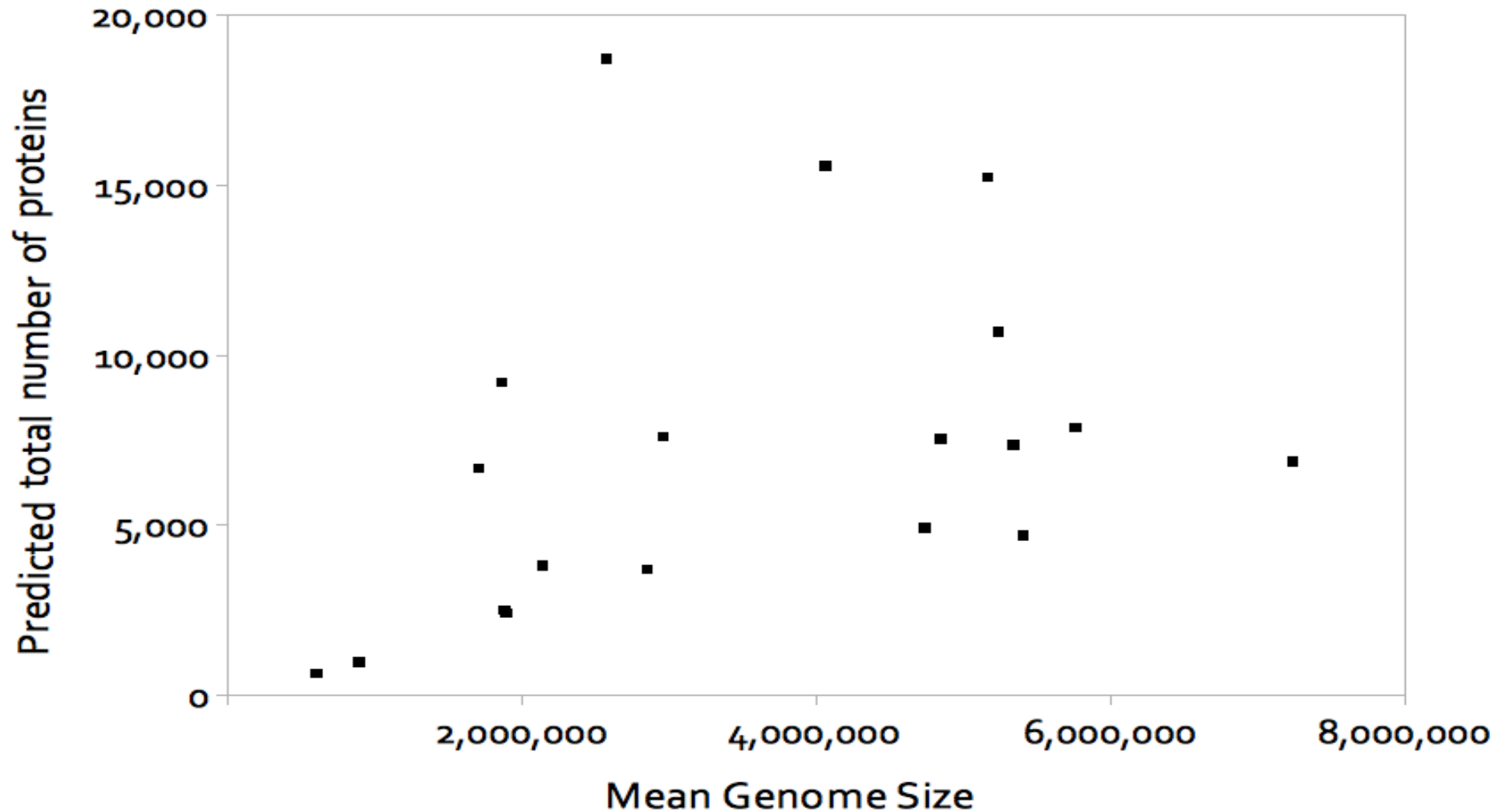
Pool Size	Probability of choosing a gene
10,372	0.022029
1,463	0.16253
575	0.44274
452	0.74235
2,380	1

# How many genes are there in different strains?

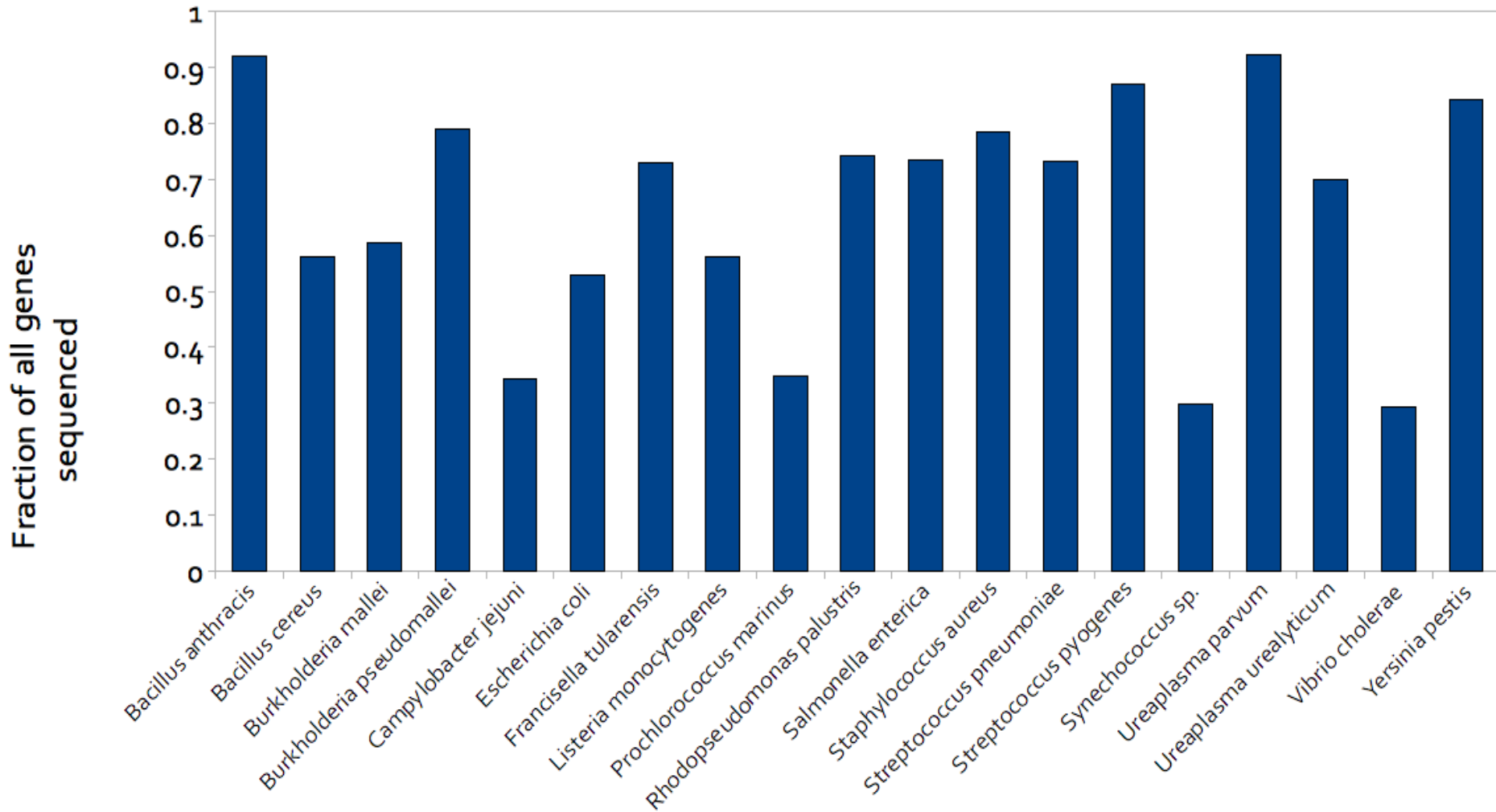




# Is pool total number of proteins dependent on genome size?



# How far along are we with sequencing?



**BACK!**