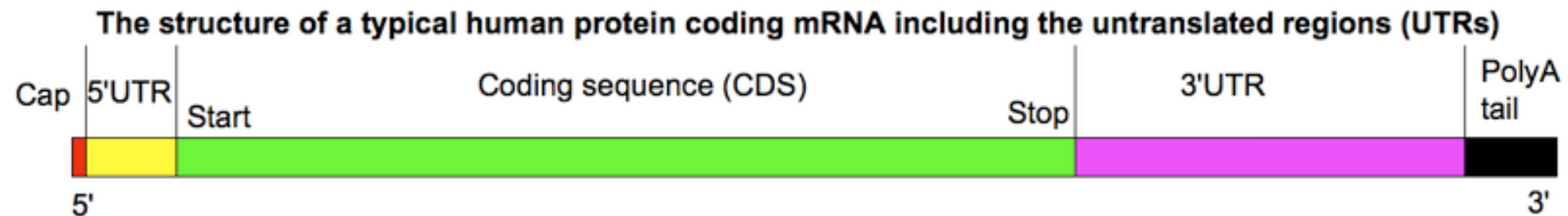


RNASeq Introduction

Malay (malay@uab.edu)

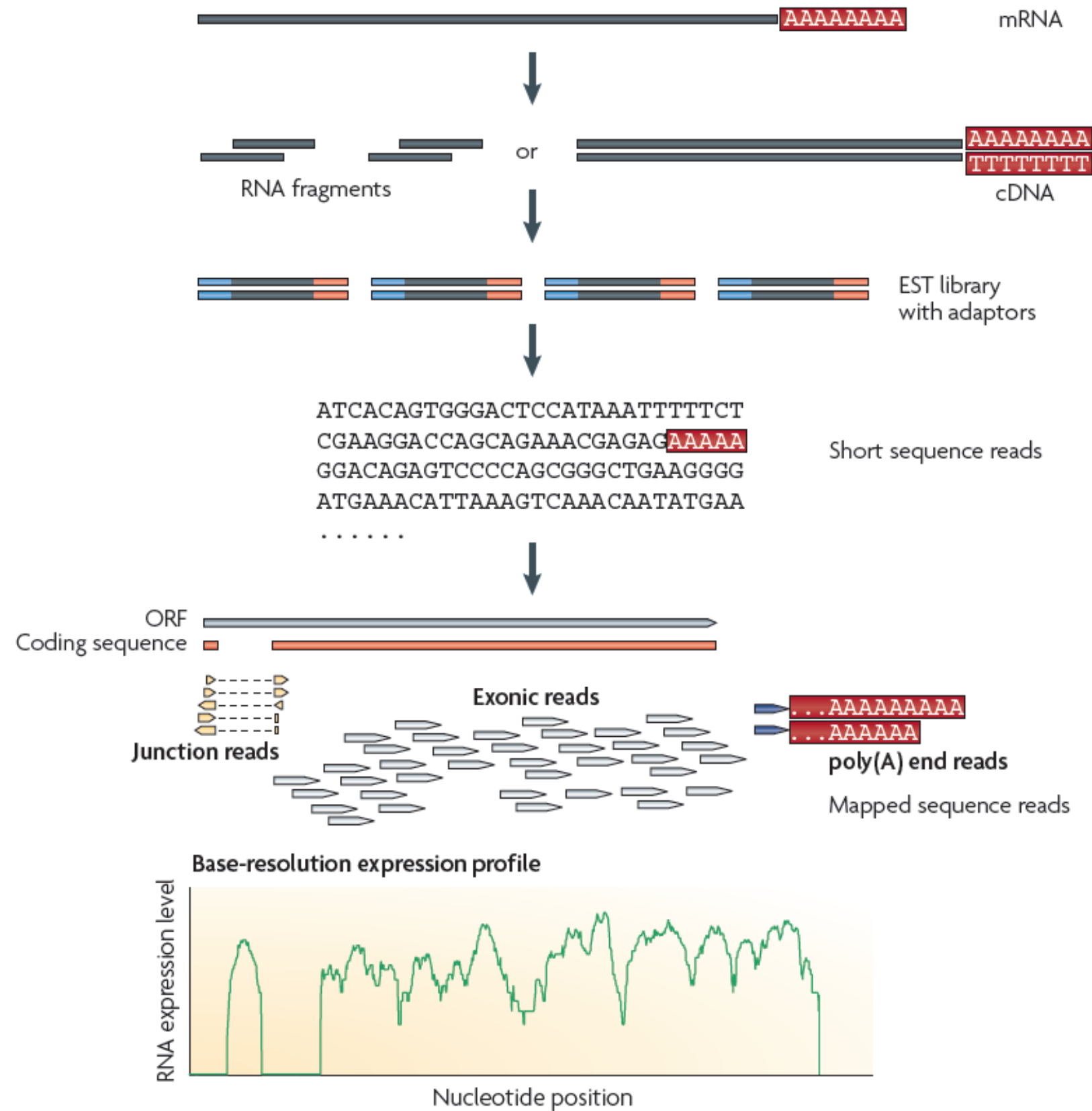
Structure of mRNA



Other RNAs:

Type	Size	Function
▪ microRNA (miRNA)	21-23 nt	regulation of gene expression
▪ small interfering RNA (siRNA)	19-23 nt	antiviral mechanisms
▪ piwi-interacting RNA (piRNA)	26-31 nt	interaction with piwi proteins/spermatogenesis
▪ small nuclear RNA (snRNA)	100-300 nt	RNA splicing
▪ small nucleolar RNA (snoRNA)	-	modification of other RNAs

RNASeq overview



Various applications of RNASeq

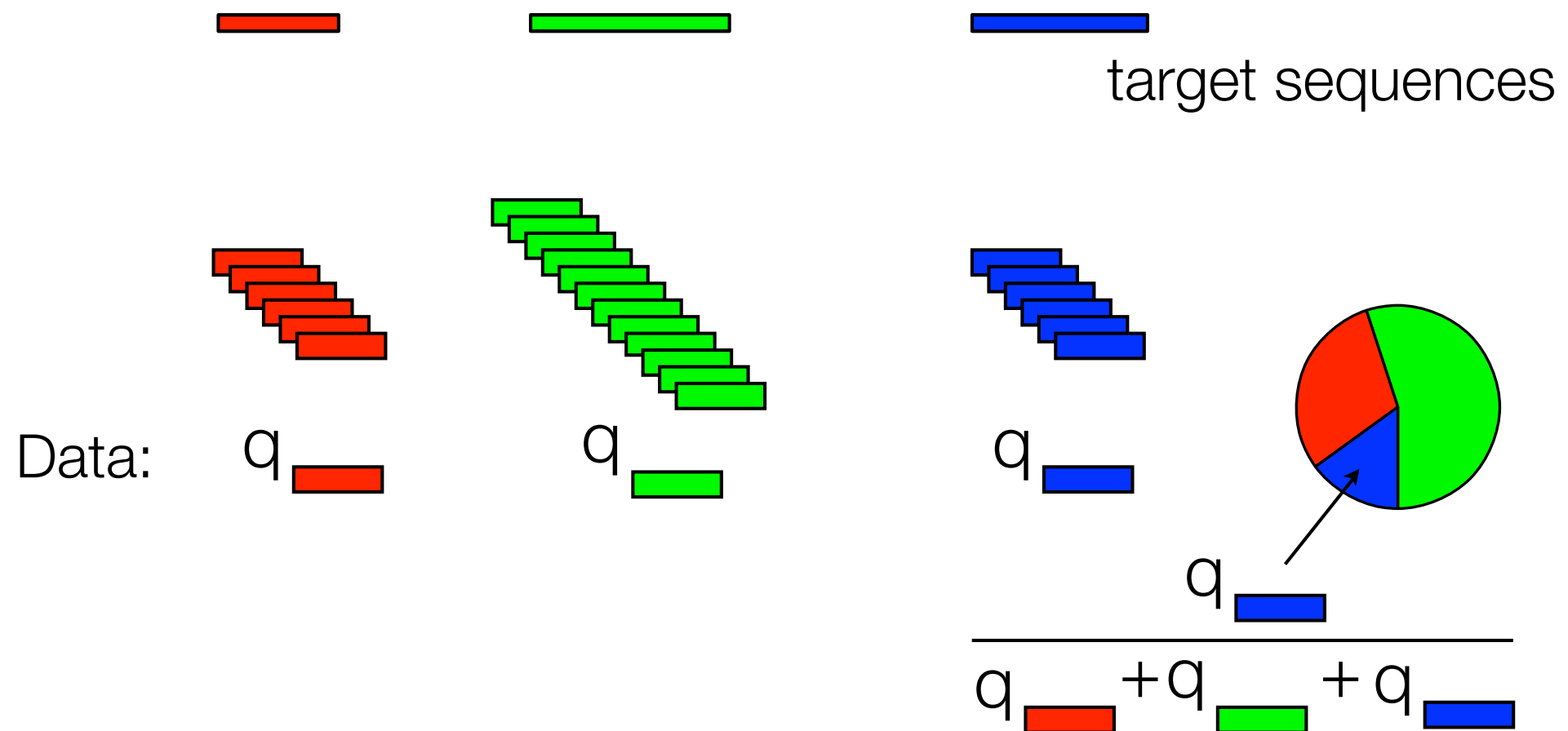
Differential expression

Novel transcript detection

Fusion transcript detection

Mutation detection (not common)

The fragment assignment problem



Example count tables:
<http://bowtie-bio.sourceforge.net/recount/>

Normalization methods in RNASeq

Upper Quartile (UQ): Very similar in principle to TC, the total counts are replaced by the upper quartile of counts different from 0 in the computation of the normalization factors.

DESeq: This normalization method is included in the DESeq

Trimmed Mean of M-values (TMM): This normalization method is implemented in the edgeR

Reads Per Kilobase per Million mapped reads (RPKM): This approach was initially introduced to facilitate comparisons between genes within a sample and combines between- and within-sample normalization.

FPKM: Same as RPKM, but for fragments.

TPM: comes from RSEM

$$RPKM = \frac{Count}{\frac{N}{10^6} \times \frac{length}{10^3}}$$

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

$$\text{TPM}_i = \left(\frac{\text{FPKM}_i}{\sum_j \text{FPKM}_j} \right) \cdot 10^6$$

Upper quartile normalization

Sample_ID gene raw_1Kb_read_count

1 Act7 1000

1 GapDH 2000

1 Sec4 500

...

1 Bglob1 10000000

TOTAL WITH Bglob1: 12,000,000

TOTAL WITHOUT: 2,000,000

2 Act7 500

2 GapDH 1000

2 Sec4 250

...

2 Bglob1 10

TOTAL WITH Bglob1: 1,000,010

TOTAL WITHOUT: 1,000,000

$$FPKM_{Act7} = \frac{1000}{\frac{1000}{10^3} \times \frac{12^6}{10^6}} = 83.3$$

$$FPKM_{Act7} = \frac{500}{\frac{1000}{10^3} \times \frac{1,000,010}{10^6}} = 499.995$$

$$FPKM_{Act7} = \frac{1000}{\frac{1000}{10^3} \times \frac{2^6}{10^6}} = 500$$

$$FPKM_{Act7} = \frac{500}{\frac{1000}{10^3} \times \frac{1^6}{10^6}} = 500$$

DESEQ normalization

$$\hat{s}_j = \operatorname{median}_i \frac{k_{ij}}{\left(\prod_{v=1}^m k_{iv} \right)^{1/m}}.$$

Geometric mean across the samples



TMM normalization (EDGER)

$$M_g = \log_2 \frac{Y_{gk} / N_k}{Y_{gk'} / N_{k'}}$$

and absolute expression levels:

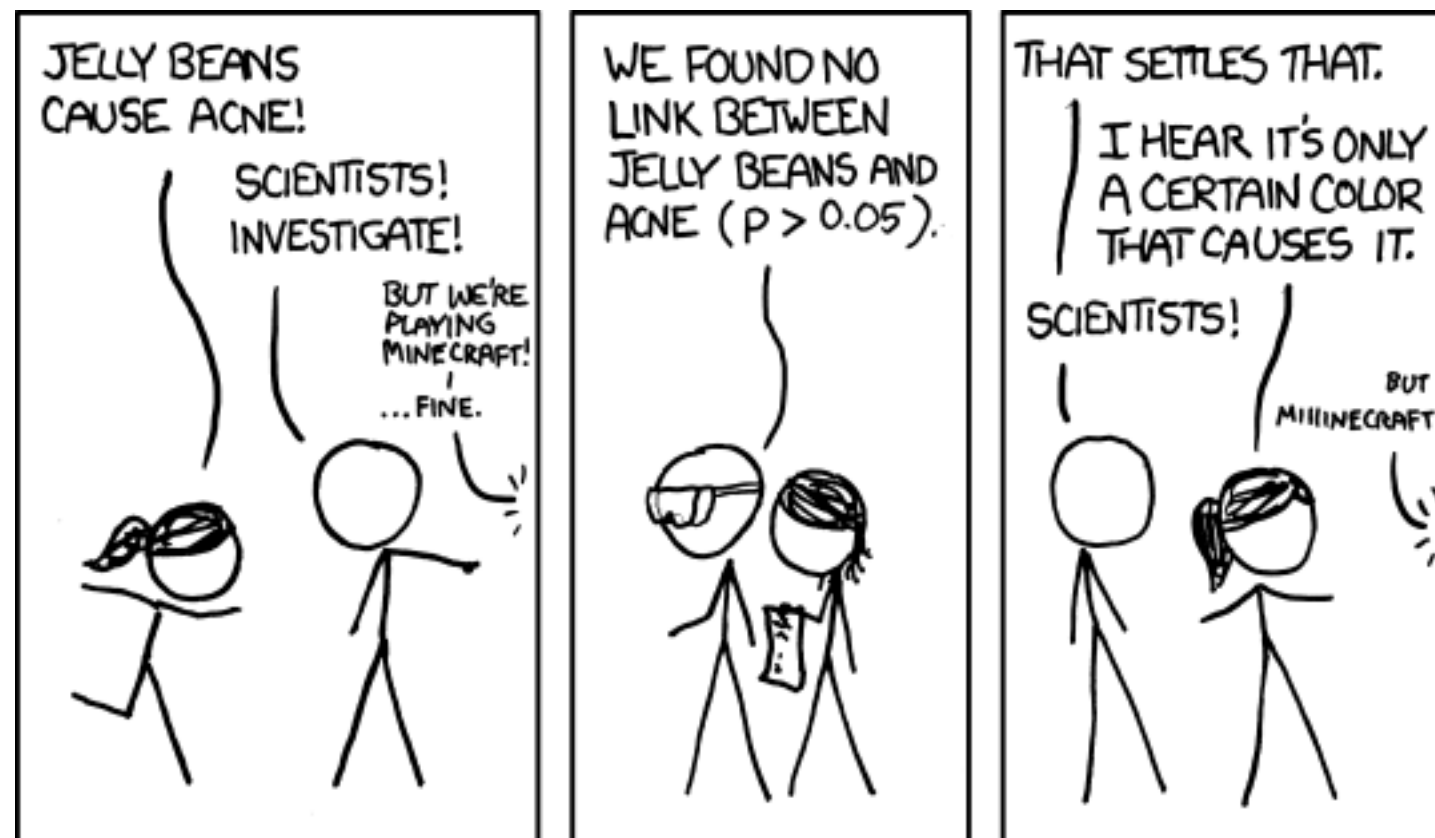
$$A_g = \frac{1}{2} \log_2 \left(Y_{gk} / N_k \bullet Y_{gk'} / N_{k'} \right) \text{ for } Y_{g\bullet} \neq 0$$

Remove top and bottom 30%

$$\log_2(TMM_k^{(r)}) = \frac{\sum_{g \in G^*} w_{gk}^r M_{gk}^r}{\sum_{g \in G^*} w_{gk}^r} \text{ where } M_{gk}^r = \frac{\log_2 \left(\frac{Y_{gk}}{N_k} \right)}{\log_2 \left(\frac{Y_{gr}}{N_r} \right)} \text{ and } w_{gk}^r = \frac{N_k - Y_{gk}}{N_k Y_{gk}} + \frac{N_r - Y_{gr}}{N_r Y_{gr}};$$

$Y_{gk}, Y_{gr} > 0.$

False Discovery Rate and q-value



WE FOUND NO
LINK BETWEEN
BLUE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
TEAL JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PURPLE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
BROWN JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PINK JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND A
LINK BETWEEN
GREEN JELLY
BEANS AND ACNE
($P < 0.05$).



WE FOUND NO
LINK BETWEEN
MAUVE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
MAGENTA JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
YELLOW JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
SALMON JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
RED JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
TURQUOISE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PEACH JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
ORANGE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
BLUE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
TEAL JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PURPLE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
BROWN JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PINK JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND A
LINK BETWEEN
GREEN JELLY
BEANS AND ACNE
($P < 0.05$).



WE FOUND NO
LINK BETWEEN
MAUVE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
MAGENTA JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
YELLOW JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
SALMON JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
RED JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
TURQUOISE JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
PEACH JELLY
BEANS AND ACNE
($P > 0.05$).



WE FOUND NO
LINK BETWEEN
ORANGE JELLY
BEANS AND ACNE
($P > 0.05$).



News

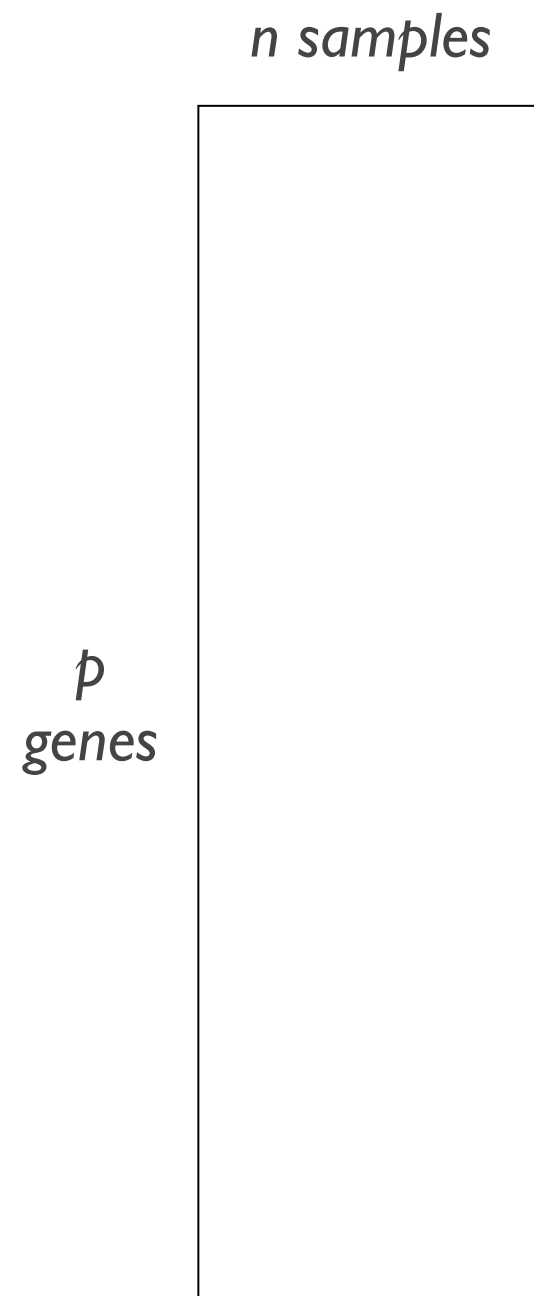
GREEN JELLY BEANS LINKED TO ACNE!

95% CONFIDENCE

ONLY 5% CHANCE OF COINCIDENCE!



SCIENTISTS...



We're doing p
simultaneous tests!

$H_1, H_2, H_3, \dots, H_p$

Bonferonni Correction

Storey's q-value

