

BIOINFORMATICS APPLIED TO THE Y-DNA HAPLOGROUP Q-M242: MALE NAVAJO GENETIC MARKER

Project ID#

Jordyn Begay, Navajo Preparatory School, New Mexico

Q1: Research Question/Engineering Goal:

- Is there a genetic marker Q-M242 which indicates that the Diné began in the Altai region of Central Asia, sandwiched between Russia, Kazakhstan, Mongolia, and China some 30,000 years ago?
- The problem is to identify a likely location and time of the origin of the Q-M242 Y-Haplogroup. This would identify where the Navajo People split off from other people and began their unique epic journey to the Four Corners Region.

Q3: Data Analysis & Results

1st base	2nd base			3rd base
T	TTT (Phe) Phenylalanine ↑	TCT (Ser) Serine ↑	TAT (Tyr) Tyrosine ↑	TGT (Cys) Cysteine ↑
T	TTC	TCC	TAC	TGC
T	TTA (Met) Methionine ↑	TGA (Stop) Stop (Amber) ↓	TAA (Stop) Stop (Ugaa) ↓	TGA (Glu) Glutamic acid ↑
T	TGG (Gly) Glycine ↑	TGG (Trp) Tryptophan ↑	TGG (Trp) Tryptophan ↑	TGG (Trp) Tryptophan ↑
C	CTT (Leu) Leucine ↑	CCT (Pro) Proline ↑	CAT (His) Histidine ↑	CGT (Arg) Arginine ↑
C	CTC	CCC	CAC	CGC
C	CTA	CCA (Pro) Proline ↑	CAA (Gln) Glutamine ↑	CSA (Arg) Arginine ↑
C	CTG (Leu) Leucine ↑	CCG	CAG	CSG
A	AAT (Asn) Asparagine ↑	AAC (Asn) Asparagine ↑	AAT (Asn) Asparagine ↑	AAT (Asn) Asparagine ↑
A	ATC (Met) Methionine ↑	ACC (Thr) Threonine ↑	AAC	AGC (Ser) Serine ↑
A	ATA	ACA	AAA	AGA (Arg) Arginine ↑
A	ATG (Met) Methionine ↑	ACG	AAG (Lys) Lysine ↑	AGG
G	GTT (Val) Valine ↑	GCT (Ala) Alanine ↑	GAT	GGT (Gly) Glycine ↑
G	GTC	GCC	GAC	GGC
G	GTA	GCA (Ala) Alanine ↑	GAA	GGA (Gly) Glycine ↑
G	GTG	GCG	GAG	GGG

Homo sapiens chromosome Y, GRCh38.p13 Primary Assembly
 NCBI Reference Sequence: NC_000024.10
[GenBank Graphics](#)

>NC_000024.10:12906492-12906657 Homo sapiens chromosome Y, GRCh38.p13 Primary Assembly

```

AACTCTTGATAAACCGTGTCTAGTTCACAGATTAAGTAGTAAATTCAGATGGCAAGATTTTAAAG
TACAGTAGTAACTTAATTGATGATCAATGAAAGTGAATAGTATCTGAACCTATATATGTAAGCTTCTA
CGSCATAGAAGTTGTGCARAAAGGTGACCAAGGTGCTTGGCATTGGCTTAAGGTGTTTTTGGAAA
AAAATCTATTTAACGTACATGGTTTTTCCGCCACCCGCCACCCCTTCAGAGTGTCTAGGTAAGG
TATTATGCTGAAAGCCCTTAAAGCGAAATAACCTTTTTCTAGTTTTAAATCATCATGATAAGGAGGC
ATGAATTGAGATTGGA
    
```

Q2: Methodology/Project Design

- Download the Forward and Reverse Primers for Q-M242
- Put Forward (AACTCTTGATAAACCGTGTCTG) and Reverse (TCCAATCTCAATTCATGCCTC) Primers of Q-M242 into Primer-Blast to identify all 366 base pairs of Q-M242.
- Use EMBOSS revseq to check the Reverse Primer.
- Apply Standard Genetic Code Table to 366 base pairs, to obtain the amino acids and stops from the DNA Codons of Q-M242. Use EMBOSS Transeq to check the frequency of occurrence of the constituent amino acids and stops.
- Show Altai region in Russia-China-Mongolia-Kazakhstan region as a possible origin of Navajo. Migration occurred across the Bering Strait to North America.
- Analyze and communicate data.

Q4: Interpretation & Conclusions

My project demonstrated that I could identify the Y-DNA Haplogroup Q-M242. Thus, my hypothesis was accepted, that my Dine ancestors began their journey about 30,000 years ago in the Altai region of Central Asia, sandwiched between Russia, Kazakhstan, Mongolia, and China. In the future, I plan to attempt to synthesize some of the 20 DNA amino acids identified in Y-Haplogroup Q-M242 and detect these synthesized amino acids via a combination of UV/Visual and deep-Infrared Spectrophotometry.