

## Automated mtDNA Data Analysis Software

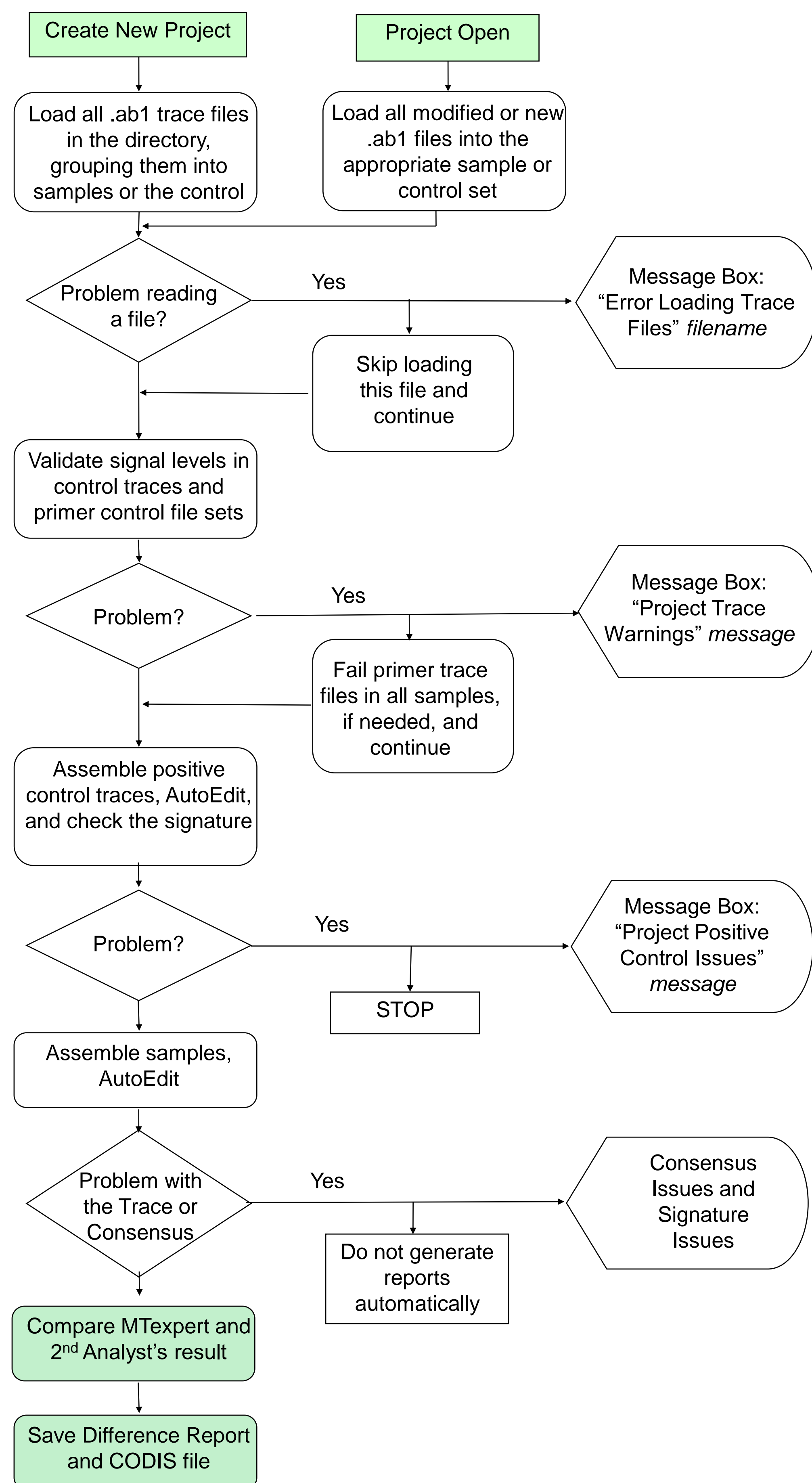
An automated software expert system for mitochondrial DNA data analysis, developed under contract with the US FBI DNAUII, allows forensic analysts to get consistent, accurate results more quickly.

In MTextpert, trace files are assembled into a consensus sequence and the mitotype is generated using the standardized Mitotyper nomenclature rules. Expert system rules and procedures are incorporated to trim trace data and resolve routine problems with the assembly. The analysis of quality control samples is also automated. Every step in the data analysis process, both automatic and manual, is tracked and logged. Project files that archive the data, the analysis process log, the assembled sample consensus sequence, and the type description are automatically generated.

The software completely automates the analysis of high quality data sets, replacing the need for one of the two forensic analysts in a two-reviewer process. When this is not possible, the software efficiently directs the analysts' attention to the issues that prevented automated completion of a project.



### Efficiency: Workflow automation reduces analysis time by up to 80%

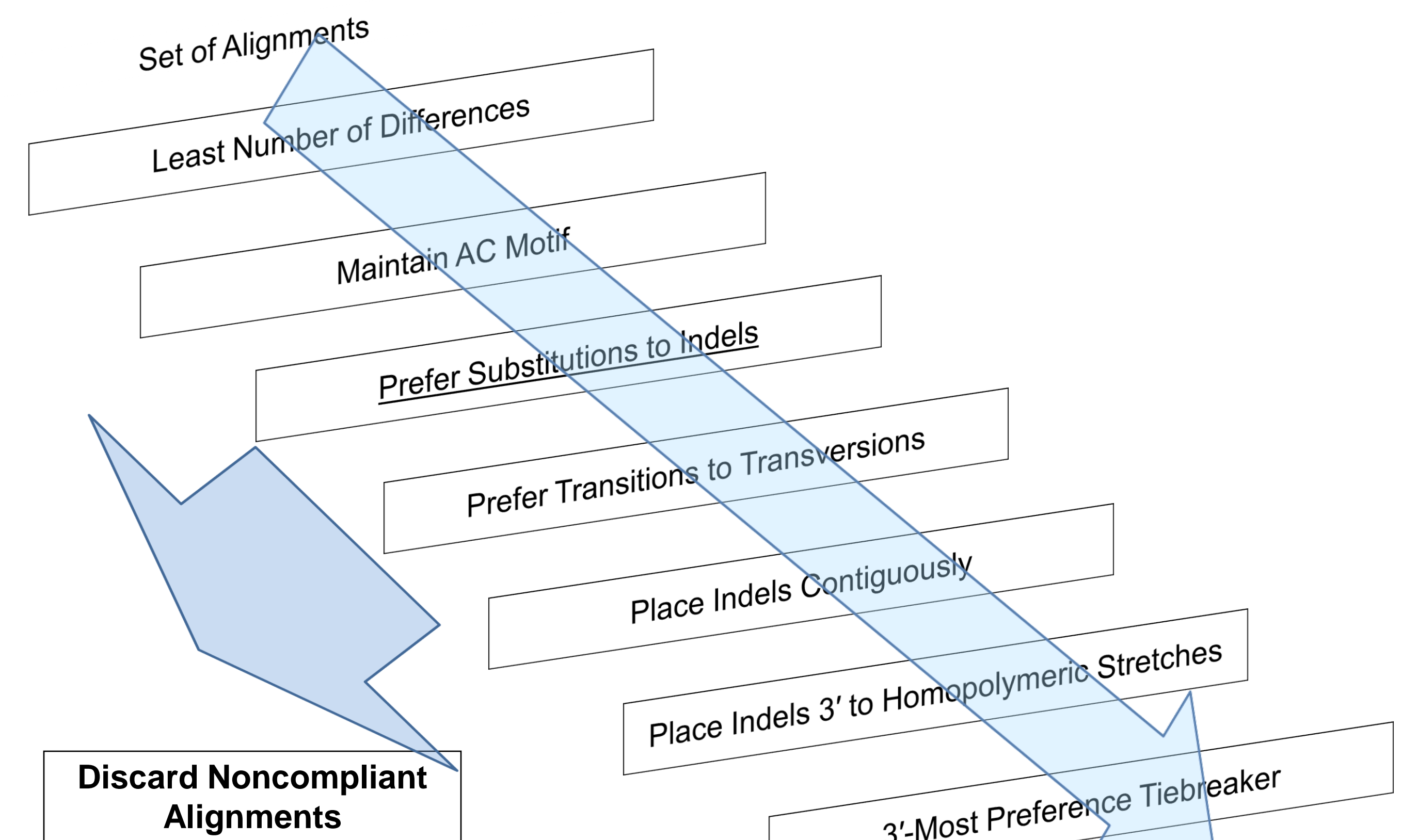


### Quality: Highlights trace and assembly issues for review

Each base in each trace and in the the assembled consensus sequence is reviewed. The operator can be alerted – and automated processing stopped – when issues are encountered. Alerts are generated for:

- Unassembled trace heteroplasmies
- Consensus heteroplasmies
- Low quality bases in the consensus
- Coverage below minimums
- Trace disagreements

### Consistency: Uses the standard Mitotyper Rules™ to create a standard difference report



The Mitotyper Rules are used to describe mtDNA sequences by listing only those sites that differ from a reference sequence. These rules were developed to reflect both historical and current operational nomenclature. In addition to encoding the standard practice, rules were created to stabilize the operational nomenclature and to ensure that only a single compliant description will result. Implementing these rules in software ensures absolute consistency and stability in describing mtDNA types. With this approach, once named and entered into the database haplotypes never need to be renamed, identical sequences will always be listed the same way, and database searching for statistical inferences will be greatly facilitated.

### Automated Edits: Routine issues are recognized and resolved

Any number of automated procedures can be created to detect and correct common issues. Procedures can be keyed on primers, dye chemistry, trace sequence and appearance, and quality values. Current automated procedures include:

- C-stretch alignment
- Low quality consensus inserts
- Primer – specific sequence distortions
- Chemistry – specific trace analysis

The Mitotyper Rules were validated using historical sequence descriptions in a wide variety of mtDNA type databases as shown in the table. Greater than 99% congruence between the software and historical results was achieved in all cases. Expert examination resolved the few discrepancies in the SWGDAM data in favor of the description produced by the Mitotyper Rules.

Dataset	Samples	Regions	Agreement	Resolution
SWGDAM	4,839	40,357	99.92%	100%
	1,254	11,303	99.82%	100%
Japanese	1,204	8,859	99.88%	
EMPOP	5,173	44,278	99.88%	