

# **ERROR CORRECTION**

**30 SEPT 15**

# **ANNOUNCEMENTS**

# **FINAL PROJECT**

- Groups of 3-5. No more than 1 grad per group.
- Can use data from published manuscript, downloaded from ENA, your own data.
- I approve topic (Oct14)
- Presentations last week of class
- Written paper due last day.

# **FINAL PROJECT EXAMPLES**

Genome

Transcriptome

Expression

Comparative Genomics

# **EXAM REVIEW**

Evolution is:

The 4 mechanisms are:

# ERROR CORRECTION



# ERROR CORRECTION

TATACAATTCTGTTTATGAAAACCTCTAAAGCAACATACTTACCAACATCCCTGGCATCGAATAACCGATTCCTATTTAGCAATTCTGTTTATGAAAACCTCTAAAGCAACATACTTACCAACATCCCTGGCATCGAATAACCGATTCCTATTTAGCA

ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGCCTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGACATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA

Consensus= ACTGTCATTGGACTA

# ERROR CORRECTION

TATACAATTTCGTTTTA TGA AAA ACTCCTAA AAGCAA ACATA TTACCA ACGA ATCCCTGGCA TACGAA ATA ACCGATTCCTATTTA GCAATTGCTCCCTA TTITTA TACAATTTCGTTTT TATGAAA ACTCCTAA AAGCAA ACATA TTACCGAA ACGA ATCCCTGGCA TACGAA ATA ACCGATTCCTATTTA AGCAATTGCTC

ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGCCTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGTCATTGGACTA  
ACTGACATTGGACTA  
ACTGACATTGGACTA  
ACTGTCATTGGACTA

Consensus= ACTG{A,T}ATTGGACTA

# ERROR CORRECTION

TATACAATTCTGTTTATGAAAACCTCTAAAGCAAACATATTACCAAGAACCTTGCAACGGAAATAACCGATTCCTATTTAAGCATTCGCTCTATTTATGAAAAACTCTAAAGCAAACATTTACCAAGAACCTTGCAACGGAAATAACCGATTCCTATTTAAGCAATGC

ACTGTCATT CGGACTA  
**GCTGATAAC CGGACTA**  
ACTGACATT CGGACTA  
ACTGTCATT CGGACTA

Consensus= ACTGTCATT CGGACTA

# **ERROR CORRECTION**

Hamming Distance:

[http://en.wikipedia.org/wiki/Hamming\\_distance](http://en.wikipedia.org/wiki/Hamming_distance)

# **ERROR CORRECTION**

3 different strategies

# **ERROR CORRECTION**

Kmer-spectra based

# **ERROR CORRECTION**

Suffix tree based

# **ERROR CORRECTION**

MSA based

# **ERROR CORRECTION**

Evaluation of Correction