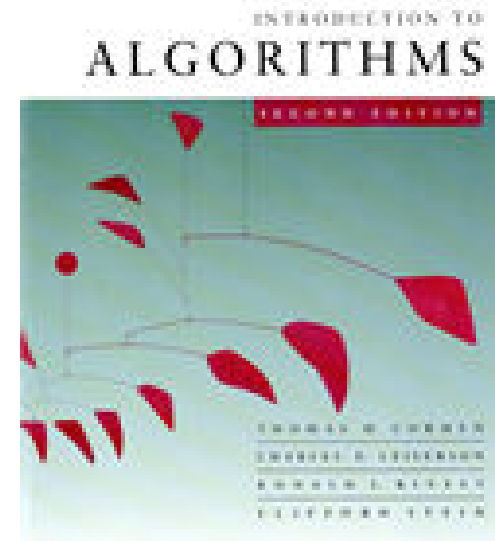




Design and Analysis of Algorithms

Introduction



Introduction to Algorithms (2nd)

by Cormen, Leiserson, Rivest and Stein
(CLRS) MIT Press

- Textbook:
 - ***Introduction to Algorithms*, 2nd edition, by T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein. The MIT Press, 2001. (2002年5月高等教育出版社出版影印版)**
- Grading Policy:
 - **Homework + Projects : 40%**
 - **Final Exam: 60%**



Analysis of algorithms

- The theoretical study of computer-program performance and resource usage.
- What's more important than performance?
 - modularity
 - correctness
 - maintainability
 - functionality
 - robustness
 - user-friendliness
 - programmer time
 - simplicity
 - extensibility
 - reliability



Why study algorithms and performance?

- Algorithms help us to understand **scalability**.
- Performance often draws the line between what is feasible and what is impossible.
- Algorithmic mathematics provides a **language** for talking about program behavior.
- Performance is the **currency** of computing.
- The lessons of program performance generalize to other computing resources.
- Speed is fun!

Two Distinct Choices



I want the
more expensive
computer

rich man

I want the better
algorithm



smart mathematician

Algorithm matters

- **Sort 10 million integers on**
 - 1 GHZ computer (1000 million instruction per second) using $2n^2$ algorithm
 - 100 MHz computer (100 million instruction per second) using $50n \log n$ algorithm

- **Supercomputer**

$$\frac{2 \cdot (10^7)^2 \text{ instructions}}{10^9 \text{ instructions/second}} = 200000 \text{ seconds} \approx 55 \text{ hours,}$$

- **Personal computer**

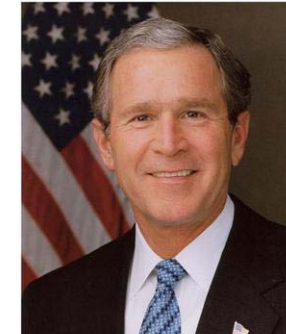
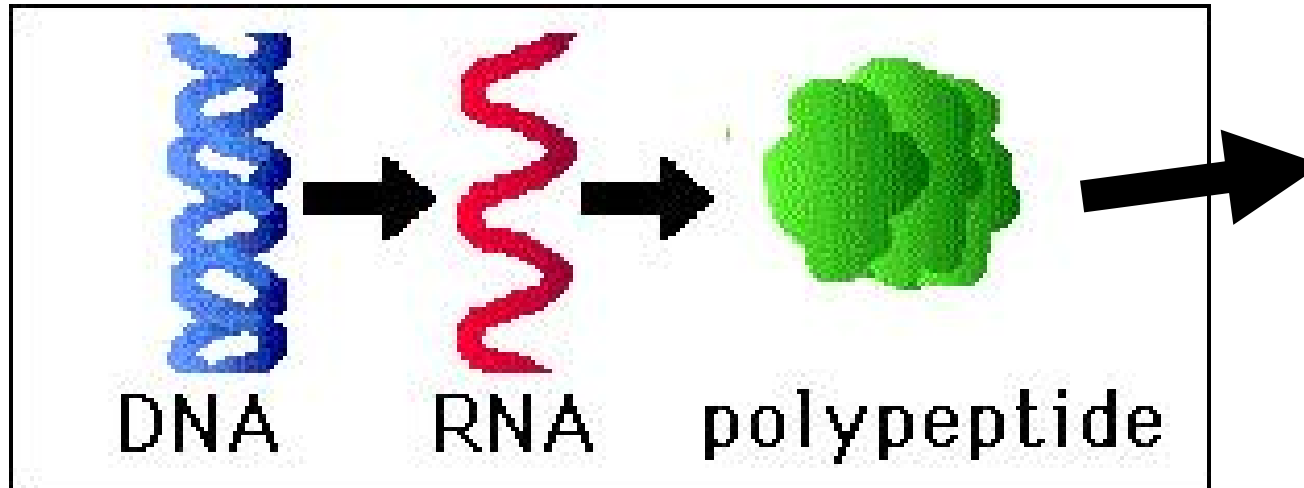
$$\frac{50 \cdot 10^7 \lg 10^7 \text{ instructions}}{10^8 \text{ instructions/second}} \approx 105 \text{ seconds.}$$



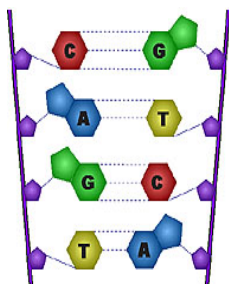
Some Applications

- **Practical App. of Algorithms are ubiquitous and Including**
 - data retrieval
 - network routing
 - games
 - human genome project
 - Internet algorithms
 - electronic commerce
 - manufacturing and other commercial settings
 - many concrete problems...

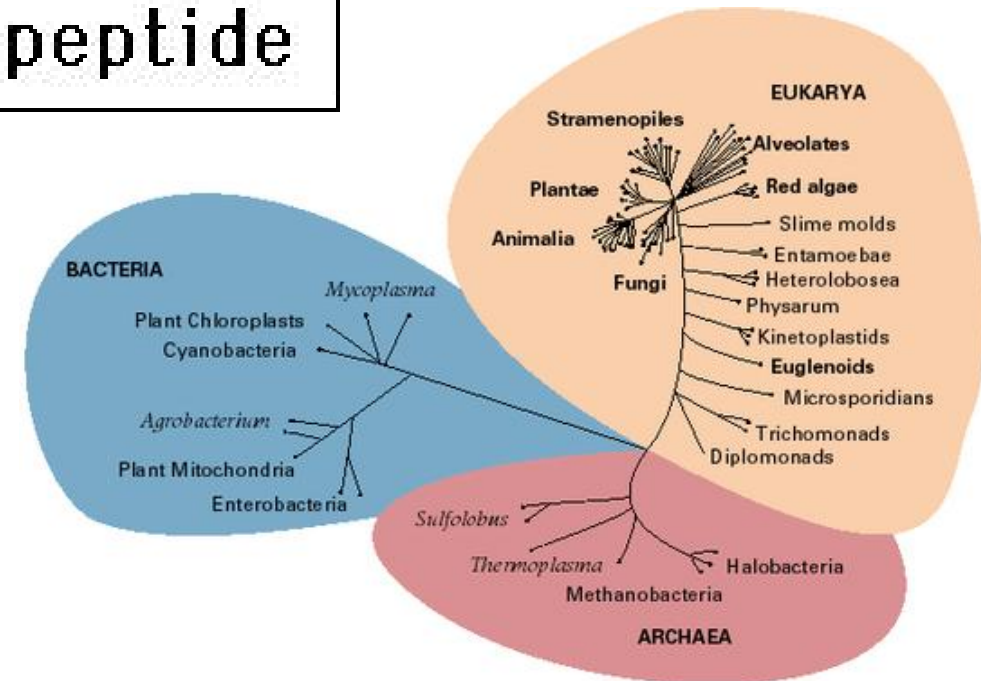
Biology in One Slide – Twentieth Century



...and today



...ACGTGACTGAGGACCGTG
CGACTGAGACTGACTGGGT
CTAGCTAGACTACGTTTTA
TATATATATACGTCGTCGT
ACTGATGACTAGATTACAG
ACTGATTTAGATACTGAC
TGATTTTAAAAAATATT...



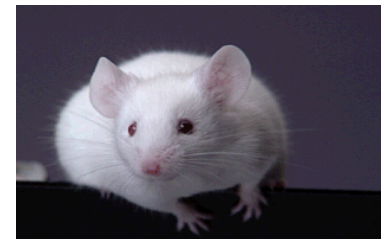
Complete DNA Sequences



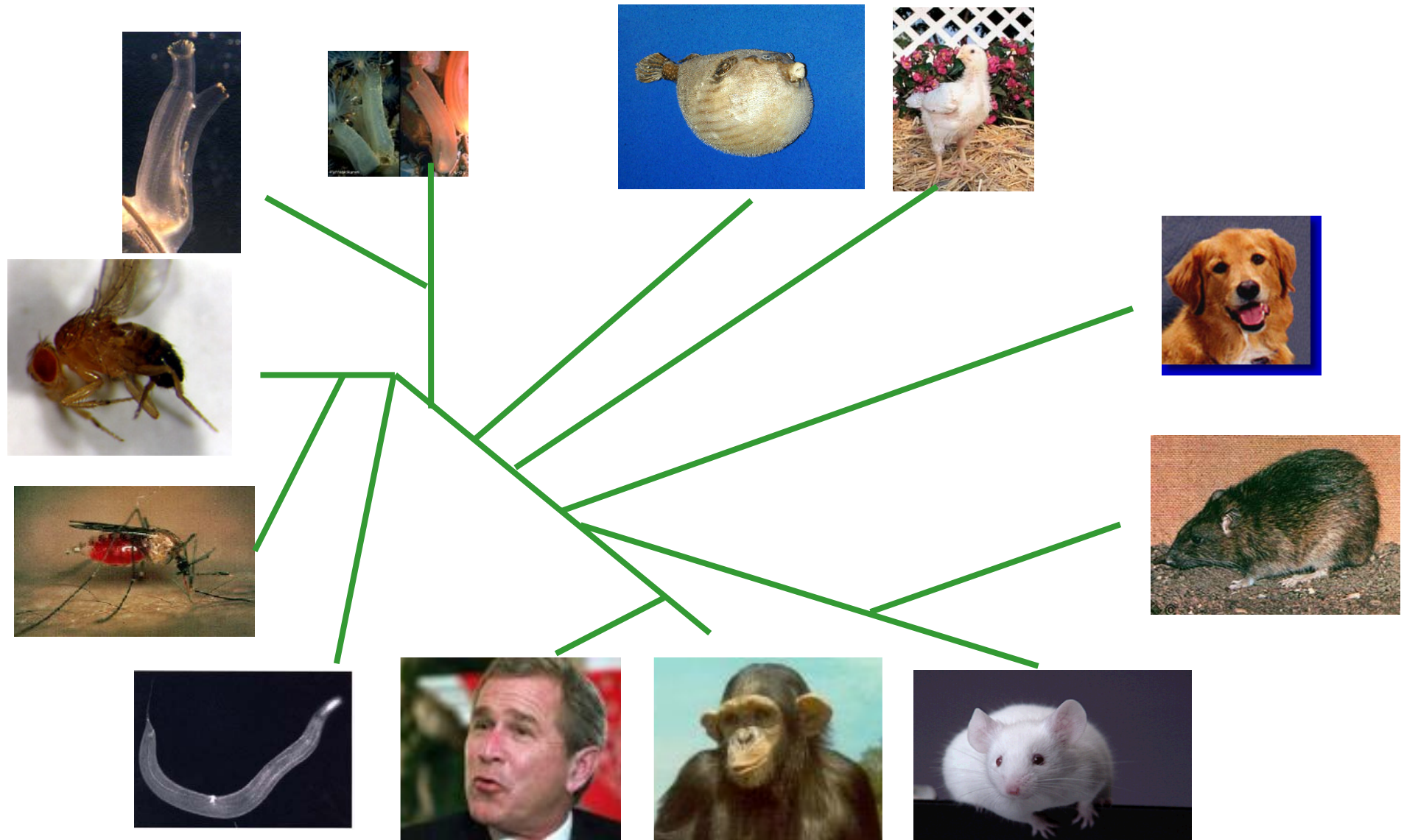
nearly 200 complete
genomes have been
sequenced



Huo Hongwei



Evolution



Huo Hongwei

TTATATTGAATTTTCAAAAATTCTTACTTTTTTTTTTGGATGGACGCAAAGAAGTTTAATAATCATATTACATGGCATTACACCA
TATACATATCCATATCTAATCTTACTTATATGTTGTGGAAATGTAAAGAGCCCATTAATCTTAGCCTAAAAAACCTTCTCTTTG
GAACCTTTAGTAATAAGCTTAACCTGCTCATTGCTATATTGAAGTACGGATTAGAAGCCGCGAGCGGGCGACAGCCCTCCGACGG
AAGACTCTCTCTCGTGGTCTCTCTTCAACCGGTGCGCTTCTGAAACGCAGATGTGCTCGCGCCGCACTGCTCCGAACAATA
AAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAATDGGCAGTAACCTGGCCCCACAAACCTTCAAATTAACGAATC
AAATTAACAACCATAGGATGATAATGCGATTAGTTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGAT
CTATTAACAGATATATAAATGGAAAAGCTGCATAACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTACTTCTTATTC
AAATGTCATAAAAAGTATCAACAAAAAATTGTTAATATACCTCTATACTTTAAOGTCAAGGAGAAAAAACTATAATGACTAAATCT
CATTGAGAAGAAGTGATTGTACCTGAGTTCAATTCTAGCGCAAAGGAATTACCAAGACCATTGGCCGAAAAGTGCCCGAGCATAA
TTAAGAAATTTATAAGCGCTTATGATGCTAAACCGGATTTTGTGTGCTAGATCGCTGGTAGAGTCAATCTAATTGGTGAAACATAT
TGATTATTGTGACTTCTCGGTTTTACCTTTAGCTATTGATTTTGATATGCTTTGCGCCGTCAAAGTTTTGAACGAGAAAAATCCA
TCCATTACCTTAATAAATGCTGATCCCAAATTTGCTCAAAGGAAGTTGATTTGCGGTTGGAOGGTTCTTATGTCACAATTGATC
CTTCTGTGTGCGGACTGGTCTAATTACTTTAAATGTGGTCTCCATGTTGCTCACTCTTTTCTAAAGAACTTGCACCGGAAAGGTT
TGCCAGTGCTCCTCTGGCCGGGCTGCAAGTCTTCTGTGAGGGTGATGTACCAACTGGCAGTGGAATTGTCTTCTTGGCCCGCATTC
ATTTGTGCGGTTGCTTTAGCTGTTGTTAAAGCGAATATGGGCCCCGGTTATCATATGTCCAAGCAAAATTTAATGCGTATTACGG
TCGTTGCAGAACATTATGTTGGTGTTAAACAATGGCGGTATGGATCAGGCTGCTCTGTTTGGCGGTGAGGAAGATCATGCTCTATA
CGTTGAGTTCAAACCGCAGTTGAAGGCTACTCCGTTTAAATTTCCGCAATTAATAAACCATGAAATTAGCTTTGTTATTGCGAAC
ACCTTGTGTATCTAAACAAGTTTGAAACCGCCCCAACCAACTATAATTTAAGAGTGGTAGAAGTCACTACAGCTGCAAATGTTT
TAGCTGCCACGTACGGTGTTGTTTTACTTTCTGGAAAAGAAGGATCGAGCACGAATAAAGGTAATCTAAGAGATTTTATGAACGT
TTATTATGCCAGATATCACAAACATTTCCACACCCTGGAACGGCGATATTGAATCOGGCATCGAACGGTTAACAAAGATGCTAGTA
CTAGTTGAAGAGTCTCTCGCCAATAAGAAAAGGGCTTTAGTGTTGACGATGTCCGCACAATCCTTGAATTGTTCTCGCGAAGAAAT
TCACAAGAGACTACTTAACAACATCTCCAGTGAGATTTCAAGTCTTAAAGCTATATCAGAGGGCTAAGCATGTGTATTCTGAATC
TTTAAGAGTCTTGAAGGCTGTGAAATTAATGACTACAGCGAGCTTTACTGCGGACGAAGACTTTTTTCAAGCAATTTGGTGCTTG
ATGAACGAGTCTCAAGCTTCTTGCGATAAACTTTACGAATGTTCTTGTCCAGAGATTGACAAAATTTGTTCCATTGCTTTGTCAA
ATGGATCATATGGTTCCCGTTTGACCGGAGCTGGCTGGGGTGTTGTACTGTTCACTTGGTTCCAGGGGGCCCCAATGGCAACAT
AGAAAAGGTAAAAGAAGCCCTTGCCAATGAGTTCTACAAGGTCAAGTACCTAAGATCACTGATGCTGAGCTAGAAAATGCTATC
ATCGTCTCTAAACCAGCATTGGGCGAGCTGTCTATATGAATTATAAGTATACTTCTTTTTTTTACTTTGTTTCAGAACAACCTCTCA
TTTTTTTTCTACTCATAACTTTAGCATCACAAAATACGCAATAATAACGAGTAGTAACACTTTTATAGTTCATACATGCTTCAACT
ACTTAATAAATGATTGTATGATAATGTTTTCAATGTAAGAGATTTGATTATOCACAACTTTAAACACAGGGACAAAATTTCTT
GATATGCTTTCAACCGCTGCGTTTTGGGATACTATTCTTGACATGATATGACTACCATTTTGTATTGTACGTGGGGCAGTTGAC
GTCTTATCATATGTCAAAGTCAATTTGCGAAGTTCTTGGCAAGTTGCCAACTGACGAGATGCAGTAAAAAGAGATTGCGGCTCTTGA
AACTTTTTGTCTTTTTTTTTTTTCCGGGACTCTACGAGAAACCTTTGTCTACTGATTAATTTTGTACTGAATTTGGACAATTCA
GATTTTAGTAGACAAGCGCGAGGAGGAAAAAGAAATGACAGAAAAATTCOGATGGACAAGAAGATAGGAAAAAAGCTTTCA
COGATTTCTAGACCGGAAAAAAGTCTGATGACATCAGAAATGAAAAATTTTCAAGTTAGACAAGGACAAAAATCAGGACAAATTTGT
AAAGATATAATAAACTATTTGATTGAGCGCAATTTGCCCTTTTCCATTTTCCATTAAATCTCTGTCTCTCTTACTTATATGAT
GATTAGGTATCATCTGTATAAAACTCTTTCTTAATTTCACTCTAAAGCATACCCATAGAGAAGATCTTTGGTTCCGAGACAT
TCTTACGCATAATAAGAATAGGAGGGAATAATGCCAGACAATCTATCATTACATTTAAGCGGCTCTTCAAAAAGATTGAACCTCTC
GCCAACTTATGGAATCTTCCAATGAGACCTTTGCGCCAAATAATGTGGATTTGGAAAAAGAGTATAAGTCATCTCAGAGTAATAT
AACTACCGAAGTTTATGAGGCATCGAGCTTTGAAGAAAAAGTAAGCTCAGAAAAOCTCAATACAGCTCATTCTGGAAGAAAAATC
TATTATGAATATGTGGTGGTTGACAAATCAATCTTGGGTGTTTCTATTCTGGATTCAATTTATGTACAACCAGGACTTGAAGCCCG
TCGAAAAAGAAAGGCGGGTTTTGGTCTTGGTACAATTAATGTTTACTTCTTGGCTTGGCTGAATGTTTCAATATCAACACTTGGCAAT

.TATIGAAATTTTCAAAAATTCTTACTTTTFTTTTGGATGGACGCAAAGAAGTTTAATAATCATATTACATGGCATTACCCACATATA
 'ATCCATATCTAATCTTAC**TTATA**TGTTGTTGGAAATGTAAAGAGCCCCATTATCTTAGCCTAAAAAACCTTCTCTTTGGAACTTTC
 'AATACGCTTAACTGCTCATTGCTATATTGAAGTA**CGG**ATTAGAAGCCGC**CGAG****CGGG**CGACAGCCCT**CCGA****CGGA**AAGACTCTCCTC
 'GCGTCCCTCGTCTTACACGGTTCGCGTTCTGAAACGCAGATGTGCCT**CGCG**CGCAGCTGCT**CGGA**ACAATAAAGATTCTACAATACT
 'TTTTATGGTTATGAAGAGGAAAAATTGGCAGTAACTTGG**CCC****CA**CAAAACCTTCAAATTAACGCTCAAATTAACAACCATAGGATG
 'ATGCGATTAGTTTCTTAGCCTTATTT**CTGGGG**TAAATTAATCAGCGAGTATGATTTTTGATCTTTAAACAGATA**TATA**AATGGAA
 'CTGCATAACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTATCTTTTATTCAAATGTCTTAAAGTATCAACAAAAAT
 'TAATAACCTCTATACCTTTAACGTCAAGGAGAAAAAACTATA**ATGACTAAAT**CTT**ATTCAGAAGAA**TGATTGTACCTGAGTTCAA
 'TAGCGCAAAGGAATTACCAAGACCATTGGCCGAAAAGTGCCCGAGCATAATTAAGAACTTTATAAGCTTTATGATGCTAAACCGG
 'TTGTTGCTAGATCGCCTGGTAGAGTCAATCTAATTGGTGAACATATTGATTATTGTGAGCTCGGTTTACCTTTAGCTATTGAT
 'GATATGCTTTTCCCGCTCAAACTTTTGAACGAGTCTATCCATCCATTACCTTAATAAATGCTATCCCACTTTTGCTCAAAGGAA
 'CGTATGATCCTTCTGTGTCCGACTGGTCTA
 'AACAGGTTTGCCAGTGCTCCTCTGGCCGGGCTG
 'ATTCATTTGTGCGTTGCTTTAGCTGTTGTTAA
 'CGTTGCAGAACATTATGTTGGTGTAAACAATG
 'GAGTTCAAACCGCAGTTGAAGGCTACTCGGTTT
 'GTATCTAACAAGTTTGAAACCGCCCCAACCAA
 'ACGGTGTGTTTCTTCTGGAAGAAGGAT
 'CACACATTTCCACACCTTGAAACGGCGATATT
 'CTAGTACTAGTTGAAGAGTCTCTTCCCAATAAGAAACAGGGCTTTAGTGTGACGATGT
 'ATTCACAAGAGACTACTTAACAAGCTCTCCAGTGAGATTTCAGTCTTAAAGCTATATCAGAGGCTAAGCATGTGTATTCTGAAT
 'TAAGAGTCTTGAAGGCTGTGAAATTCTGACTACAGCGAGCTTTACTGCGAOGAAGACTTTTAAAGCAATTTGGTGCCTTGATG
 'GAGTCTCAAGCTTCTTGCGATAAACTTACGAATGTTCTTGTCAGAGATTGACAAAATTTGTCCATTGCTTTGTCAAATGGATC
 'TGTTTCCCGTTTGACCGGAGCTGGCTGGTGGTTGTAAGTCTGTTCACTTGGTTCCAGGGGGGCTTAAATGGCAACATAGAAAAGGTAA
 'AAGCCCTTGCCAATGAGTTCTACAAGGTCTGTACCCCTAAGATCACTGATGCTGAGCTAGCAATGCTATCATCGTCTCTAAACCA
 'TTGGGCAGCTGTCTATATGAATTATAAGTATCTTCTTTTACTTTGTTTCAAGAACATCTCTCATTTTTTTCTACTCATAACT
 'GCATCACAAAATACGCAATAATAACGAGTAGTCTCACTTTTATAGTTTATACATGCTTCTACTACTTAATAAATGATTGTATGATA
 'TTTTCAATGTAAGAGATTTOGATTATCCACAAATTTAAACACAGGGACAAAATTTCTATATGCTTTCAACCGCTGCGTTTGG
 'CCTATTCTTGACATGATATGACTACCATTTTGTTTGTGACGTGGGGCAGTTGACGTTTATCATATGTCAAAGTCATTGCGAAG
 'TTGGCAAGTTGCCAACTGACGAGATGCAGTAAAAAGCTATTGCCGTCTTGAAACTTTGTGCTTTTCTTCTCCGGGGACTCTAC
 'AAC**CCCTTTGT**CTTACTGATTAA**TTTTGTACT**GAAATT**TTACAAT**TCAGATTTTACAGACAAGCGCGAGGAGGAAAAGAAATGACA
 'AAATTCOGATGGACAAGAAGATAGGAAAAAAGAGCTACACCGATTCTCTTCCGGAAAAAAGTCTGATGACATCAGAATGA
 'ATTTTCAAGTTAGA**CAAGGAC**AAAATCAGGACAAATTGTATGATATAATAAATATTGATTTCAGCGCCAATTTGCCCTTTTCCA
 'TCCATTAAATCTCTGTCTCTCTTACTTATATGATGATTAGTATCATCTG**TATA**AAACTCCTTTCTTAATTTCACTCTAAAGCAT
 'CCATAGAGAAGATCTTTCCGTTTGAAGACATTCCTACGCATTAAGAATAGGAGGGAATA**ATGCCAGACAATCTATCATTACATT**
 'GCGGCTCTTCAAAAAGATTGAACTCTCGCCAACCTTATGGAATCTTCAATGAGACCTTTGCGCCAAATAATGTGGATTGGAAGAA
 'TATAAGTCATCTCAGAGTAATATACTACCGAAGTTTATGAGGCTTGGAGCTTTGAAGAAAAAGTAAGCTCAGAAAAACCTCAATA
 'CTCATTCTGGAAGAAAATCTATTATGAATATGTGGTGGTTGACAAATCAATCTTGGGTGTTTCTATTCTGGATTCATTTATGTACA
 'AGGACTTGAAGCCCGTCGAAAAAGAAAGGCGGGTTTGGTCTGGTACAATTATTGTTACTTCTGGCTTGGTGAATGTTTCAATATC
 'ACTTGGCAAAATTGCAGCTACAGGTCTACAACCTGGGTCTAAATTGGTGGCAGTGTGGATAACAATTTGGATTGGGTACGGTTTCGT
 'TGCTTTTGTGTTGTTTGGCCTCTAGAGTTGGATCTGCTTATCATTTGTCTATCCCTATATCATCTAGAGCATCATTCGGTATTTCT



.TATTGAATTTTCAAAAATTCCTTACTTTTFTTTTGGATGGACGCAAAGAAGTTTAATAATCATATTACATGGCATTACCCACATATA
 ATCCATATCTAATCTTACTTATAATGTTGTGGAAATGTAAAGAGCCCCATTATCTTAGCCTAAAAAACCTTCTCTTTGGAACTTTC
 AATACGCTTAACTGCTCATGTCTATATTGAAGTAAGGATTAGAAGCCGCGAGCGGGCGACAGCCCTCCGACGGGAAGACTCTCCTC
 GCGTCCCTCGTCTTCAACGGTGGCGTTCTGAAAACGAGATGTGCCTCGCGCGCACTGCTCCGAACAATAAAGATTCTACAATACT
 TTTTATGGTTATGAAGAGGAAAAATTGGCAGTAACTGGCCCACAAACCTTCAAATTAACGAATCAAAATTAACAAACCATAGGATG
 ATGCGATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTAAACAGATAATAAATGGAA
 CTGCATAACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTACTTCTTATTCAAATGTCATAAAAGTATCAACAAAAAT
 TAATATACCTCTATACCTTTAACGTCAGGAGAAAAAACTATAATGACTAAATCTCATTGAGAAGAAGTGATTGTACCTGAGTTCAA
 TAGCGCAAAGGAATTACCAAGACCATTGGCCGAAAAGTGCCGAGCATAATTAAGAAATTTATAAGCGCTTATGATGCTAAACCGG
 TTGTTGCTAGATCGCCTGGTAGAGTCAATCTAATTGGTGAACATATTGATTATTGTGACTTCTCGGTTTTACCTTTAGCTATTGAT
 GATATGCTTTGCGCCGTCAAAGTTTTGAACGAGAAAAATCCATCCATTACCTTAATAAATGCTGATCCCAAATTTGCTCAAAGGAA
 CGATTTGCCGTTGGACGGTTCTTATGTCACAATTGATCCTTCTGTGTGGACTGGTCTAATTACTTTAAATGTGGTCTCCATGTTG
 ACTCTTTTCTAAAGAACTTGCACCGGAAAGGTTTGCCAGTGCTCCTCTGGCGGGCTGCAAGTCTTCTGTGAGGGTGATGTACCA
 GGCAGTGGATTGTCTTCTCGGCCGCAATTCATTTGTGCGTTGCTTTAGCTGTTGTTAAAGCGAATATGGGCCCTGGTTATCATAT
 CAAGCAAAATTTAATGCGTATTACGGTGGTGCAGAACATTATGTTGGTGTAAACAAATGGCGGTATGGATCAGGCTGCCCTCTGTTT
 GTGAGGAAGATCATGCTCTATACGTTGAGTTCAAACCGCAGTTGAAGGCTACTCCGTTTAAATTTCCGCAATTAACAAACCATGAA
 AGCTTTGTTATTGCGAACACCCTTGTGTTATCTAACAAAGTTTGAAACCGCCCCAACCACTATAATTTAAGAGTGGTAGAAGTCAAC
 AGCTGCAATGTTTTAGCTGCCACGTACGGTGTGTTTTACTTTCTGGAAAAGAAGGATOGAGCACGAATAAAGGTAATCTAAGAG
 TCATGAACGTTTTATTATGCCAGATATCACAACATTTCCACACCCTGGAACGGGCGATAATTGAATCCGGCATOGAACGGTTAACAAAG
 CTAGTACTAGTTGAAGAGTCTCTCGCCAATAAGAAACAGGGCTTTAGTGTGACGATGTGACACAATCCTTGAATTGTTCTGCGA
 ATTCACAAGAGACTACTTAACAACATCTCCAGTGAGATTTCAAGTCTTAAAGCTATATCAGAGGGCTAAGCATGTGTATTCTGAAT
 TAAGAGTCTTGAAGGCTGTGAAATTAATGACTACAGCGAGCTTTACTGCCGACGAAGACTTTTTCAAGCAATTTGGTGCCTTGATG
 GAGTCTCAAGCTTCTTGCGATAAACTTTACGAATGTTCTTGTCCAGAGATTGACAAAATTTGTTCCATTGCTTTGTCAAATGGATC
 TGGTTCCCGTTTGACCGGAGCTGGCTGGGGTGGTTGTACTGTTCACTTGGTTCCAGGGGGCCCAATGGCAACATAGAAAAGGTAA
 AAGCCCTTGCCAATGAGTTCTACAAGGTCAAGTACCCTAAGATCACTGATGCTGAGCTAGAAAATGCTATCATCGTCTCTAAACCA
 TTGGGCGAGCTGTCTATATGAATTATAAGTATACTTCTTTTTTTTACTTTGTTTCAGAACAACTTCTCATTTTTTTCTACTCATAACT
 GCATCACAAAATACGCAATAATAACGAGTAGTAACACTTTTATAGTTTATACATGCTTCAACTACTTAATAAATGATTGTATGATA
 TTTTCAATGTAAGAGATTTGATTTATCCACAACTTTAAACACAGGGACAAAATTTCTTGATATGCTTTCAACCGCTGCGTTTTGG
 CCTATTCTTGACATGATATGACTACCATTTTGTATTGTACGTGGGGCAGTTGACGTCTTATCATATGTCAAAGTCATTTGCGAAG
 TTGGCAAGTTGCCAACTGACGAGATGCAGTAAAAAGAGATTGCCGTCTTGAAACTTTTGTCTTTTTTTTTTTCGGGGGACTCTAC
 AACCCCTTTGTCTTACTGATTAAATTTTGTACTGAAATTTGGACAATTCAGATTTTAGTAGACAAGCGCGAGGAGGAAAAGAAATGACA
 AAATTCOGATGGACAAGAAGATAGGAAAAAAGCTTTACCGATTTCTAGACCGGAAAAAGTCGTATGACATCAGAATGA
 ATTTTCAAGTTAGAAGGACAAAATCAGGACAAATTTGTAAAGATATAATAAATCTATTGATTTCAGCGCCAATTTGCCCTTTTCCA
 TCCATTAAATCTCTGTTCTCTCTTACTTATATGATGATTAGGTATCATCTGTATAAACTCCTTTCTTAATTTCACTCTAAAGCAT
 CCATAGAGAAGATCTTTGCGTTCCGAAGACATTCCTACGCATAATAAGAATAGGAGGGAATAATGCCAGACAATCTATCATTACATT
 GCGGCTCTTCAAAAAGATTGAACTCTCGCCAACTTATGGAATCTTCCAATGAGACCTTTGCGCCAAATAATGTGGATTTGGAAAAA
 TATAAGTCATCTCAGAGTAATATACTACCGAAGTTTATGAGGCATCGAGCTTTGAAGAAAAAGTAAGCTCAGAAAAACCTCAATA
 CTCAATTCTGGAAGAAAATCTATTATGAATATGTGGTGGTTGACAAATCAATCTTGGGTGTTTCTATTCTGGATTCAATTTATGTACA
 AGGACTTGGAAGCCCGTCGAAAAAGAAAGGCGGGTTTGGTCTGTTACAATTATTGTTACTTCTGGCTTGCTGAATGTTTCAATATC
 ACTTGGCAAATTTGCAGCTACAGGTCTACAACCTGGGTCTAAATTTGGTGGCAGTGTGGGATAACAATTTGGATTGGGTACGGTTTCGT
 TTGCTTTTGTGTTTTTGGCCTCTAGAGTTGGATCTGCTTATCATTGTCATTCCTATATCATCTAGAGCATCATTCGGTATTTTTCT

[illegible]

Challenges in Computational Biology

