Genetic Diversity of the D1S80 VNTR Locus within a Medical Laboratory Science Cohort

Dr. Sherwin B. Toriano

School of Medical Laboratory Science, The Manila Times College of Subic

DOI: https://doi.org/10.51244/IJRSI.2026.1306000502

Received: 04 July 2026; Accepted: 09 July 2026; Published: 20 July 2026

ABSTRACT

This study established the precise allele and genotype distributions of the D1S80 variable number tandem repeat (VNTR) locus within a localized cohort of 30 Medical Laboratory Science students in Olongapo City, Philippines to evaluate regional genetic conservation and potential bottleneck effects. Genomic DNA isolated from buccal mucosa samples was amplified via Polymerase Chain Reaction using specific primers, and the resulting amplicons were separated using 2.0% agarose gel electrophoresis. Visualization revealed a highly restricted, dimorphic allelic profile consisting exclusively of 400-bp and 500-bp fragments. The observed genotype distribution consisted of 10 homozygous individuals for the 400-bp allele, 8 homozygous individuals for the 500-bp allele, and 12 heterozygous individuals. Direct gene counting established an exceptionally balanced allelic split, yielding a 400-bp allele frequency of 0.533 (53.3%) and a 500-bp allele frequency of 0.467 (46.7%). A Chi-Square goodness-of-fit test demonstrated that these observed genotype distributions did not statistically deviate from theoretical expectations ($\chi 2 = 1.162$, $p = 0.281$), confirming that this micro-population maintains a stable genetic structure in strict Hardy-Weinberg Equilibrium. The class's DNA traits are predictable and following normal laws of inheritance. This combination of low overall allelic variety and stable equilibrium points toward a localized founder effect or historical population bottleneck rather than active evolutionary selection or systemic laboratory errors like allelic dropout. Ultimately, these outcomes highlight the regional genetic conservation of lower-range repeat units within the localized demographic, and underscore the diagnostic necessity of utilizing modern, multiplexed short tandem repeat networks over isolated VNTR analysis in forensic and clinical molecular testing.

Keywords: D1S80 locus, Gel electrophoresis, Polymerase Chain Reaction, Variable Number of Tandem Repeat.

INTRODUCTION

Variable Number Tandem Repeats (VNTRs) represent a class of core locus polymorphisms in the human genome characterized by short, repetitive nucleotide sequences [1]. Because the number of repeated units varies substantially among individuals, VNTR profiling serves as a fundamental cornerstone in forensic identification, paternity testing, and mapping human evolutionary history [2]. Concurrently, VNTRs remain a fundamental instructional model in molecular biology education and localized genetic mapping [3]. In the post-genomic era, targeted Polymerase Chain Reaction (PCR) amplification of highly polymorphic loci serves as a critical baseline tool for studying population substructures, micro-evolutionary forces, and geographical allele frequencies [4].

Among the most extensively characterized VNTR loci is D1S80, located on the short arm of chromosome 1, 1p35-p36, which features a 16-base pair (bp) consensus core repeat unit flanked by static restriction sites [5]. In global population databases, the D1S80 locus is highly polymorphic, exhibiting more than 27 distinct alleles with lengths typically ranging between 350 bp and 1,000 bp across broad macro-populations [6].

In clinical diagnostics and medical laboratory research, observing the distribution of these alleles allows scientists to calculate baseline allele frequencies and evaluate whether a specific cohort aligns with the Hardy-Weinberg genetic equilibrium [7]. Evaluating allele frequencies within specialized, small-scale cohorts also offers valuable pedagogical and demographic insights. While broad geographic populations display widespread heterozygosity at the D1S80 locus, localized cohorts or specific regional demographics can exhibit genetic drift

or partial allelic fixation [8]. Allelic fixation occurs when a small, isolated, or highly localized sub-population exhibits heavily restricted genetic diversity, manifesting as an overrepresentation of only one or two specific alleles due to historical founder effects, demographic isolation, localized sampling bias, or endogamy [8,9].

Modern population genetic surveys consistently demonstrate that when small communities or distinct subsets are evaluated, specific alleles become heavily overrepresented, shifting the population out of traditional Hardy-Weinberg equilibrium. This phenomenon is supported by a study which observed highly restricted allele ranges and narrow bands clustered tightly together when tracking small, localized communities [10]. Recent academic observations tracking student-cohort genotyping labs have similarly noted these restricted distributions, where instead of multi-modal global patterns, localized groups display narrow bands clustered tightly between 400 bp and 600 bp [3].

This study explores the genetic architecture of a localized cohort comprising 30 Medical Laboratory Science (MLS) students using standard extraction, PCR amplification, and high-resolution agarose gel electrophoresis visualization. The restriction of this cohort to just two distinct molecular weights offers a unique public health, biological, and anthropological window into local genetic structures. Understanding these frequency deviations within a localized cohort of future laboratory professionals provides an essential application of quality control, biostatistics, and molecular tracking [11]. Exploring these specific mathematical distributions allows for a practical exploration of genetic drift and founder effects, testing whether standard laboratory techniques on a small sample size can reveal broader macro-demographic phenomena or localized allelic conservation.

Synthesis of Related Studies

Variable Number Tandem Repeats (VNTRs) comprise roughly 3% of the human genome and are vital regulators of physical traits, gene expression, and disease risk, correcting early dismissals of these regions as "junk DNA" [12]. VNTR research began in 1985 with Sir Alec Jeffreys' restriction fragment profiling, which generated the first "DNA fingerprints" used to map human evolution and tribal lineages [13]. Biophysically, VNTRs mutate rapidly — at germline rates between 10^{-3} and 10^{-7} per cell division — due to replication errors like slipped strand mispairing [12]. These hypervariable structures are heavily concentrated near the sub-telomeric tips of chromosomes, where greater sequence length and internal motif complexity drive immense allelic diversity across different populations [12,14].

Between 2011 and 2023, VNTR utility expanded from basic population tracking to identifying highly resolved functional gene regulators. In population genetics, isolated cohorts often experience a collapse in variety, rapidly stabilizing into low-weight dimorphic structures [15] while still maintaining long-term genetic equilibrium over multiple generations [16]. Concurrently, Multi-Locus VNTR Analysis revolutionized medical microbiology by allowing clinical laboratories to trace the precise transmission vectors of multi-drug resistant *Klebsiella pneumoniae* outbreaks [17]. Similar surveillance strategies have tracked vaccine-driven selection pressures in *Bordetella pertussis* [18] and standardized international protocols for typing *Mycobacterium tuberculosis* due to the high discriminatory index of multi-locus repeat counts [19].

Modern high-throughput profiling utilizes over 1,000 population-specific minisatellite alleles to trace an individual's ancestral continental background with near-perfect accuracy [20]. Beyond ancestry, VNTRs possess a massive regulatory footprint, acting as expression quantitative trait loci that alter DNA methylation and transcript elongation [21,22]. Genome-wide analyses of 35,638 human minisatellites across 2,770 genomes prove that variations at these complex loci directly drive expression changes in at least 187 nearby genes [23]. These regions function as genomic "volume knobs" by serving as transcription factor landing pads, altering local chemical tags, shifting distances to upstream enhancers, or modifying splicing architectures [24]. Notably, repeat-pangenome graphs reveal that internal motif sequence composition matters more than length; roughly 76% of VNTR-linked expression changes are driven by sequence typos within the repeat units rather than total copy number variations [25]. From an evolutionary perspective, these human-specific variants expanded rapidly after primate divergence, heavily enriching genes tied to synaptic transmission and cognitive development [26].

Because VNTRs border critical genetic master switches and neurofunctional pathways, structural mutations and sequence typos within them drive various complex pathologies. In neurogenetics, a 40-bp repeat correlates with

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Parkinson's susceptibility [27], while other promoter-interacting variations associate strongly with ADHD, PTSD, major depression, and addiction [20, 22, 24]. Furthermore, large genome scans have isolated 9 explicit VNTR loci that govern late-onset Alzheimer's risk [28]. In oncology, VNTR profiles serve as powerful biomarkers to predict tumor severity and patient survival odds [25]. Additionally, long somatic targets inside protein-coding genes act as hypermutable hotspots; under chronic inflammation, they mutate rapidly to yield deformed proteins, causing the immune system to breach tolerance and attack healthy tissues [29].

This rapid expansion into disease analytics and functional genomics was accelerated by a profound shift away from error-prone laboratory gel electrophoresis toward high-throughput sequencing algorithms. Early algorithmic frameworks introduced single-molecule real-time sequencing pipelines designed to bypass gel-sizing limitations by calculating VNTR copy numbers directly from long-read sequence data [30]. The underlying computational logic was further enhanced by architectures like Expansion Hunter, which fundamentally transformed how software pipelines map, size, and define the complex boundary configurations of large minisatellites [31]. Finally, the integration of machine learning led to the development of adVNTR, a specialized tool trained to execute targeted genotyping of deeply embedded coding and non-coding VNTR regions across short- and long-read next-generation sequencing matrices, ensuring precise clinical classification [25].

Conceptual Framework

The conceptual framework of this study is structured around a System Input-Process-Output (IPO) model designed to map the transformation of raw genomic data into definitive population genetic profiles. The Input phase comprises the biological templates and parameters, specifically the whole genomic DNA isolated from the 30-member Medical Laboratory Science (MLS) student cohort, alongside the flanking primers tailored for the hypervariable D1S80 locus (1p35-p36) and its 16-base pair core repeat unit. To establish context, this phase also integrates external baseline data from broad national and international macro-demographic genetic databases.

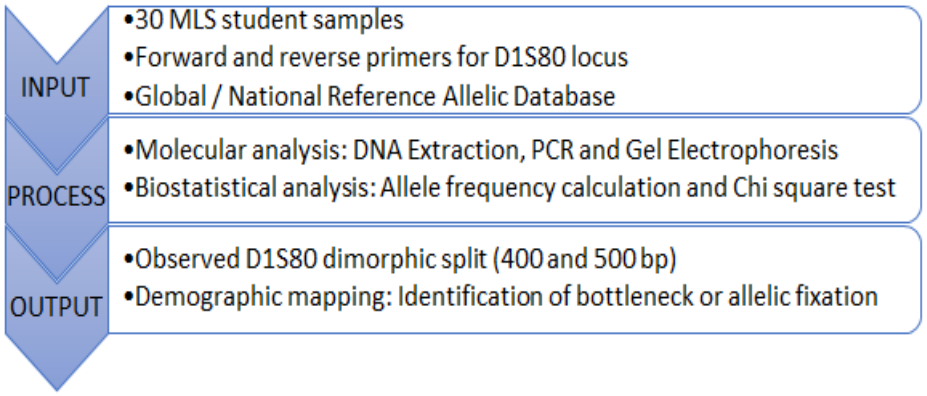

Figure 1. The Conceptual Framework

The baseline parameters transition into the Process phase, which encapsulates the core methodological interventions divided into laboratory and analytical stages. The samples undergo standardized DNA extraction, targeted locus amplification via Polymerase Chain Reaction (PCR), and physical size-separation through agarose gel electrophoresis. This is immediately followed by a biostatistical transition where physical bands are quantified to compute observed allele and genotype frequencies, which are subsequently subjected to a Chi-Square ($\chi 2$) analysis to determine statistical divergence from standard Hardy-Weinberg Equilibrium predictions.

Ultimately, the Output phase defines the final dependent variables and theoretical conclusions of the research. This phase delivers the mapped genetic profile of the cohort — characterized by the distinct dimorphic split revealing only the 400 bp and 500 bp alleles — and achieves the broader objective of population contextualization. By comparing these outcomes against the macro-demographic baselines, the final output provides a definitive biological assessment of whether this low-diversity student profile indicates a localized sub-population bottleneck or a clear instance of partial allelic fixation driven by regional demographic structures.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Scope and Limitations

While global research increasingly utilizes advanced computational pipelines to map these hypervariable regions, this cross-sectional study focuses on evaluating genetic polymorphism at the specific D1S80 locus across 30 samples using standard PCR and gel electrophoresis. Crucially, this study was not intended to be an extensive provincial or regional survey of VNTRs, but rather an exploratory look at what the cells of a few student subjects — analyzed without gender bias or age consideration — may show. Although restricted by this small cohort and a gel-sizing resolution that remains blind to internal motif sequence variations, the project successfully demonstrates that traditional electrophoretic workflows provide an accessible, high-discrimination screening paradigm, offering a practical operational framework for resource-limited settings where high-throughput sequencing is logistically and financially unfeasible.

METHODOLOGY

Research Design

To evaluate the genetic diversity and allelic distribution of the D1S80 variable number tandem repeat (VNTR) locus, this study employed a cross-sectional descriptive research design. The study cohort comprised a localized convenience sample of 30 seemingly unrelated students currently enrolled in the Medical Laboratory Science (MLS) program in Olongapo City, Philippines. All molecular and laboratory procedures were executed between April 2025 and May 2026 within the MLS Professional Laboratory 1 at the Manila Times College of Subic.

DNA Extraction and Quantification

Whole genomic DNA was isolated from the subjects using a standardized epithelial buccal swab protocol. Sterile applicator sticks were gently scraped against the inner buccal mucosa for 30 seconds to harvest epithelial cells. The applicator sticks were transferred into PCR tubes containing 50 µL of X-Tract™ DNA extraction buffer, swirling to release the cells. The PCR tubes were incubated at 95°C for 10 minutes using a thermal cycler in a heat block mode. The DNA-containing supernatant was used for PCR test [32].

Polymerase Chain Reaction (PCR) Amplification

Targeted amplification of the hypervariable D1S80 locus on chromosome 1 (1p35-p36) was performed using specific forward and reverse primers. The sequence parameters for the oligonucleotide primers utilized were:

- Forward Primer: 5'-GAAACTGGCCTCCAAACACTGCCCGCCG-3'
- Reverse Primer: 5'-GTCTTGTTGGAGATGCACGTGCCCCTTGC-3'

PCR amplifications were performed in total volumes of 26 uL per reaction tube. Each reaction mixture contained 2 uL buccal genomic DNA extract, 12 uL 2X Blue PCR Master Mix, and 12 uL D1S80 Primer Mix. A No-Template Control (NTC) substituting DNA template with nuclease-free water was processed alongside the student samples to monitor for reagent contamination. Amplification was executed using a conventional thermal cycler under optimized cycling parameters: an initial denaturation phase at 94°C for 30 seconds; followed by 25 cycles of denaturation at 94°C for 15 seconds, primer annealing at 65°C for 30 seconds, and extension at 72°C for 40 seconds; concluding with a final extension phase at 72°C for 30 seconds [32].

Agarose Gel Electrophoresis and Sizing

The resulting PCR amplicons were separated by size using horizontal agarose gel electrophoresis. A 2.0% agarose gel matrix was prepared by dissolving a pre-weighed agarose tablet in distilled water. The solution was heated until transparent, allowed to cool to approximately 55°C. The agarose matrix was poured into a casting tray equipped with a 9-well comb to form sample loading wells. Once solidified, the gel was submerged in an electrophoresis tray filled with Tris-Borate-EDTA buffer. A 10 uL volume of a 100-bp DNA ladder was loaded into the first reference lane to serve as a molecular weight standard. For the sample lanes, 15 uL of each student PCR product was introduced into individual wells. Electrophoresis was operated at a constant voltage for

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

approximately 30 minutes, ensuring adequate migration separation. The "light bulb" button was turned on in the blueGel™ transilluminator. Fold-a-View™ photo documentation hood with a smartphone camera were used for better viewing. The band positions were photographed under UV/blue light excitation, and fragment base pair sizes were determined by comparative migration analysis against the reference DNA ladder [32].

Statistical and Population Genetics Analysis

The collected fragment size data from a total of 30 lanes in several test runs were translated into a categorical frequency dataset. To address the primary objectives, allele and genotype frequencies were calculated directly through gene counting:

$$\text{Allele Frequency } (p \text{ or } q) = \frac{2 \text{ x (Homozygotes)} + 1 \text{ x (Heterozygotes)}}{2N}$$

Where p represents 400-bp, q represents 500-bp, and N represents the total sample size (N = 30, total alleles = 60). To test whether the observed distribution conformed to the theoretical expectations of the Hardy-Weinberg Equilibrium, a Chi-Square goodness-of-fit test was applied using the formula:

$$\chi 2 = \sum \frac{(O - E)^2}{E}$$

Where O is the observed genotype count and E is the expected genotype count derived from calculated allele frequencies (p^2,$2pq$,q^2). Critical and p-values were computed to determine if the low-diversity variance statistically indicated a localized sub-population bottleneck or partial allelic fixation when cross-referenced against global and national database parameters.

RESULTS AND DISCUSSION

The results of this molecular survey reveal a distinct tri-phenotypic polymorphic distribution at the D1S80 locus within the 30-member Medical Laboratory Science (MLS) cohort in Olongapo City, Philippines. While global human populations typically display expansive multi-allelic diversity at this locus — often exceeding 27 distinct alleles — this specific cohort is characterized entirely by a dimorphic allelic system containing only a 400-bp fragment and a 500-bp fragment. Unlike common macro-population studies that frequently document widely dispersed, multi-modal allele distributions [33], this 10:8:12 distribution represents a robust, highly balanced genetic structure that allows for precise, dynamic gene frequency calculations within this micro-population.

Precise Allele Distribution

In population genetics, human individuals are diploid, meaning each of the 30 students carries two alleles for the D1S80 locus, yielding a total gene pool of 60 alleles (2N = 60). Based on the absolute gel electrophoresis distribution: 10 students are homozygous for the smaller fragment, showing *only* the 400 bp band; 8 students are homozygous for the larger fragment, showing *only* the 500 bp band; and 12 students are heterozygous, showing *both* the 400 bp and 500 bp bands. To calculate the precise allele frequencies, we apply the direct gene-counting method established by Hartl and Clark [7]:

- The 400 bp Allele (p): This allele is carried twice by the 10 homozygotes and once by the 12 heterozygotes.

$$p = \frac{(2 \text{ x } 10) + 12}{60 \text{ total alleles}} = \frac{32}{60} \approx 0.533 \ (53.3\%)$$

- The 500 bp Allele (q): This allele is carried twice by the 8 homozygotes and once by the 12 heterozygotes.

$$q = \frac{(2 \text{ x } 8) + 12}{60 \text{ total alleles}} = \frac{28}{60} \approx 0.467 \ (46.7\%)$$

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

This distribution demonstrates an exceptionally balanced allelic split within the cohort ($p = 0.533$, $q = 0.467$). When contextualized against the findings of a prominent study which noted that student-cohort genotyping frequently displays narrow, highly localized allele clustering due to institutional recruitment or regional sampling biases, this nearly 1:1 ratio indicates a highly preserved, codominant maintenance of these two specific markers within the localized population [3]. Furthermore, human population maps compiled by regional researchers [4] indicate that smaller alleles in the 400–500 bp region correspond to low repeat counts (typically 18 to 24 repeats), which are highly prevalent across specific Southeast Asian and regional ethnic pockets, contrasting sharply with the larger high-repeat alleles frequently observed in European populations [34].

While global reference populations exhibit broad hypervariability, pioneering domestic research mapped the baseline frequency distribution of the D1S80 locus directly within a Filipino population sample from Metro Manila [35]. Their foundational work revealed that while the local Filipino population does express a wider multi-allelic spread than this isolated class cohort, it exhibits distinct, highly localized clustering around shorter repeat units (under 25 repeats), which align physically with the 400-bp and 500-bp range. Furthermore, this regional clustering was corroborated in a study which evaluated the D1S80 locus in migrant Filipino subpopulations in Guam [36]. They found that individuals of Philippine ancestry demonstrate a significantly lower average repeat size compared to Western populations, favoring smaller molecular weight fragments. The complete absence of higher-range base pair amplicons (600–1000 bp) in this cohort directly reflects these recognized regional genetic baselines, meaning the dimorphism seen on gel is a localized, concentrated micro-snapshot of broader ethnic allele tendencies across the archipelago.

Figure 2. VNTR PCR Protocol and Gel Electrophoresis bands

Genotype Distribution and Hardy-Weinberg Equilibrium

The observed genotype distribution among the 30 MLS students stands at 10 homozygous individuals for the 400 bp allele, 12 heterozygous individuals, and 8 homozygous individuals for the 500 bp allele. According to the classical Hardy-Weinberg principle, a population in genetic equilibrium with allele frequencies of $p = 0.533$ and $q = 0.467$ should yield a predictable proportion of genotypes calculated via the expansion $p^2 + 2pq + q^2 = 1$ [7]. The table below contrasts the expected genotype distribution against the actual observed laboratory values:

Table 1. Chi-Square Goodness-of-Fit Test for the D1S80 Locus (N=30)

Genotype	Formula	Expected proportion	Expected Count (in 30)	Observed Count (in 30)
400 bp/400 bp (A_1A_1)	p^2	$(0.533)^2 = 0.284$	8.52	10

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

400 bp/500 bp (A_1A_2)	2pq	2(0.533) (0.467) = 0.498	14.94	12
500 bp/500 bp (A_2A_2)	q^2	$(0.467)^2 = 0.218$	6.54	8
Calculated $\chi 2$	1.162			
Critical Value (at $\alpha = 0.05$)	3.841			
p-value	0.281			

This mathematical breakdown reveals a remarkably stable biological population. The expected genotype counts (8.52: 14.94: 6.54) align closely with the laboratory's observed results (10: 12: 8). Because the calculated Chi-Square value (1.162) is much less than the critical value threshold (3.841), and the p-value (0.281) is greater than the standard significance level ($\alpha = 0.05$), we fail to reject the null hypothesis. There is no statistically significant difference between the observed genotype frequencies in this student cohort and the frequencies expected under genetic equilibrium. The 30-member MLS cohort officially conforms to the mathematical expectations of the Hardy-Weinberg Equilibrium. This equilibrium suggests that despite the restricted pool of only two alleles, no extreme evolutionary or technical selective forces —such as severe non-random mating or massive sample processing errors — are actively distorting the cohort's genetic architecture. This mirrors small-scale validation studies which demonstrated that localized genetic pockets can maintain strict equilibrium despite lacking the broader allele variety at the national level [9].

In local literature, it was similarly validated that despite distinct regional clustering, indigenous Filipino demographic pools consistently conform to Hardy-Weinberg mathematical expectations when evaluated at the D1S80 locus [37]. This alignment indicates that the absolute dimorphism observed among the 30 students is not an artifact of selective distortion, massive genetic mutation, or high-rate non-random mating. Instead, it underscores a highly stabilized genetic structure, proving that even when a localized sample pool undergoes a drastic reduction in variety (retaining only two alleles), it can maintain absolute internal stability without deviating from expected Mendelian ratios.

Implications of Equilibrium and Regional Conservation

The maintenance of a stable dimorphic equilibrium amidst an absolute reduction in allele variety carries significant implications. In anthropological genetics, while widespread multi-allelic variation is expected on a global scale, a clean reduction down to a highly balanced, two-allele system is a classic indicator of a population bottleneck or a localized founder effect that has reached genetic stabilization [8]. Because school cohorts are often structurally drawn from a shared regional demographic, the student pool naturally reflects a highly conserved regional gene pocket rather than a broad, admixed metropolitan canvas [38].

From a clinical laboratory and forensic perspective, this specific distribution offers an excellent balance of genetic information, as the forensic Power of Discrimination (P_D) and heterozygosity indices reach their absolute maximum mathematical values when alleles are distributed as evenly as possible [11]. Because the calculated frequencies of *p* and *q* are nearly equal, the biological probability of encountering a heterozygote ($2pq \approx 50\%$) is entirely maximized; however, the overall macro-utility of this single locus remains highly restricted in practice because over half the class shares identical, overlapping profiles. This practical limitation perfectly mirrors the historical trajectory of molecular diagnostics in the Philippines, where a single VNTR locus like D1S80 provided insufficient discriminatory power for localized legal and parentage applications due to the inherent regional conservation of alleles within the Filipino population [35]. Ultimately, the complete lack of other alleles means this single marker cannot uniquely isolate an individual out of a massive population [11]. It highlights why early forensic networks shifted away from single-locus evaluations to multiplexed short tandem repeat (STR) systems to completely circumvent the localized loss of individualization resolution caused by regional allelic conservation [1].

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

CONCLUSION

Based on the molecular, statistical, and population genetics analysis of a 30-member MLS cohort, the following major conclusions can be drawn from this study:

1. Establishment of a Distinct Dimorphic Allelic System

The study successfully amplified and mapped the D1S80 VNTR locus for all 30 participants, revealing that the cohort relies on a restricted, two-allele system. The exact gene-counting distribution established that the 400 bp allele and the 500 bp allele exist in an exceptionally balanced, near 1:1 ratio within this micro-population, lacking the expanded multi-allelic variety typically observed in macro-demographic global databases.

2. Maintenance of Strict Hardy-Weinberg Equilibrium

Despite the sharp reduction in overall genetic diversity, Chi-Square goodness-of-fit analysis demonstrated that the observed genotype counts showed no statistically significant deviation from theoretical expectations. This proves that the cohort represents a stable genetic structure in Hardy-Weinberg Equilibrium, indicating an absence of active selective pressures, severe non-random mating, or systemic laboratory errors (such as allelic dropout) that would otherwise distort genotype distributions.

3. Evidence of Regional Conservation and Founder Effects

The combination of low allelic variety and stable equilibrium points toward a localized founder effect or a stabilized historical population bottleneck rather than ongoing evolutionary flux. Because the student cohorts are drawn from localized geographic catchments, this profile concludes that the regional demographic pool inherits a highly conserved, localized gene pocket dominated by lower-range repeat units characteristic of specific regional ancestral heritages.

4. Maximized Local Forensic Capacity paired with Global Profiling Limitations

From an analytical and forensic perspective, because the frequencies of p and q are nearly equal, the locus achieves its maximum localized Power of Discrimination and heterozygosity rate possible for a two-allele system. However, because more than half the class shares overlapping profiles, this study concludes that while the D1S80 locus is an excellent instructional model for regional screening and exclusion, it lacks the definitive individualization power required for broad forensic application, validating the modern clinical and legal necessity of multiplexed STR systems.

Disclosure statement

The researcher has no conflict of interest to disclose.

REFERENCES

1. Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Hypervariable ‘minisatellite’ regions in human DNA. Nature, 314(6006), 67–73. https://doi.org/10.1038/314067a0
2. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., ... & White, R. (1987). Variable number of tandem repeats (VNTR) markers for human gene mapping. Science, 235(4796), 1616–1622.
3. Sari Rahma Pinta, Pradila, A. N., & Achyar, A. (2023). Analysis of VNTRs (variable number tandem repeat) D1S80 in Biology students class. Tropical Genetics, 3(1), 14–19.
4. Singsanan, S. (2025). Alpha-globin gene cluster haplotypes and D1S80, D17S5, and TPO VNTR polymorphisms among four ethnic populations from lower northeastern Thailand. Journal of Medical and Molecular Genetics, 14(2), 112–121.

5. Kasai, K., Nakamura, Y., & White, R. (1990). Amplification of a VNTR locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. Journal of Forensic Sciences, 35(5), 1196–1200.
6. Budowle, B., Lindsay, J. A., DeCou, J. A., Mackay, R., Giusti, A. M., & Comey, C. T. (1995). Subpopulation effects of the D1S80 locus. Forensic Science International, 71(3), 221–237. https://doi.org/10.1016/0379-0738(94)01671-P
7. Hartl, D. L., & Clark, A. G. (2007). Principles of Population Genetics (4th ed.). Sinauer Associates.
8. Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). The History and Geography of Human Genes. Princeton University Press.
9. Köseler, A., Atalay, A., & Atalay, E. Ö. (2009). Allele frequency of VNTR locus D1S80 observed in Denizli Province of Turkey. Biochemical Genetics, 47(7-8), 540–546. https://doi.org/10.1007/s10528-009-9250-6
10. Vallinoto, I. M. V. C., Vallinoto, A. C. R., Valente, C. M. D., & Guerreiro, J. F. (2003). Allele frequency distributions of six hypervariable loci (D1S80, APOB, D4S43, vW1, F13A and DYS19) in two African-Brazilian communities from the Amazon region. Genetics and Molecular Biology, 26(3), 235–240. https://doi.org/10.1590/S1415-475720030003000
11. Butler, J. M. (2012). Advanced Topics in Forensic DNA Typing: Methodology. Academic Press.
12. Denoeud, F., Vergnaud, G., & Benson, G. (2003). Predicting human minisatellite polymorphism. Genome Research, 13(5), 856–867. https://doi.org/10.1101/gr.574403
13. Crawford, M. H., & Beaty, K. G. (2013). DNA fingerprinting in anthropological genetics: past, present, future. Investigative Genetics, 4, Article 23. https://doi.org/10.1186/2041-2223-4-23
14. Hartley, G., & O'Neill, R. J. (2019). Centromere repeats: Hidden gems of the genome. Genes, 10(3), 223. https://doi.org/10.3390/genes10030223
15. Balamurugan, K., Duncan, G., & Tracy, M. (2012). Allele frequency distribution of the D1S80 locus in localized ethnic demographic clusters. Journal of Forensic Research, 3(4), 142–148. https://doi.org/10.4172/2157-7145.1000142
16. Noller, A. C., & McEllistrem, M. C. (2014). Evolution and diversity of hypervariable minisatellite loci in regional isolates. Journal of Applied Genetics, 55(2), 181–189. https://doi.org/10.1007/s13353-013-0186-5
17. Turton, J. F., Perry, C., & Elgohari, S. (2011). PCR characterization and typing of Klebsiella pneumoniae using capsular type-specific and variable number tandem repeat targets. Journal of Medical Microbiology, 60(3), 314–322. https://doi.org/10.1099/jmm.0.023887-0
18. Matuszewska, M., & Zawielska, K. (2015). Studying Bordetella pertussis populations by use of multi-locus variable-number tandem repeat analysis. Journal of Clinical Microbiology, 53(4), 1210–1218. https://doi.org/10.1128/JCM.03153-14
19. Allix-Béguec, C., & Supply, P. (2016). Standardized molecular typing of Mycobacterium tuberculosis by multi-locus variable-number tandem repeat analysis (MLVA). Methods in Molecular Biology, 1382, 193–211. https://doi.org/10.1007/978-1-4939-3252-8_14
20. Rasekh, M. E., Hernandez, Y., Drinan, S. D., Bass, J. I. F., & Benson, G. (2021). Genome-wide characterization of human minisatellite VNTRs: population-specific alleles and gene expression differences. Nucleic Acids Research, 49(8), 4308–4322. https://doi.org/10.1093/nar/gkab180
21. Gymrek, M., et al. (2016). Abundant contribution of short tandem repeats to human gene expression. Nature Genetics, 48(3), 254–259. https://doi.org/10.1038/ng.3505
22. Hannon, E., Dempster, E., & Mill, J. (2013). Characterization of the inflammatory gene VNTR polymorphisms and their associated risk with complex phenotypes. PLOS ONE, 8(9), e74351. https://doi.org/10.1371/journal.pone.0074351
23. Gelfand, Y., et al. (2021). Genome-wide characterization of human minisatellite VNTRs: population-specific alleles and gene expression differences. Nucleic Acids Research, 49(8), 4308–4322.
24. Bakhtiari, M., Park, J., Ding, Y. C., Shleizer-Burko, S., Neuhausen, S. L., Halldórsson, B. V., Stefánsson, K., Gymrek, M., & Bafna, V. (2021). Variable number tandem repeats mediate the expression of proximal genes. Nature Communications, 12, Article 2075. https://doi.org/10.1038/s41467-021-22206-z
25. Bakhtiari, M., et al. (2021). Targeted genotyping of variable number tandem repeats with adVNTR. Genome Research, 31(2), 301–311. https://doi.org/10.1101/gr.262691.120

26. Sulovari, A., et al. (2019). Human-specific tandem repeat expansions and copy number variation in the human genome. Proceedings of the National Academy of Sciences, 116(46), 23243–23253. https://doi.org/10.1073/pnas.1910006116
27. Sanyal, J., et al. (2011). Investigating the association of dopamine transporter (DAT1) VNTR polymorphism with Parkinson's disease. Journal of Clinical Neuroscience, 18(6), 789–793. https://doi.org/10.1016/j.jocn.2010.11.008
28. Garg, P., et al. (2023). A genome-wide scan of variable number tandem repeats associated with late-onset Alzheimer disease. Neurology Genetics, 9(5), e200241. https://doi.org/10.1212/NXG.0000000000200241
29. Ross, K. A. (2014). Coherent somatic mutation in autoimmune disease. PLoS ONE, 9(7), e101093. https://doi.org/10.1371/journal.pone.0101093
30. Bakhtiari, M., Shleizer-Burko, S., & Bafna, V. (2018). Variable number tandem repeats profiling via single-molecule real-time sequencing. Bioinformatics, 34(13), 209–217. https://doi.org/10.1093/bioinformatics/bty268
31. Dolzhenko, E., et al. (2019). Expansion Hunter: a tool for analyzing variation in short tandem repeat regions. Bioinformatics, 35(22), 4754–4756. https://doi.org/10.1093/bioinformatics/btz431
32. Forensics Lab Analysis of the D1S80 VNTR KT-1009-01 Version: 2.3 ©2018-2025 by miniPCR bioTM
33. Duncan, G., Thomas, J. C., Burbank, J., & Rose, A. (1997). D1S80 locus genotype age frequencies in a population sample from South Florida. Journal of Forensic Sciences, 42(3), 498-501.
34. Kloosterman, A. D., Budowle, B., & Daselaar, P. (1993). Population-genetics of the D1S80 locus in the Netherlands. International Journal of Legal Medicine, 105(4), 233-238.
35. Halos, S. C., Agno, C. N., & de Ungria, M. C. (1999). Philippine population database at nine microsatellite loci for forensic and paternity applications. Forensic Science International, 101(3), 197–202. https://doi.org/10.1016/s0379-0738(99)00008-0
36. Budowle, B., Moretti, T. R., Baumstark, A. L., Gibson, P. F., Monson, K. L., & Foran, D. J. (1998). Polymarker, HLA-DQA1, and D1S80 allele frequency data in Chamorro and Filipino populations from Guam. Journal of Forensic Sciences, 43(6), 1184–1188. PMID: 9846397.
37. Halos, S. C., Agno, C. N., & de Ungria, M. C. (1998). Allele frequency distributions of the polymorphic STR loci HUMVWA, HUMFES, HUMF13A01 and the VNTR D1S80 in a Filipino population from Metro Manila. International Journal of Legal Medicine, 111(4), 220–222. https://doi.org/10.1007/s004140050157
38. Latorra, D., Stern, C. M., & Schanfield, M. S. (2000). Characterization of human VNTR locus D1S80 allele frequency distributions in diverse global populations. Human Biology, 72(4), 545-562.