**DNA Analysis Roadmap Using Your 5 Algorithms**

| **Algorithm** | **Best Used For** | **Type** | **Use Case (Class-Level)** | **Insights Gained** |
| --- | --- | --- | --- | --- |
| **1. Backtracking** | **- All possible LCS / motif discovery - Exhaustive exploration** | **Pairwise / small group** | **Intra-class (top N sequences)** | **- Common subsequences - Conserved motifs** |
| **2. Divide and Conquer** | **- Efficient LCS/LCP on long sequences - Memory optimization** | **Pairwise** | **Inter-class comparisons** | **- Structural comparison between classes** |
| **3. Dynamic Programming** | **- Standard LCS calculation - Similarity & mutation analysis** | **Pairwise** | **Intra & Inter-class** | **- Similarity measure - Avg. conservation** |
| **4. Greedy** | **- DNA assembly simulation - Overlap merging** | **Multi-sequence** | **Intra-class or synthetic mixes** | **- Sequence reconstruction - Fragment merging** |
| **5. Trie** | **- Common prefix/suffix - Fast pattern search / repetition detection** | **Single/multi sequence** | **Intra-class** | **- Class signature motifs - Unique identifiers** |

**🧪 Biological Tasks & Algorithm Match**

| **Biological Task** | **Use This Algorithm** | **Why?** |
| --- | --- | --- |
| **Compare sequences within a class (conserved regions)** | **DP or Backtracking** | **DP is fast for LCS length; Backtracking gives patterns** |
| **Compare different classes (evolutionary diff.)** | **DP + Divide & Conquer** | **Fast for long DNA pairs; space-efficient** |
| **Extract all LCS strings from top similar sequences** | **Backtracking** | **Finds all possible subsequences** |
| **Reconstruct sequences from fragments** | **Greedy** | **Merges based on maximal overlaps** |
| **Find shared prefixes/suffixes across a class** | **Trie** | **Trie structure is ideal for prefix queries** |
| **Create unique class signatures** | **Trie + DP** | **Combine character frequency + LCS results** |

**🧬 Suggested Workflow by Stage**

**🔹 1. Intra-Class Analysis (Within each class)**

* **✅ Use DP to calculate average LCS length across sequences**
* **🔍 Use Backtracking on top 10 to find common LCS strings**
* **🌱 Build a Trie to detect shared motifs or prefixes**

**🔹 2. Inter-Class Comparison (Across classes)**

* **⚡ Use Divide and Conquer for efficient LCS on longer DNA**
* **📏 Compare avg LCS length across classes using DP**

**🔹 3. Pattern Discovery & Analysis**

* **🌟 Backtrack for all LCS variants**
* **🧬 Use Trie to highlight most common starting patterns or repeats**

**🔹 4. DNA Assembly Simulation**

* **🧩 Break sequences into fragments (simulate sequencing)**
* **🔗 Reassemble using Greedy overlap-based algorithm**

**🔹 5. Benchmarking & Visualization**

* **📊 Compare algorithm results (DP vs Divide & Conquer)**
* **🧮 Show differences in computation time and accuracy**
* **📈 Plot LCS length distributions and motif frequencies**

**📊 Optional Visualizations**

| **Visualization** | **Based On** | **Tool Idea** |
| --- | --- | --- |
| **Heatmap of class similarity** | **DP or Divide & Conquer** | **LCS matrix between classes** |
| **Motif cloud per class** | **Trie / Backtracking** | **Most common subsequences** |
| **DNA reconstruction graph** | **Greedy** | **Sequence overlap chains** |
| **Prefix tree visualization** | **Trie** | **Highlight common entry points** |