

Rosalind Python Exam Quick Cheatsheet

1. DNA Counting - Count A, C, G, T in string.
2. RNA Transcription - Replace T with U.
3. Reverse Complement - Reverse DNA, swap A<->T, C<->G.
4. GC Content - Find % of G and C in sequence.
5. Hamming Distance - Count differing positions of two strings.
6. Mendel's Law - Use probability of dominant/recessive crosses.
7. Fibonacci Rabbits - Classic recurrence: $F(n)=F(n-1)+k \cdot F(n-2)$.
8. Mortal Fibonacci - Track rabbit lifespans with DP.
9. Protein Translation - Convert codons (3 RNA) to amino acids.
10. RNA Splicing - Remove introns, then translate.
11. Substring Locations - Find all indexes of substring in DNA.
12. Consensus & Profile - Build profile matrix, consensus string.
13. Overlap Graph - Draw edges if suffix==prefix.
14. Longest Common Substring - Use suffix arrays/LCS DP.
15. k-mers - Count frequency of substrings of length k.
16. Motif Finding - Find positions where motif occurs in DNA set.
17. Translation Table - RNA codon -> Protein mapping.
18. Point Mutations - Same as Hamming distance.
19. Protein Mass - Sum monoisotopic weights of amino acids.
20. Spectrum Graph - Build graph of peptide mass differences.
21. Perfect Matching - Matchings in basepairs: factorial count.
22. RNA Secondary Structure - Catalan/DP count of matchings.
23. Partial Permutations - Formula $P(n,k) = n!/(n-k)!$
24. Log Probability DNA - $\log_{10}(\text{probability})$ for DNA string.
25. Signed Permutations - Generate permutations with +/- signs.
26. Subsequence - Find indices in s where t is subsequence.
27. Transition/Transversion Ratio - Count purine vs pyrimidine swaps.
28. Tree Connections - Minimum edges = components-1.
29. Noncrossing Matchings - Catalan recursion.
30. Error Correction - Reads differ by Hamming=1 fix.

31. Binary Tree Internal Nodes - Formula: $n-2$.
32. k-mer Composition - Count all k-length substrings lexicographically.
33. Failure Array - KMP prefix table of string.
34. LCS - Longest Common Subsequence via DP.
35. Lexicographic Strings - Generate all strings from alphabet $\leq n$.
36. Max Matchings - Maximum RNA pairings possible.
37. Distance Matrix - p-distance for DNA set.
38. Reversal Distance - Count minimal reversals to match perms.
39. Random Motifs - Probability random DNA has motif.
40. Counting Subsets - Answer is $2^n \bmod 1e6$.
41. Alternative Splicing - Sum $\text{binomial}(n,k)$ for $k \geq m$.
42. Edit Distance - DP for insert/replace/delete.
43. Expected Restriction Sites - Expected count = $\text{prob} \times \text{length}$.
44. Motzkin Numbers - DP for RNA noncrossing matchings.
45. Distances in Trees - Parse Newick format, find path length.