Geographical Mapping of SNP Variants in Vitamin D Metabolism Genes Across the Caucasus: Bioinformatic Analysis of Genetic Data

Ilia State University
Faculty of Natural Sciences and Medicine Applied Genetics

Supervisor
Dr. David Tarkhnishvili
Student
Dr. Mariam Berikashvili

2025

List of Abbreviations:

CSV - comma separated values

CYP27A1 – Cytochrome P450 Family 27 Subfamily A Member 1

CYP27A1 – Cytochrome P450 Family 27 Subfamily A Member 1 gene

CYP2R1 – Cytochrome P450 Family 2 Subfamily R Member 1

CYP2R1 – Cytochrome P450 Family 2 Subfamily R Member 1 gene

DHCR7 - 7 dehydrocholesterol reductase enzyme

DHCR7 - 7 dehydrocholesterol reductase gene

GC – Group-Specific Component

GC – Group-Specific Component gene

NADSYN1 – NAD Synthetase 1

NADSYN1 – NAD Synthetase 1 gene

PCA - Principal Component Analysis

VCF - Variant Call Format

SNP - single nucleotide Polymorphism

UV - Ultraviolet

Table of Contents:

Research Issue

The south Caucasus region is a uniquely diverse ethnic groups and varying environmental factors. In the last several years physicians in Georgia have expressed the concern regarding high prevalence of vitamin D deficiency in the population. Vitamin D plays a crucial role in regulating various physiological processes. Its synthesis in the body is largely driven by ultraviolet (UV) exposure, which triggers a complex enzymatic process involving key genes such as DHCR7, GC, NADSYN1, and CYP genes.

While DHCR7 encodes 7-dehydrocholesterol reductase, essential in converting cholesterol precursors to active vitamin D, CYP genes, such as CYP2R1 and CYP27B1, are critical for the synthesis of vitamin D. Single nucleotide polymorphisms (SNPs) within these genes have been shown to influence the efficiency of this process [3]. Given the diverse environmental factors in the Caucasus region, including varying altitudes, latitudes, and UV exposure, the interaction between genetic predispositions and environmental influences is crucial in determining vitamin D status. The genetic diversity of the Caucasus region can influence the prevalence of genetic variants, including those within critical vitamin D metabolism genes such as DHCR7, GC, NADSYN1, and CYP genes, all of which play pivotal roles in vitamin D synthesis and regulation. Variants in these genes, particularly single nucleotide polymorphisms (SNPs), may impact the efficiency of vitamin D production and its availability in the body. The central position of the study is taken by DHCR7 gene which encodes the enzyme 7-dehydrocholesterol reductase and is essential for converting cholesterol precursors into active vitamin D. While GC encodes the vitamin D-binding protein responsible for its transport, and NADSYN1 is involved in coenzyme biosynthesis, which may influence vitamin D metabolism indirectly.

Given the crucial roles of these genes in modulating vitamin D levels, understanding them geographical distribution within the Caucasus region is vital. Environmental factors such as latitude, altitude, and sunlight exposure also significantly affect vitamin D synthesis. The wide range of latitudes and altitudes in the Caucasus leads to variations in ultraviolet (UV) radiation exposure, which may interact with genetic predispositions to influence vitamin D status. Despite the apparent impact of genetic and environmental factors on vitamin D metabolism, comprehensive studies exploring the distribution of DHCR7, GC, and NADSYN1 gene variants in the Caucasus are still limited. Addressing this research gap is essential for developing personalized healthcare strategies and targeted interventions to mitigate the health risks associated with vitamin D deficiency in this genetically and environmentally diverse region.

Abstract

The DHCR7 gene, encoding the enzyme 7-dehydrocholesterol reductase, plays a crucial role in the synthesis of vitamin D. Variants within the DHCR7 gene, particularly single nucleotide polymorphisms (SNPs), have been implicated in altering the efficiency of vitamin D synthesis. This study aims to investigate the geographical distribution of DHCR7 gene variants in the South Caucasus region and assess their potential association with vitamin D metabolism.

Genetic data from individuals across different regions, including, but not exclusively, Abkhazeti, Adjara, and Tusheti, were collected using the FTDNA Family Finder platform. The dataset comprised of SNPs within the DHCR7, GC, NADSYN1,CYP2R1 and CYP27B1 genes from 95 individuals. Quality control measures were applied to ensure the integrity of the data. Principal Component Analysis (PCA) was performed to explore the genetic structure of the samples, revealing distinct clustering patterns indicating regional genetic differentiation.

Additionally, haplotype analysis highlighted several common haplotypes; however, comprehensive analysis was limited by the sample size. Future studies with larger sample sizes and detailed phenotypical data were necessary to fully understand the interactions between genetics and environment in influencing vitamin D synthesis in the South Caucasus region.