GeneCrafter v1.1 Report for COL1A1

Result Summary

Metric	Before	After
CAI Score	N/A	**0.8135**
GC Content (%)	76.8	**62.8** ▼
Complex DNA Repeats	2900	**8** ▼
Immunogenicity Score	0.09	**0.08**

Interpretation

- ℹ️ **GC Content Adjusted:** The GC content moved from 76.8% to 62.8%, getting closer to the target of 50.0%.
- ℹ️ **Improvement:** The 'Complex DNA Repeats' score was reduced from 2900 to 8.
- ℹ️ **Improvement:** The 'Immunogenicity Score' score was reduced from 0.09 to 0.08.

Optimization Options

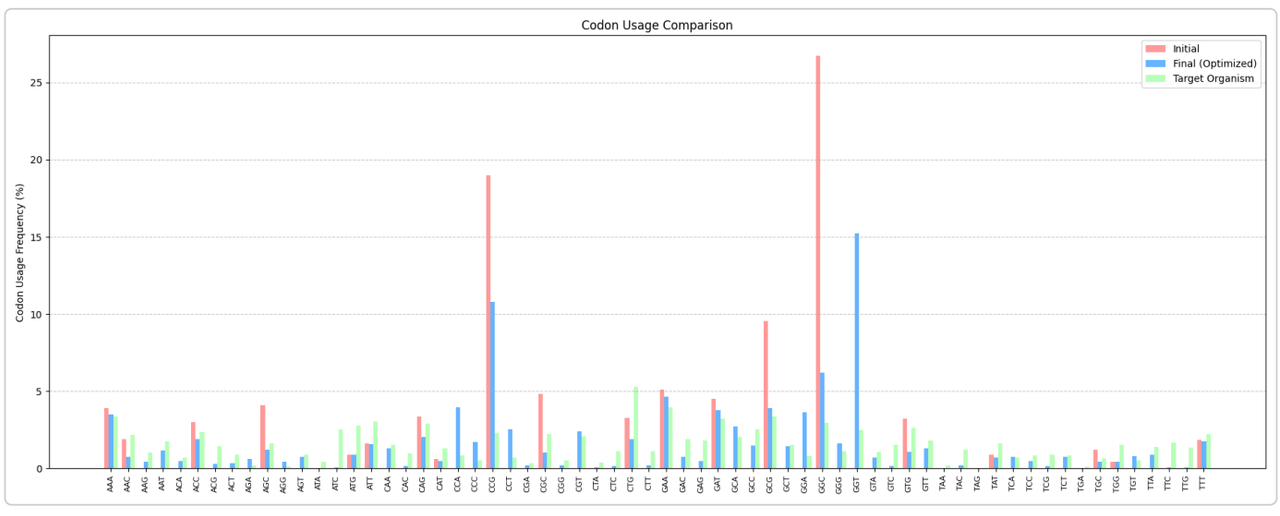
Option	Value
Organism	Escherichia coli str. K-12 substr. MG1655
Optimizer	GA
Goal	Maximize
Targets	cai, gc, cpb, tfbs
Target GC %	50.0 (User Input)
Avoided Enzymes	NdeI(CATATG), XhoI(CTCGAG)

Detailed Analysis

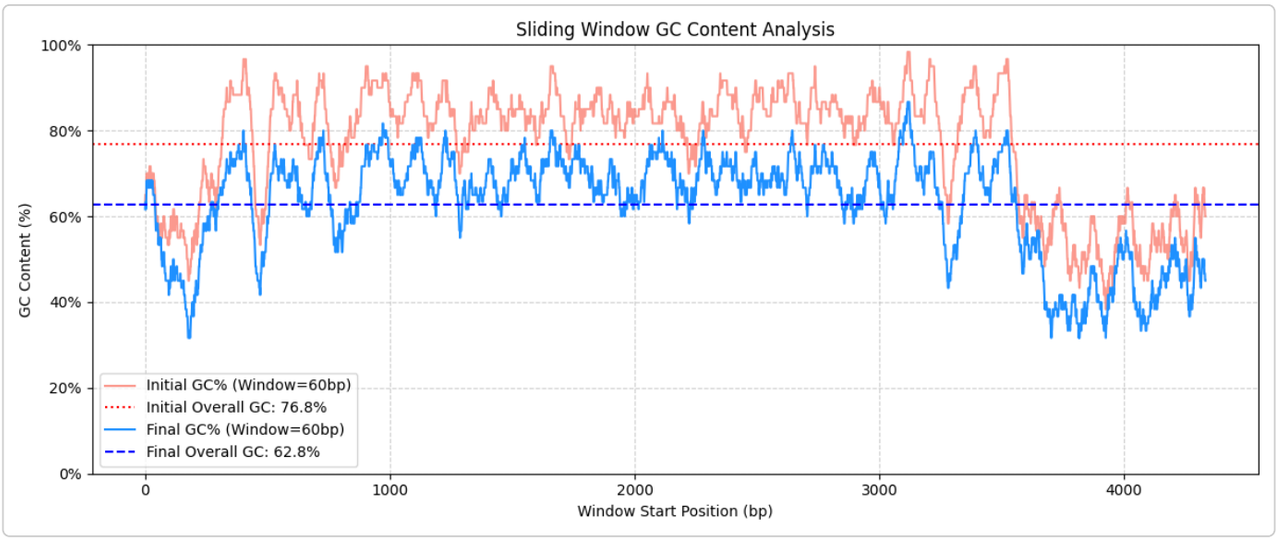
Metric	Initial	Final (Optimized)
Sequence Length	4392	4392
Gc Content Percent	76.8	62.8
Cai Score	N/A	0.8135
Cpb Score	0.0000	0.0000
Tfbs Count	1	1
Restricted Enzyme Sites	0	0
Five Prime Mfe Kcal Mol	-14.90	-12.65
Cpg Dinucleotide Count	761	392
Upa Dinucleotide Count	39	118
Immunogenicity Score	0.09	0.08
Immunogenicity Risk	Low	Low
Complex Dna Repeats Count	2900	8

Visual Analysis

Codon Usage Analysis

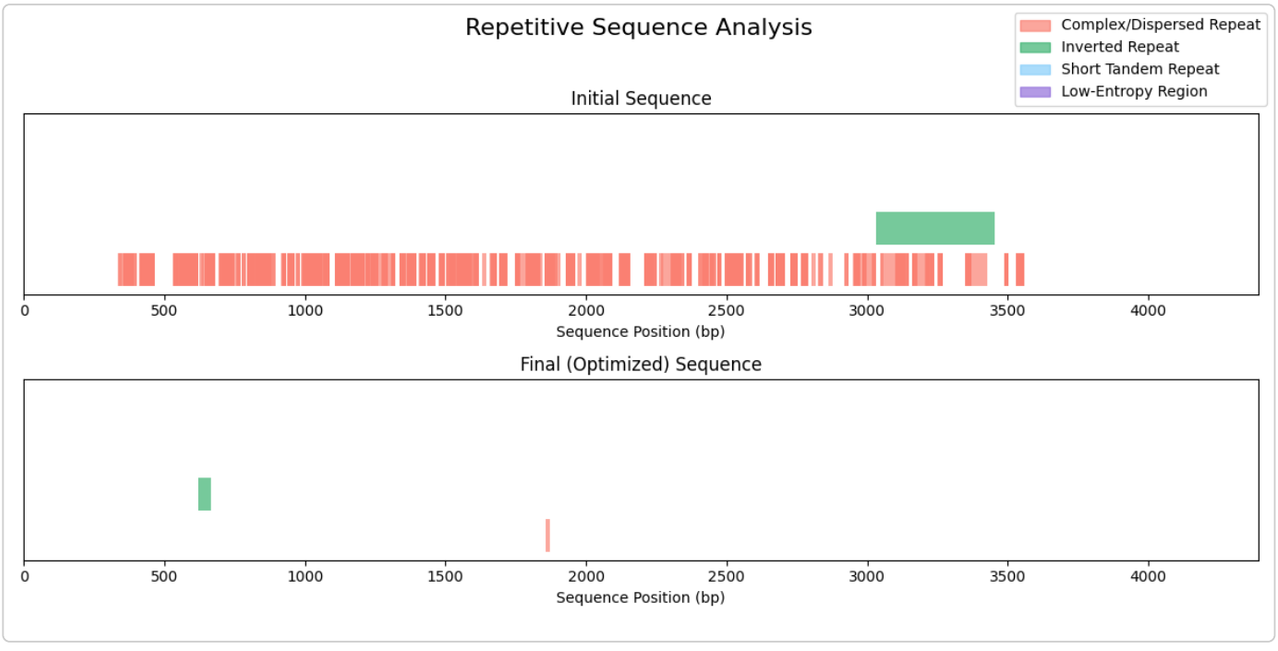


GC Content Distribution



Repetitive Sequence Analysis

This plot shows locations of complex repeats. Effective optimization should eliminate these regions.



RNA Secondary Structure (2D Visualization)

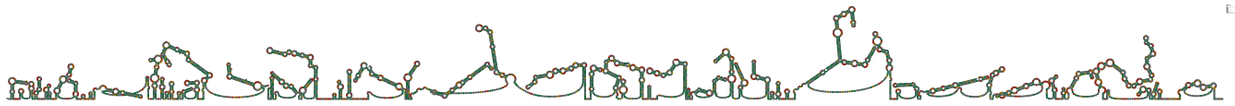
DMS-reactivity-style coloring: Green=base-paired A/C, Orange-Red=unpaired A/C, Grey=U/G (not DMS-reactive). Top: **Before Optimization** | Bottom: **After Optimization**.

Before Optimization (COL1A1)

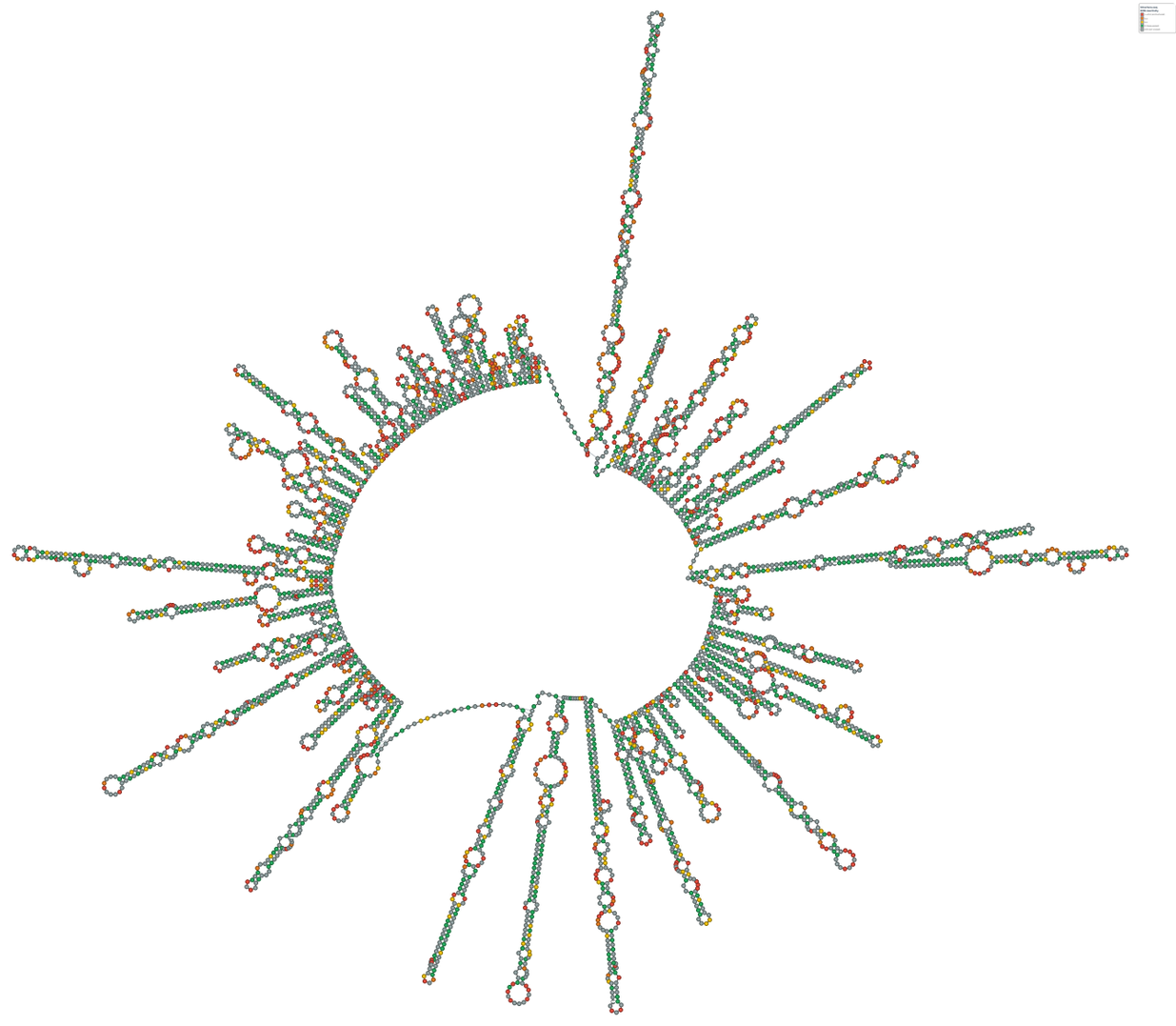
Seeded LinearFold | Layout: RNAplot Puzzler + Refine

MFE: -2386.45 kcal/mol | **Shape:** Multi-domain (Dumbbell) | **Paired:** 2,774 / 4,392 nt (63.2%) | **Max depth:** 78 | (20.00s)

RNAplot Puzzler + Refine



RNAplot Naview + Refine

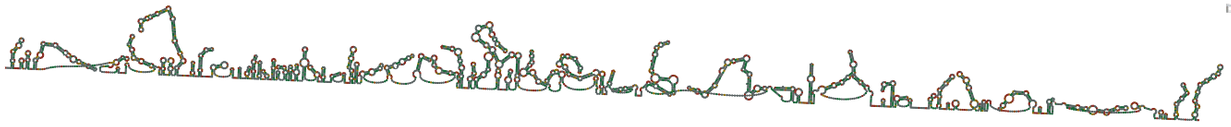


After Optimization (COL1A1)

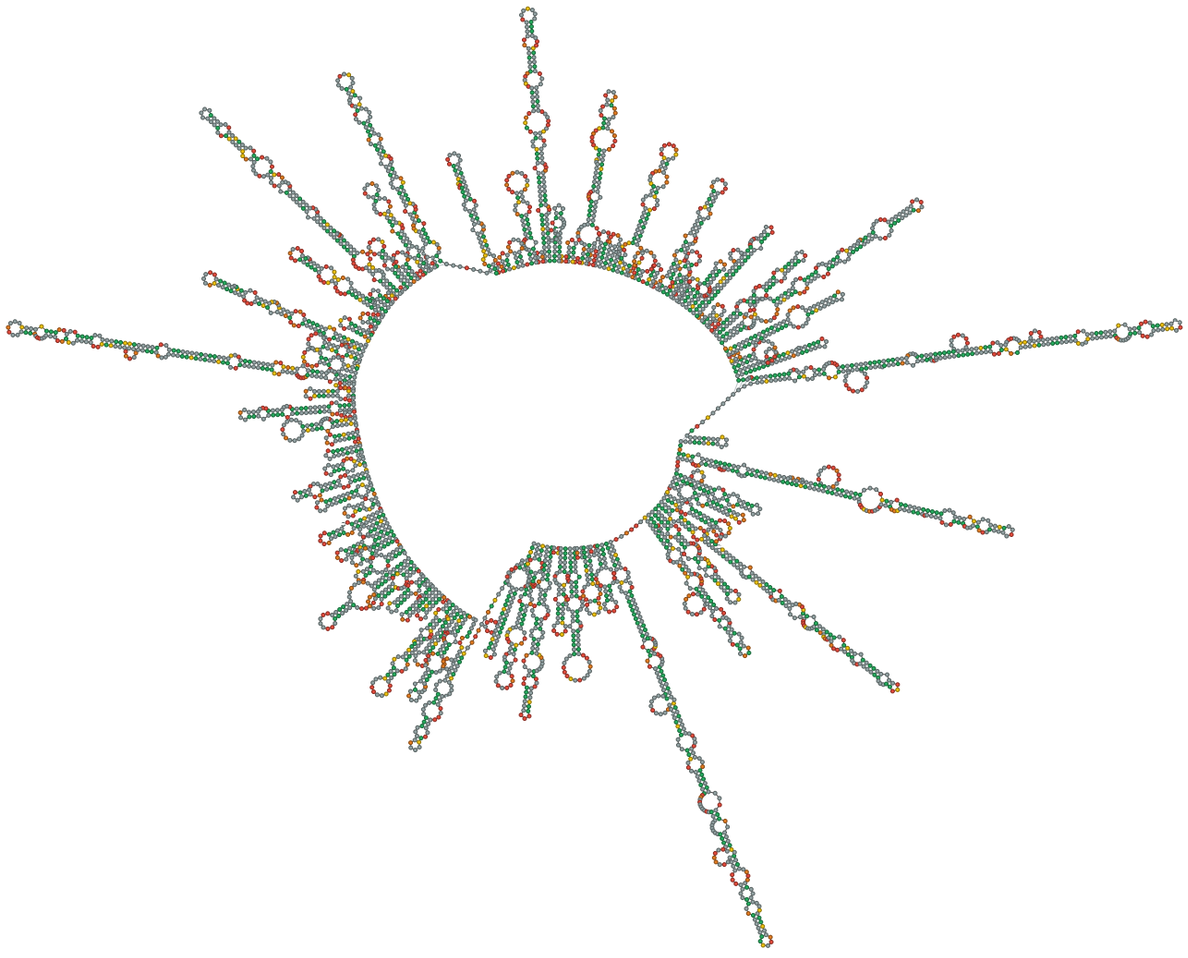
Seeded LinearFold | Layout: RNAplot Puzzler + Refine

MFE: -1786.15 kcal/mol | **Shape:** Multi-domain (Dumbbell) | **Paired:** 2,630 / 4,392 nt (59.9%) | **Max depth:** 80 | (20.17s)

RNAplot Puzzler + Refine



RNAplot Naview + Refine



Input Sequence (FASTA)

```
>Initial_Sequence
ATGTTTAGCTTTGTGGATCTGCGCCTGCTGCTGCTGCTGGCGGCGACCGCGCTGCTGACCCATGGCCAGGAAGAAGGCCAGGTGGAAGGCCAGGA
TGAAGATATTCCGCCGATTACCTGCGTGCAGAACGGCCTGCGCTATCATGATCGCGATGTGTGGAAACCGGAACCGTGCCGCATTTGCGTGTGCG
ATAACGGCAAAGTGCTGTGCGATGATGTGATTTGCGATGAAACCAAAAACTGCCCGGGCGCGGAAGTGCCGGAAGGCGAATGCTGCCCGGTGTGC
CCGGATGGCAGCGAAAGCCCGACCGATCAGGAAACCACCGGCGTGGAAGGCCCGAAAGGCGATACCGGCCCGCGCGGCCCGCGCGGCCCGGCGGG
CCCGCCGGGCCGCGATGGCATTCCGGGCCAGCCGGGCCTGCCGGGCCCGCCGGGCCCGCCGGGCCCGCCGGGCCCGCCGGGCCTGGGCGGCAACT
TTGCGCCGCAGCTGAGCTATGGCTATGATGAAAAAAGCACCGGCGGCATTAGCGTGCCGGGCCCGATGGGCCCGAGCGGCCCGCGCGGCCTGCCG
GGCCCGCCGGGCGCGCCGGGCCCGCAGGGCTTTCAGGGCCCGCCGGGCGAACCGGGCGAACCGGGCGCGAGCGGCCCGATGGGCCCGCGCGGCCC
GCCGGGCCCGCCGGGCAAAAACGGCGATGATGGCGAAGCGGGCAAACCGGGCCGCCCGGGCGAACGCGGCCCGCCGGGCCCGCAGGGCGCGCGCG
GCCTGCCGGGCACCGCGGGCCTGCCGGGCATGAAAGGCCATCGCGGCTTTAGCGGCCTGGATGGCGCGAAAGGCGATGCGGGCCCGGCGGGCCCG
AAAGGCGAACCGGGCAGCCCGGGCGAAAACGGCGCGCCGGGCCAGATGGGCCCGCGCGGCCTGCCGGGCGAACGCGGCCGCCCGGGCGCGCCGGG
CCCGGCGGGCGCGCGCGGCAACGATGGCGCGACCGGCGCGGCGGGCCCGCCGGGCCCGACCGGCCCGGCGGGCCCGCCGGGCTTTCCGGGCGCGG
TGGGCGCGAAAGGCGAAGCGGGCCCGCAGGGCCCGCGCGGCAGCGAAGGCCCGCAGGGCGTGCGCGGCGAACCGGGCCCGCCGGGCCCGGCGGGC
GCGGCGGGCCCGGCGGGCAACCCGGGCGCGGATGGCCAGCCGGGCGCGAAAGGCGCGAACGGCGCGCCGGGCATTGCGGGCGCGCCGGGCTTTCC
GGGCGCGCGCGGCCCGAGCGGCCCGCAGGGCCCGGGCGGCCCGCCGGGCCCGAAAGGCAACAGCGGCGAACCGGGCGCGCCGGGCAGCAAAGGCG
ATACCGGCGCGAAAGGCGAACCGGGCCCGGTGGGCGTGCAGGGCCCGCCGGGCCCGGCGGGCGAAGAAGGCAAACGCGGCGCGCGCGGCGAACCG
GGCCCGACCGGCCTGCCGGGCCCGCCGGGCGAACGCGGCGGCCCGGGCAGCCGCGGCTTTCCGGGCGCGGATGGCGTGGCGGGCCCGAAAGGCCC
GGCGGGCGAACGCGGCAGCCCGGGCCCGGCGGGCCCGAAAGGCAGCCCGGGCGAAGCGGGCCGCCCGGGCGAAGCGGGCCTGCCGGGCGCGAAAG
GCCTGACCGGCAGCCCGGGCAGCCCGGGCCCGGATGGCAAAACCGGCCCGCCGGGCCCGGCGGGCCAGGATGGCCGCCCGGGCCCGCCGGGCCCG
CCGGGCGCGCGCGGCCAGGCGGGCGTGATGGGCTTTCCGGGCCCGAAAGGCGCGGCGGGCGAACCGGGCAAAGCGGGCGAACGCGGCGTGCCGGG
CCCGCCGGGCGCGGTGGGCCCGGCGGGCAAAGATGGCGAAGCGGGCGCGCAGGGCCCGCCGGGCCCGGCGGGCCCGGCGGGCGAACGCGGCGAAC
AGGGCCCGGCGGGCAGCCCGGGCTTTCAGGGCCTGCCGGGCCCGGCGGGCCCGCCGGGCGAAGCGGGCAAACCGGGCGAACAGGGCGTGCCGGGC
GATCTGGGCGCGCCGGGCCCGAGCGGCGCGCGCGGCGAACGCGGCTTTCCGGGCGAACGCGGCGTGCAGGGCCCGCCGGGCCCGGCGGGCCCGCG
CGGCGCGAACGGCGCGCCGGGCAACGATGGCGCGAAAGGCGATGCGGGCGCGCCGGGCGCGCCGGGCAGCCAGGGCGCGCCGGGCCTGCAGGGCA
TGCCGGGCGAACGCGGCGCGGCGGGCCTGCCGGGCCCGAAAGGCGATCGCGGCGATGCGGGCCCGAAAGGCGCGGATGGCAGCCCGGGCAAAGAT
GGCGTGCGCGGCCTGACCGGCCCGATTGGCCCGCCGGGCCCGGCGGGCGCGCCGGGCGATAAAGGCGAAAGCGGCCCGAGCGGCCCGGCGGGCCC
GACCGGCGCGCGCGGCGCGCCGGGCGATCGCGGCGAACCGGGCCCGCCGGGCCCGGCGGGCTTTGCGGGCCCGCCGGGCGCGGATGGCCAGCCGG
GCGCGAAAGGCGAACCGGGCGATGCGGGCGCGAAAGGCGATGCGGGCCCGCCGGGCCCGGCGGGCCCGGCGGGCCCGCCGGGCCCGATTGGCAAC
GTGGGCGCGCCGGGCGCGAAAGGCGCGCGCGGCAGCGCGGGCCCGCCGGGCGCGACCGGCTTTCCGGGCGCGGCGGGCCGCGTGGGCCCGCCGGG
CCCGAGCGGCAACGCGGGCCCGCCGGGCCCGCCGGGCCCGGCGGGCAAAGAAGGCGGCAAAGGCCCGCGCGGCGAAACCGGCCCGGCGGGCCGCC
CGGGCGAAGTGGGCCCGCCGGGCCCGCCGGGCCCGGCGGGCGAAAAAGGCAGCCCGGGCGCGGATGGCCCGGCGGGCGCGCCGGGCACCCCGGGC
CCGCAGGGCATTGCGGGCCAGCGCGGCGTGGTGGGCCTGCCGGGCCAGCGCGGCGAACGCGGCTTTCCGGGCCTGCCGGGCCCGAGCGGCGAACC
GGGCAAACAGGGCCCGAGCGGCGCGAGCGGCGAACGCGGCCCGCCGGGCCCGATGGGCCCGCCGGGCCTGGCGGGCCCGCCGGGCGAAAGCGGCC
GCGAAGGCGCGCCGGGCGCGGAAGGCAGCCCGGGCCGCGATGGCAGCCCGGGCGCGAAAGGCGATCGCGGCGAAACCGGCCCGGCGGGCCCGCCG
GGCGCGCCGGGCGCGCCGGGCGCGCCGGGCCCGGTGGGCCCGGCGGGCAAAAGCGGCGATCGCGGCGAAACCGGCCCGGCGGGCCCGGCGGGCCC
GGTGGGCCCGGTGGGCGCGCGGCCCGGCGGGCCCGCAGGGCCCGCGCGGCGATAAAGGCGAAACCGGCGAACAGGGCGATCGCGGCATTAAAG
GCCATCGCGGCTTTAGCGGCCTGCAGGGCCCGCCGGGCCCGCCGGGCAGCCCGGGCGAACAGGGCCCGAGCGGCGCGAGCGGCCCGGCGGGCCCG
CGCGGCCCGCCGGGCAGCGCGGGCGCGCCGGGCAAAGATGGCCTGAACGGCCTGCCGGGCCCGATTGGCCCGCCGGGCCCGCGCGGCCGCACCGG
CGATGCGGGCCCGGTGGGCCCGCCGGGCCCGCCGGGCCCGCCGGGCCCGCCGGGCCCGCCGAGCGCGGGCTTTGATTTTAGCTTTCTGCCGCAGC
CGCCGCAGGAAAAAGCGCATGATGGCGGCCGCTATTATCGCGCGGATGATGCGAACGTGGTGCGCGATCGCGATCTGGAAGTGGATACCACCCTG
AAAAGCCTGAGCCAGCAGATTGAAAACATTCGCAGCCCGGAAGGCAGCCGCAAAAACCCGGCGCGCACCTGCCGCGATCTGAAAATGTGCCATAG
CGATTGGAAAAGCGGCGAATATTGGATTGATCCGAACCAGGGCTGCAACCTGGATGCGATTAAAGTGTTTTGCAACATGGAAACCGGCGAAACCT
GCGTGTATCCGACCCAGCCGAGCGTGGCGCAGAAAAACTGGTATATTAGCAAAAACCCGAAAGATAAACGCCATGTGTGGTTTGGCGAAAGCATG
ACCGATGGCTTTCAGTTTGAATATGGCGGCCAGGGCAGCGATCCGGCGGATGTGGCGATTCAGCTGACCTTTCTGCGCCTGATGAGCACCGAAGC
GAGCCAGAACATTACCTATCATTGCAAAAACAGCGTGGCGTATATGGATCAGCAGACCGGCAACCTGAAAAAAGCGCTGCTGCTGCAGGGCAGCA
ACGAAATTGAAATTCGCGCGGAAGGCAACAGCCGCTTTACCTATAGCGTGACCGTGGATGGCTGCACCAGCCATACCGGCGCGTGGGGCAAAACC
GTGATTGAATATAAAACCACCAAAACCAGCCGCCTGCCGATTATTGATGTGGCGCCGCTGGATGTGGGCGCGCCGGATCAGGAATTTGGCTTTGA
TGTGGGCCCGGTGTGCTTTCTG
```

Final Optimized Sequence (FASTA) (Identity: 81.1%)

```
>Optimized_COL1A1
ATGTTTAGCTTTGTGGACCTGCGACTGCTGCTGCTGGCCGCAACCGCGCTGCTGACCCATGGCCAGGAAGAAGGCCAGGTGGAAGGTCAAGA
CGAAGATATTCCGCCGATTACCTGTGTGCAAAATGGTCTGCGATATCATGATCGTGATGTATGGAAACCGGAACCGTGCCGTATTTGCGTTTGTG
ATAACGGTAAAGTGTTATGTGATGATGTTATTTGTGATGAGACTAAAAACTGTCCGGGTGCAGAAGTACCGGAAGGTGAATGTTGTCCGGTGTGC
CCAGATGGTAGCGAATCCCCGACCGACCAGGAAACCACCGGCGTTGAAGGACCAAAGGGTGATACGGGCCCGAGAGGCCCGCGTGGACCTGCAGG
ACCGCCCGGTCGCGATGGTATTCCAGGGCAGCCGGGACTGCCGGGACCGCCAGGGCCACCCGGTCCTCCGGGACCTCCGGGTTTAGGTGGTAACT
TTGCGCCCCAACTGTCCTATGGTTACGATGAAAAAAGTACAGGTGGTATTAGTGTGCCGGGCCCCATGGGTCCGAGCGGTCCGCGTGGACTGCCG
GGTCCACCTGGTGCACCAGGACCACAGGGCTTTCAGGGCCCGCCGGGTGAACCGGGTGAACCAGGTGCAAGCGGTCCGATGGGTCCACGCGGACC
TCCGGGTCCACCGGGTAAAAACGGTGATGATGGTGAAGCAGGCAAGCCGGGTAGACCCGGTGAGCGCGGCCCTCCGGGTCCGCAGGGTGCACGTG
GTCTGCCGGGCACAGCGGGTCTGCCGGGTATGAAAGGTCACCGCGGTTTTAGCGGTTTAGATGGTGCTAAAGGTGATGCGGGACCAGCCGGTCCT
AAAGGTGAACCGGGTTCGCCGGGTGAAAATGGTGCTCCAGGTCAGATGGGTCCTCGTGGTCTTCCGGGAGAACGCGGCCGCCCGGGTGCGCCTGG
TCCAGCTGGGGCCAGGGGTAATGATGGGGCGACCGGAGCGGCGGGTCCGCCTGGCCCAACCGGCCCCGCGGGTCCTCCGGGTTTTCCGGGTGCGG
TAGGCGCTAAAGGCGAAGCCGGTCCGCAAGGTCCTCGGGGTAGCGAAGGACCTCAGGGTGTTCGCGGAGAACCGGGTCCGCCCGGTCCGGCTGGT
GCGGCGGGTCCGGCAGGTAATCCGGGCGCAGATGGCCAGCCGGGCGCGAAAGGTGCAAACGGCGCACCAGGAATTGCAGGTGCCCCTGGCTTTCC
AGGAGCACGTGGTCCGTCCGGTCCGCAGGGCCCGGGTGGCCCGCCGGGGCCGAAAGGTAATTCTGGTGAACCGGGCGCGCCTGGCAGTAAAGGCG
ATACCGGTGCTAAAGGAGAACCGGGCCCGGTGGGTGTTCAAGGCCCGCCAGGGCCTGCAGGTGAAGAAGGTAAACGTGGTGCGAGGGGTGAGCCA
GGTCCTACCGGCCTGCCGGGGCCACCGGGTGAACGTGGTGGTCCAGGTTCACGTGGATTTCCAGGTGCAGATGGTGTGGCCGGCCCGAAAGGCCC
GGCGGGTGAAAGAGGTTCACCCGGTCCAGCAGGCCCGAAGGGGTCACCGGGTGAGGCGGGTCGCCCGGGTGAAGCTGGCCTCCCGGGTGCGAAAG
GTCTTACCGGTAGCCCAGGTAGCCCAGGCCCGGATGGCAAAACAGGACCGCCGGGTCCGGCGGGTCAGGACGGTCGTCCAGGTCCCCCGGGACCC
CCGGGAGCACGTGGTCAGGCGGGCGTAATGGGCTTTCCGGGCCCGAAAGGTGCGGCGGGTGAGCCTGGCAAAGCAGGCGAACGCGGCGTACCGGG
ACCGCCCGGCGCTGTTGGTCCCGCGGGCAAAGATGGAGAAGCGGGAGCGCAAGGTCCGCCAGGTCCTGCAGGTCCCGCCGGTGAAAGAGGAGAAC
AAGGACCGGCGGGTAGTCCGGGCTTTCAAGGCCTGCCAGGGCCGGCTGGTCCGCCTGGCGAAGCTGGTAAACCCGGCGAACAAGGTGTTCCGGGT
GATTTAGGTGCACCGGGCCCGAGTGGCGCAAGAGGGGAGCGTGGTTTTCCCGGAGAAAGAGGTGTACAGGGTCCTCCGGGTCCAGCCGGCCCGCG
TGGTGCTAACGGCGCGCCGGGTAACGATGGCGCCAAGGGGGATGCGGGTGCGCCGGGAGCACCGGGCAGTCAGGGCGCGCCGGGTTTACAGGGTA
TGCCGGGAGAACGCGGTGCCGCGGGCTTACCAGGCCCGAAAGGAGATAGGGGTGACGCTGGACCGAAAGGGGCGGATGGAAGCCCGGGAAAAGAT
GGTGTGCGCGGCTTAACCGGCCCGATCGGCCCGCCGGGTCCGGCTGGCGCGCCGGGAGACAAAGGAGAATCAGGTCCATCGGGCCCAGCGGGTCC
AACGGGTGCTCGTGGTGCGCCGGGAGATCGTGGTGAACCGGGGCCTCCAGGACCGGCCGGCTTTGCAGGTCCACCGGGTGCAGATGGCCAGCCAG
GCGCAAAAGGTGAACCCGGTGATGCGGGCGCCAAAGGGGATGCAGGACCGCCTGGGCCAGCAGGTCCGGCAGGACCACCCGGACCGATTGGTAAT
GTTGGTGCGCCGGGTGCCAAAGGCGCCCGTGGCTCTGCGGGACCGCCAGGTGCTACTGGTTTTCCGGGTGCAGCTGGTCGCGTAGGTCCGCCGGG
TCCCTCTGGCAATGCAGGTCCGCCGGGCCCCCCTGGGCCGGCCGGTAAAGAAGGAGGCAAAGGTCCGCGTGGCGAAACAGGTCCAGCTGGGCGTC
CGGGTGAAGTTGGTCCACCGGGCCCGCCAGGTCCGGCCGGCGAAAAAGGTAGCCCGGGTGCGGATGGTCCGGCGGGTGCACCGGGTACCCCAGGT
CCGCAGGGTATTGCGGGTCAGCGTGGTGTGGTCGGTTTGCCGGGTCAACGTGGTGAAAGAGGCTTTCCGGGTCTGCCGGGTCCGAGTGGTGAACC
GGGTAAACAAGGTCCGTCAGGTGCGTCTGGTGAACGGGGTCCTCCGGGGCCGATGGGCCCGCCAGGTCTTGCCGGTCCGCCTGGTGAATCTGGCC
GTGAAGGTGCACCTGGTGCAGAAGGTTCACCAGGACGTGATGGCAGCCCGGGCGCGAAAGGGGATAGGGGAGAAACCGGCCCGGCAGGCCCACCG
GGGGCGCCGGGTGCCCCGGGTGCCCCAGGTCCGGTGGGTCCCGCGGGTAAATCTGGCGATCGCGGCGAAACTGGTCCTGCGGGACCGGCGGGACC
TGTTGGTCCGGTGGGTGCACGTGGTCCCGCAGGGCCACAGGGTCCGCGTGGTGATAAAGGAGAAACCGGTGAACAGGGTGACCGTGGTATTAAAG
GTCATAGAGGTTTTTCAGGTCTACAGGGTCCGCCTGGTCCACCAGGTTCTCCGGGTGAACAGGGTCCTTCAGGTGCGAGCGGCCCCGCTGGGCCT
CGTGGCCCTCCGGGCAGCGCAGGAGCGCCTGGCAAAGATGGTCTGAATGGCCTGCCCGGTCCCATTGGACCTCCGGGCCCAAGGGGTCGTACCGG
GGATGCGGGTCCAGTGGGACCGCCAGGTCCGCCAGGTCCGCCGGGACCCCCTGGTCCGCCGAGCGCGGGTTTTGATTTTAGTTTTCTGCCGCAGC
CGCCGCAAGAAAAAGCACATGACGGTGGACGTTATTATCGTGCCGATGACGCGAATGTGGTACGCGATCGTGACCTGGAAGTGGATACCACCCTG
AAAAGCTTATCACAGCAAATTGAAAATATTAGATCACCTGAAGGAAGTCGTAAAAATCCGGCACGTACTTGTCGTGATTTAAAAATGTGTCATTC
CGATTGGAAATCTGGTGAATATTGGATTGACCCAAATCAGGGTTGCAACTTAGATGCTATTAAAGTTTTTTGTAACATGGAAACAGGGGAAACCT
GCGTATACCCAACCCAACCGAGTGTTGCGCAAAAAAATTGGTATATTAGCAAAAATCCGAAGGATAAACGTCATGTTTGGTTCGGTGAAAGCATG
ACCGATGGTTTTCAGTTTGAATACGGTGGCCAGGGTTCTGATCCGGCGGATGTGGCTATTCAACTGACCTTTCTGCGGTTAATGAGTACCGAAGC
GTCTCAAAATATTACCTATCATTGTAAAAACTCCGTTGCATATATGGATCAGCAGACTGGTAATCTGAAAAAAGCCTTACTGCTGCAAGGTTCTA
ATGAGATTGAAATTCGCGCCGAAGGCAACTCCCGATTTACCTATAGCGTTACGGTTGATGGTTGCACATCACACACCGGTGCGTGGGGTAAAACG
GTTATTGAATATAAAACCACCAAGACATCCAGGCTGCCAATTATTGATGTTGCTCCGTTAGATGTCGGTGCGCCGGATCAAGAATTTGGCTTTGA
TGTAGGTCCAGTTTGTTTTCTC
```