



Hypothesis

Perspective for CAR T-Cell Therapy in Underrepresented Populations: A Hypothesis-Generating CD19 Genomic Analysis

Maysa Al-Hussaini , Anas Al Okaily and Osama Alsmadi *

Department of Cell Therapy and Applied Genomics, King Hussein Cancer Center, Amman 11941, Jordan; mhussaini@khcc.jo (M.A.-H.); aa.12682@khcc.jo (A.A.O.)

* Correspondence: oa.12163@khcc.jo; Tel.: +962-770336330

Abstract

CD19-directed chimeric antigen receptor (CAR) T-cell therapy has fundamentally transformed the treatment landscape for relapsed and refractory B-cell malignancies, yet antigen escape remains a persistent therapeutic challenge that limits long-term remission durability. While antigen loss is typically considered a somatic event acquired during tumor evolution under therapeutic selective pressure, germline CD19 polymorphisms could theoretically influence CAR-binding kinetics, alter epitope presentation, and modulate therapeutic outcomes in ways that remain largely not characterized. Unfortunately, Middle Eastern populations are underrepresented in pharmacogenomic databases and CAR-T clinical trials, creating a knowledge gap that may perpetuate global health disparities in access to precision immunotherapy. We analyzed publicly available whole-exome sequencing data from 1196 individuals of Arab origin to comprehensively characterize CD19 variants with potential relevance to CAR T-cell immunotherapy. The L174V (rs2904880) variant stood out, and showed the Valine/Valine (V/V) genotype frequency was 65.3%, corresponding to a V174 allelic frequency of 76.6%, while the minor allele, L174, has a frequency of 23.4%. The missense mutation (c.520C > G) responsible for this variant results in a leucine-to-valine (L174V) substitution at position 174 of the CD19 protein, relative to the reference genome. The cohort genotypes (CC, CG, and GG) exhibited a significant deviation from Hardy–Weinberg equilibrium ($p < 0.00001$). While this deviation is consistent with the high consanguinity rates (25–60%) amongst Arab populations, it remains not fully explained, and may be attributed to population structure, relatedness, or technical factors. We further emphasize that our computational analysis cannot establish any direct clinical or functional impact due to this variant, and therefore we refrain from suggesting any specific actions at the current time. In light of these findings, we hypothesize that the distinctive genetic architecture of consanguineous populations should not be viewed as a confounding variable. Instead, it presents a unique opportunity to investigate the clinical relevance of germline variation in the context of precision oncology, particularly at therapy-relevant loci, pending functional validation.



Academic Editor: Jorge Luis Espinoza

Received: 7 April 2026

Revised: 17 June 2026

Accepted: 18 June 2026

Published: 25 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](#).

Keywords: CAR T-cell therapy; CD19; pharmacogenomics; population genetics; precision medicine; Middle East; health disparities; consanguinity; founder effect; antigen escape

1. Introduction

Chimeric Antigen Receptor T-cell (CAR-T) therapy is a transformative advance in oncology, offering durable remissions for relapsed/refractory B-cell malignancies [1]. Second-generation CD19-targeting CARs (tisagenlecleucel, axicabtagene ciloleucel) have demon-

strated remarkable efficacy in B-ALL and DLBCL [2,3]. They achieve 80–90% complete remission rates in pediatric/young adult B-ALL [4].

Mechanistically, autologous T cells are harvested and genetically modified to express a chimeric receptor with an extracellular scFv recognizing CD19, a CD3 ζ signaling domain, and costimulatory domains (CD28 or 4-1BB) that enhance persistence [5]. Upon reinfusion, they proliferate exponentially and kill CD19+ B cells via perforin/granzyme and cytokine pathways [6].

Despite these successes, 30–50% of patients relapse, with CD19 modulation accounting for many resistance cases [7]. Antigen escape occurs via multiple mechanisms: transcriptional downregulation, alternative splicing, lineage switching, and, most relevant here, missense mutations disrupting the CAR binding epitope [8,9].

Most FDA-approved CAR-T products use the FMC63 scFv, which recognizes a specific epitope within the CD19 extracellular domain [10]. Single amino acid substitutions at this interface can reduce or abolish CAR-T recognition, favoring tumor clones with such mutations. Clinical studies have documented acquired somatic CD19 mutations in patients relapsing after FMC63-based therapy, including L174 mutations that impair binding [11,12].

1.1. The L174V Polymorphism: A Germline Variant in the CAR Binding Domain

The L174V variant (rs2904880) is a leucine-to-valine substitution at position 174 of CD19, within the minimal recognition domain of the FMC63 scFv [10]. Structural and functional studies suggest it may influence binding kinetics, but clinical significance remains undefined [13,14].

While somatic L174 mutations are acquired escape mutations post-CAR-T, the L174V variant also exists as a germline polymorphism in the general population [15]. Germline variants are present from birth and may affect CD19 structure/density across all B cells, whereas somatic mutations arise in malignant clones under selective pressure. The clinical significance of germline L174V for CAR-T outcomes remains unexplored, representing today a critical knowledge gap.

Middle Eastern populations are underrepresented in major genomic databases and CAR-T trials. In gnomAD v4.1, Middle Eastern individuals constitute <1% of samples despite representing ~5% of the global population [16]. This disparity is even more pronounced in pharmacogenomic studies examining therapy-relevant loci, where Arab individuals are frequently absent.

This gap matters due to unique demographics: Arab populations have high consanguinity rates (25–60%) [17], leading to pronounced founder effects where specific variants become concentrated [17]. The Qatar Genome Program for example found that 62.5% of Qatari participants carry at least one recessive pathogenic variant, with homozygosity 5.8-fold higher than in non-Arab ancestries [18]. Thus, variants rare in Europeans may be common and clinically significant in Arab populations, and vice versa.

1.2. Hypothesis and Study Rationale

We hypothesize that Arab populations, with high consanguinity and founder effects, may exhibit distinct allele frequencies and homozygosity patterns for the CD19 L174V polymorphism compared to European and East Asian populations that dominate genomic databases and CAR-T trials. Additionally, consanguinity-driven elevated homozygosity could amplify functional effects of homozygous variant carriage, making Arab cohorts uniquely informative for studying germline variation at this locus.

This study addresses this gap by characterizing CD19 variation in a large, publicly available Arab cohort, with the following objectives: (1) determining allele frequencies and genotype distributions for all non-synonymous CD19 variants, (2) comparing these

frequencies to global reference populations, (3) assessing Hardy–Weinberg equilibrium deviation as a reflection of population structure, and (4) generating testable hypotheses on the potential impact of germline L174V homozygosity on CAR-T therapy outcomes.

2. Materials and Methods

2.1. Data Source and Cohort Description

We analyzed publicly available whole-exome sequencing data from individuals of Arab descent, predominantly Qatari nationals, accessed through the National Center for Biotechnology Information Sequence Read Archive (SRA) under accessions SRP060765, SRP061943, and SRP061463. These datasets were originally generated as part of Qatar Genome Program pilot studies and represent one of the largest publicly available genomic resources for an underrepresented population [19].

The final analytic cohort consisted of 1196 individuals. All included individuals were self-identified as having Arab family origins. The cohort was approximately balanced for gender (52% male, 48% female) with ages ranging from 18 to 85 years. Individuals with known malignancies, including hematological disorders, were excluded in these original studies, making this cohort representative of the general population for the purpose of hematological disorders. The original studies received ethical approval from the Qatar Biomedical Research Institute Institutional Review Board, and all participants provided written informed consent for genomic research and public data deposition.

2.2. Quality Control and Ancestry Inference

All samples underwent rigorous quality control prior to analysis. Variants were filtered to retain only those with read depth ≥ 30 , genotype quality ≥ 30 , and missingness $< 5\%$. Relatedness analysis was conducted by the authors [19] using KING; however, for the purpose of this report, this is a limitation that we acknowledge. For downstream statistical analysis, the final cohort comprises 1196 individuals.

Beginning with SAM files (imported from the Qatari Genome Project) as input, the workflow systematically removed group headers and filtered alignments by excluding unmapped reads, retaining only concordant (properly paired) read pairs, and selecting the identified unique mapping tags to eliminate ambiguous multi-mapping events. The filtered alignments were subsequently converted to BAM format using SAMtools (v1.21), sorted by genomic coordinates, and conditionally modified to add “chr” prefixes to chromosome identifiers for standardized genome reference compatibility, with SAMtools also employed for BAM indexing to enable efficient random access during variant calling.

The pipeline then completed the variant calling using the Pisces software (v5.2.11.163), which analyzes the curated high-quality alignments to generate a VCF file containing identified raw genomic variants. No post-calling filters were applied to the variants in the generated VCF file, as these represent germline variants rather than somatic mutations. Lastly, the annotation was performed using Variant Studio V3 to functionally characterize the identified variants and assess their potential biological significance [20].

2.3. Population Comparisons

To ensure comparability with reference databases, we applied identical quality filters to our data and to public data extracts: depth ≥ 30 , genotype quality ≥ 30 , and missingness $< 5\%$. We acknowledge that differences in sequencing technology (Illumina HiSeq for our data vs. varied platforms for reference databases) and variant calling algorithms (Pisces v5.2 for our data vs. GATK for gnomAD) may introduce bias. Therefore, we present comparisons as descriptive enrichment ratios rather than statistically precise differences.

For population frequency comparisons, we utilized publicly available data from gnomAD v4.1 (genome aggregation database, containing 730,947 exomes and 76,215 genomes across global populations), the 1000 Genomes Project Phase 3 (2504 individuals from 26 populations), and ExAC (Exome Aggregation Consortium, 60,706 individuals) [16,21–23]. We specifically extracted allele frequencies for CD19 variants in the following super-populations: European (non-Finnish), African/African American, Latino/Admixed American, East Asian, South Asian, and other (including Middle Eastern where available).

2.4. Hardy–Weinberg Equilibrium and Inbreeding Analysis

Genotype frequencies for the CD19 variant (C/C, C/G, G/G) were tallied from observed counts. Allele frequencies were calculated by direct gene allele counts. Expected genotype counts under Hardy–Weinberg equilibrium (HWE) were derived using the standard equations: $E_{LL} = p^2 \times N$, $E_{LV} = 2pq \times N$, and $E_{VV} = q^2 \times N$, where p is the frequency of the L (C) allele, q is the frequency of the V (G) allele, and N is the total number of individuals. A chi-square goodness-of-fit test with one degree of freedom was applied to compare observed and expected genotype counts, with statistical significance set at $p \leq 0.05$.

To assess whether any observed deviation from HWE could be attributed to consanguinity (non-random mating), the inbreeding coefficient F was estimated as $F = 1 - (H_{obs}/H_{exp})$, where H_{obs} is the observed heterozygosity and $H_{exp} = 2pq$ is the expected heterozygosity under HWE. Expected genotype counts under inbreeding were then recalculated using the adjusted equations: $LL = p^2 + pqF$, $LV = 2pq(1 - F)$, and $VV = q^2 + pqF$. For the R515H variant (rs34763945), 95% confidence intervals for allele frequencies were calculated using the Wilson score method and are provided in Table 1.

Table 1. Observed and expected genotype counts for CD19 L174V (rs2904880) under Hardy–Weinberg equilibrium (HWE) and after adjustment for consanguinity (inbreeding coefficient $F = 0.367$)⁴.

Genotype	Observed (n [%]) ¹	Expected Under HWE (n [%]) ^{2,3}	Expected Under HWE + Consanguinity (n [%]) ⁵
L/L (C/C)	144 [12.04%]	65.32 [5.46%]	144 [12.04%]
L/V (C/G)	271 [22.66%]	428.36 [35.82%]	271 [22.66%]
V/V (G/G)	781 [65.30%]	702.32 [58.72%]	781 [65.30%]
Total	1196 [100%]	1196.00 [100%]	1196 [100%]

¹ Allele frequencies: $p = \text{freq}(L) = 0.2337$; $q = \text{freq}(V) = 0.7663$. ² HWE expectations: $E_{LL} = p^2N$; $E_{LV} = 2pqN$; $E_{VV} = q^2N$.

³ Chi-square test for HWE ($df = 1$): $\chi^2 = 161.4$; $p < 0.00001$. ⁴ Inbreeding coefficient: $F = 1 - (H_{obs}/H_{exp}) = 0.3674$.

⁵ Consanguinity-adjusted expectations: $LL = p^2 + pqF$; $LV = 2pq(1 - F)$; $VV = q^2 + pqF$.

3. Results

Analysis of the CD19 coding sequence in 1196 Arab individuals identified 21 non-synonymous variants (Supplementary Table S1) distributed across the extracellular domain, transmembrane domain, and cytoplasmic tail. Of these, 17 were missense variants and 4 were nonsense variants predicted to result in premature truncation. Notably, all variants except c.520C > G (L174V) were observed exclusively in the heterozygous state, with no homozygous alternate individuals identified for any of the remaining 20 variants (Supplementary Table S1). The median variant allele frequency across all non-synonymous variants was 0.8% (range 0.04% to 23.4% for the minor allele of rs2904880), consistent with the expected distribution of rare and common variation.

Two variants of potential clinical relevance were prioritized for detailed analysis based on their location within functional domains and prior literature implicating them in CAR-T resistance mechanisms. L174V (rs2904880) within the FMC63 scFv binding domain and

R515H (rs34763945) within the transmembrane domain have been shown to affect CD19 surface expression and internalization dynamics [24,25].

3.1. The L174V Variant (rs2904880)

The L174V variant exhibited a V174 allele frequency of 76.6% in our Arab cohort; the minor allele (L174) frequency was 23.4%. Throughout this manuscript, L174 refers to the reference allele (leucine) and V174 refers to the variant allele (valine). This V174 allele frequency is notably higher than in European populations (68.6% in non-Finnish Europeans, representing a 1.11-fold enrichment) and comparable to frequencies observed in South Asian (83.2%) and African (94.0%) populations. Regarding reference allele designation: In gnomAD v4.1, the reference allele is C (encoding Leucine, L174); in 1000 Genomes Phase 3, the reference allele is also C; in ExAC, the reference allele is C. V174 (G allele) is consistently the alternate allele across all three databases.

The genotype distribution was: L/L (homozygous reference, leucine/leucine) = 144 individuals (12.0%); L/V (heterozygous) = 271 individuals (22.7%); and V/V (homozygous genotype, valine/valine) = 781 individuals (65.3%) (Table 1). The V/V rate of 65.3% attributed to homozygosity (GG) in our Arab cohort is 1.24-fold higher than in non-Finnish Europeans (52.5% homozygosity for V/V) and 1.80-fold higher than in Latino/Admixed American populations (36.3%) as reported by the 1000 Genomes Project [21].

Hardy–Weinberg equilibrium analysis for rs2904880 revealed highly significant deviation from expected genotype counts ($\chi^2 = 161.4$, $df = 1$, $p < 0.00001$), characterized by a marked heterozygote deficit (observed 271 [22.66%] vs. expected 428.36 [35.82%]) and a reciprocal homozygote excess. This pattern is consistent with non-random mating. Estimation of the inbreeding coefficient yielded $F = 0.367$, suggests a high rate of consanguinity. After adjusting expected genotype counts using this F value, the observed counts were reproduced nearly exactly (Table 1, column 4). Despite the observed statistical significance, this HWE deviation remains not fully explained; the contribution of other factors such population structure, relatedness, and technical factors including duplicate samples cannot be ruled out, and that requires further computing verifications. The original data source (Qatar Genome Project) group applied the standard data quality checks of their own, however we cannot independently verify these at the current time. We further presented the inbreeding coefficient ($F = 0.367$) as a descriptive statistic rather than a definitive estimate of consanguinity.

3.2. The R515H Variant (rs34763945)

The R515H variant (rs34763945), located within the CD19 transmembrane domain, was observed at a MAF of 1.34% in our Arab cohort (Table 2).

This frequency is significantly lower than the European frequency of 6.7% and similar to frequencies observed in Amish populations (1.54%). The R515H substitution changes a positively charged arginine to a neutral histidine within the transmembrane domain, hypothetically affecting CD19 membrane orientation, homodimerization, or association with signaling partners such as CD21 and CD81 [26–29].

The observation that population-specific variation is bidirectional, with L174V more common in Arabs than Europeans but R515H less common, demonstrates that genetic differences at therapy-relevant loci cannot be predicted from global averages or from any single reference population. This underscores the necessity of direct empirical characterization of understudied populations rather than relying on imputation/extrapolation from centric databases alone.

Table 2. rs34763945 allelic frequencies in the Arab cohort compared to other populations with 95% confidence intervals and sample denominators.

Genetic Ancestry Group	Allele Count	Allele Number	Sample Denominator (Individuals) *	Homozygotes	Allele Frequency (%)	95% CI (Lower)	95% CI (Upper)
European (non-Finnish)	78,926	1,179,742	589,871	2747	6.69%	6.65%	6.73%
European (Finnish)	4120	63,982	31,991	126	6.44%	6.25%	6.63%
South Asian	2084	91,074	45,537	45	2.29%	2.19%	2.39%
Admixed American	1327	59,990	29,995	13	2.21%	2.10%	2.33%
Ashkenazi Jewish	597	29,606	14,803	11	2.02%	1.86%	2.18%
Middle Eastern	117	6060	3030	5	1.93%	1.61%	2.31%
Amish	14	908	454	0	1.54%	0.92%	2.58%
African/African American	743	75,040	37,520	5	0.99%	0.92%	1.06%
East Asian	5	44,888	22,444	0	0.01%	0.01%	0.03%
Our study	32	2392	1196	0	1.34%	0.95%	1.88%

* Sample Denominator is the number of diploid individuals contributing to each estimate, calculated as the allele number divided by 2 (each individual carries two alleles at this locus).

3.3. Comparison with Recently Published Real-World Data

Hanbali et al. reported the first Middle Eastern real-world experience with tisagenlecleucel in 20 adolescent and young adult B-ALL patients from Saudi Arabia [30]. At 12-month median follow-up, 1-year RFS and OS were 56% and 74%, respectively; loss of B-cell aplasia occurred in 35% of patients at a median of 5 months, and 35% proceeded to allogeneic stem cell transplantation. Severe cytokine release syndrome (grade III–IV) occurred in 5% of patients [30], compared to reported background rates of 20–30 in other cohorts [31]. The Saudi study did not include CD19 genotyping, and therefore, the potential effect of the L174V polymorphism on outcomes remains unknown. Our genomic findings lay the foundation for future prospective studies integrating genotyping into CAR-T trials in Middle Eastern populations.

4. Discussion

This study provides the first comprehensive characterization of CD19 genetic variation in a large Arab cohort, addressing a critical gap in pharmacogenomic knowledge that has persisted despite the global expansion of CAR-T therapy. The principal finding of 76.6% V174 allele frequency (minor allele frequency of L174 = 23.4%) and 65.3% V/V homozygosity positions Arab populations at the upper end of the global spectrum for this potentially therapy-relevant variant. This finding generates testable hypotheses for future clinical investigations.

The location of L174V within the minimal recognition domain of the FMC63 scFv is particularly noteworthy. Structural studies have demonstrated that the V174 substitution induces conformational changes approximately 23 Å from the binding interface, potentially affecting binding kinetics through allosteric mechanisms rather than direct steric hindrance [25,27]. More compellingly, Zhang et al. demonstrated that FMC63 CAR-T cells effectively eliminated wild-type CD19-expressing cells but showed reduced killing of cells harboring CD19 mutations, while an alternative CAR (21D4) targeting a distinct membrane-proximal epitope retained efficacy against mutated targets [32].

A critical distinction is required here. These studies examined somatic mutations acquired during tumor evolution under therapeutic pressure, not germline polymorphisms present from birth in all B cells. The clinical significance of germline L174V homozygosity remains entirely unexplored. Our finding that 65% of this Arab cohort are V/V homozygotes compared to 52.5% of Europeans raises fundamental questions that remain unanswered pending functional and clinical studies.

4.1. Mechanistic Hypotheses and Testable Predictions

The following hypotheses are purely speculative and are not supported by experimental data from this study. They are presented solely as a framework to guide future research. We have not conducted binding assays, CD19 expression experiments, or structural modeling.

Based on the available structural and functional evidence, we propose three non-mutually exclusive mechanistic hypotheses regarding the potential impact of germline L174V status on CAR-T therapy outcomes. These hypothetical structural insights may partially explain the suboptimal clinical outcomes attributed to an altered scFv-CD19 interface, reflecting a subtle rather than complete CD19 epitope loss. Three testable hypotheses are therefore proposed as a framework for future investigations.

Hypothesis 1 (Binding Affinity Hypothesis). *The V174 isoform might have reduced binding affinity for the FMC63 scFv compared to the L174 isoform. This hypothesis would predict that V/V genotypes could have higher thresholds for CAR-T activation, though the opposite effect is also possible. This hypothesis can be tested through surface plasmon resonance binding studies using recombinant CD19 isoforms and FMC63 scFv, as well as cellular avidity assays comparing CAR-T recognition of L174-expressing versus V174-expressing target cells [33,34].*

Hypothesis 2 (Epitope Accessibility Hypothesis). *The substitution of valine which has a shorter side chain relative to leucine (reference residue) is predicted to weaken the CD19 hydrophobic packing against critical CDR residues, thereby exerting a modulatory (rather than absolute) effect on scFv-CD19 binding. The V174 substitution could potentially alter the conformational presentation of the FMC63 epitope. This could be tested using CD19 expression titration experiments and competition binding assays with conformation-specific antibodies [35].*

Hypothesis 3 (Antigen Escape Dynamics Hypothesis). *Germline V174 homozygosity might alter the fitness landscape for somatic escape mutations. This can be tested through longitudinal sequencing of CD19 in patients relapsing after CAR-T therapy, stratifying by germline genotype [36,37].*

4.2. The Research Opportunity of Consanguineous Population Structure

Though statistically significant, the Hardy–Weinberg deviation may be partly explained by Arab demographic features: consanguinity rates of 25–60% and strong founder effects [17,38]. If biological in origin, elevated homozygosity in consanguineous populations could amplify any phenotype associated with homozygous variant carriage. If V174 homozygosity affects CAR-T outcomes, detection would be easier in Arab (65% homozygotes) than European (52% homozygotes) populations. Statistical power is substantially higher with greater homozygosity, especially for recessive/additive models [39].

4.3. Implications for Global Health Equity in Immunotherapy

Underrepresentation of Middle Eastern populations in CAR-T trials is both a missed scientific opportunity and an equity concern. As of 2025, <2% of enrolled participants in CD19 CAR-T trials are of Middle Eastern/North African ancestry, despite these regions

comprising ~5% of the global population [40]. This disparity persists despite increasing regional availability: King Faisal Specialist Hospital has treated >200 patients with CAR-T and launched local manufacturing [<https://www.kfshrc.edu.sa/en/news/2025/10/kfshrc-treats-200-patients-with-cart-therapy-and-launches-local-manufacturing>, accessed on 1 June 2026], and King Hussein Cancer Center has started CAR-T trials with 17 enrolled patients (unpublished data). Arab populations exhibit distinct genetic profiles at a therapy-relevant locus. Whether these differences translate into differential outcomes remains unknown, as no binding, structural, or clinical data support such a claim. This is a potential source of health inequity if V174 homozygosity affects CAR-T efficacy or toxicity. Current trials have not included the population where this effect would be most detectable.

5. Limitations

First, our analysis uses public genomic data without clinical outcomes; no CAR-T trial has incorporated CD19 genotyping in this population. We cannot establish a direct link but provide a genomic foundation. Second, despite relatedness filtering (π -hat > 0.125) and PCA, cryptic relatedness or substructure may persist; technical artifacts cannot be ruled out. Third, comparisons with reference databases are not harmonized due to differences in sequencing, variant calling, and coverage; we present descriptive comparisons. Fourth, while $n = 1196$ is substantial for an underrepresented population, it remains modest versus gnomAD; validation in independent Arab cohorts is needed. Fifth, functional studies of L174V's effect on FMC63 binding kinetics are encouraged; we present no binding, expression, or clinical data.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jpm16070343/s1>, Table S1, title: (CD19 coding variants identified in 1196 Arab individuals, ranked by alternate allele frequency).

Author Contributions: Conceptualization, O.A.; methodology, O.A. and M.A.-H.; software, A.A.O.; validation, A.A.O. and O.A.; formal analysis, A.A.O. and O.A.; investigation, O.A. and M.A.-H.; resources, O.A., A.A.O., and M.A.-H.; data curation, A.A.O.; writing original draft preparation, O.A.; writing review and editing, M.A.-H., A.A.O. and O.A.; visualization, M.A.-H., A.A.O. and O.A.; supervision, O.A.; project administration, O.A.; funding acquisition, O.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Institutional Review Board Statement: This study analyzed de-identified publicly available sequencing data originally collected under protocols approved by the Qatar Biomedical Research Institute Institutional Review Board. As a secondary analysis of publicly available data, Scientific Approval was granted by the Research Council office, chaired by Dr. Amal Al-Omari, on 8 May 2019, code no. 19KHCC21. The IRB approval was granted on 17 December 2019 by IRB Chair Dr. Maysa Al-Hussaini.

Informed Consent Statement: To this study, the patients whom the data was derived from were consented under the Qatar genome project, and the corresponding sequencing data was made accessible following a de-identification process by the investigators, whereby participants are unidentifiable. All QBB participants at the time signed an Informed Consent Form before their participation.

Data Availability Statement: All sequencing data analyzed in this study are publicly available through the NCBI Sequence Read Archive under accessions SRP060765, SRP061943, and SRP061463. Analysis scripts and derived variant call files have been deposited at Zenodo (doi <https://doi.org/10.1038/hgv.2016.16>). The original data is publicly available [doi <https://doi.org/10.1038/hgv.2016.16>].

Acknowledgments: The authors acknowledge the participants of the Qatar Genome Program whose data made this analysis possible. We thank the open science community for maintaining publicly

accessible genomic databases that enable research on underrepresented populations. This study was conducted as part of a hypothesis-generating initiative to promote diversity in precision immunotherapy research. The authors also wish to thank the leadership and management of KHCC for their support. We also thank the research office and IRB office for their support and approval of this research.

Conflicts of Interest: The authors declare no conflicts of interest. No pharmaceutical or biotechnology company had any role in the design, analysis, or interpretation of this study.

References

1. June, C.H.; Sadelain, M. Chimeric Antigen Receptor Therapy. *N. Engl. J. Med.* **2018**, *379*, 64–73. [[CrossRef](#)] [[PubMed](#)]
2. Maude, S.L.; Laetsch, T.W.; Buechner, J.; Rives, S.; Boyer, M.; Bittencourt, H.; Bader, P.; Verneris, M.R.; Stefanski, H.E.; Myers, G.D.; et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *N. Engl. J. Med.* **2018**, *378*, 439–448. [[CrossRef](#)] [[PubMed](#)]
3. Neelapu, S.S.; Locke, F.L.; Bartlett, N.L.; Lekakis, L.J.; Miklos, D.B.; Jacobson, C.A.; Braunschweig, I.; Oluwole, O.O.; Siddiqi, T.; Lin, Y.; et al. Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *N. Engl. J. Med.* **2017**, *377*, 2531–2544. [[PubMed](#)]
4. Testa, U.; Sica, S.; Pelosi, E.; Castelli, G.; Leone, G. CAR-T Cell Therapy in B-Cell Acute Lymphoblastic Leukemia. *Mediterr. J. Hematol. Infect. Dis.* **2024**, *16*, e2024010. [[CrossRef](#)] [[PubMed](#)]
5. van der Stegen, S.J.C.; Hamieh, M.; Sadelain, M. The pharmacology of second-generation chimeric antigen receptors. *Nat. Rev. Drug Discov.* **2015**, *14*, 499–509. [[CrossRef](#)] [[PubMed](#)]
6. Benmebarek, M.R.; Karches, C.H.; Cadilha, B.L.; Lesch, S.; Endres, S.; Kobold, S. Killing Mechanisms of Chimeric Antigen Receptor (CAR) T Cells. *Int. J. Mol. Sci.* **2019**, *20*, 1283. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
7. Shah, N.N.; Fry, T.J. Mechanisms of resistance to CAR T cell therapy. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 372–385. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
8. Sotillo, E.; Barrett, D.M.; Black, K.L.; Bagashev, A.; Oldridge, D.; Wu, G.; Sussman, R.; LaNauze, C.; Ruella, M.; Gazzara, M.R.; et al. Convergence of Acquired Mutations and Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy. *Cancer Discov.* **2015**, *5*, 1282–1295. [[CrossRef](#)] [[PubMed](#)]
9. Orlando, E.J.; Han, X.; Tribouley, C.; Wood, P.A.; Leary, R.J.; Riester, M.; Levine, J.E.; Qayed, M.; Grupp, S.A.; Boyer, M.; et al. Genetic mechanisms of target antigen loss in CAR19 therapy of acute lymphoblastic leukemia. *Nat. Med.* **2018**, *24*, 1504–1506. [[CrossRef](#)] [[PubMed](#)]
10. Zola, H.; MacArdle, P.J.; Bradford, T.; Weedon, H.; Yasu, H.; Kurosawa, Y. Preparation and characterization of a chimeric CD19 monoclonal antibody. *Immunol. Cell Biol.* **1991**, *69*, 411–422. [[CrossRef](#)] [[PubMed](#)]
11. Huang, J.; Huang, X.; Yu, D. CD19-negative relapse after CAR-T cell therapy: Mechanisms of antigen escape and lineage switch. *Front. Immunol.* **2026**, *17*, 1692287. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
12. He, C.; Mansilla-Soto, J.; Khanra, N.; Hamieh, M.; Bustos, V.; Paquette, A.J.; Angus, A.G.; Shore, D.M.; Rice, W.J.; Khelashvili, G.; et al. CD19 CAR antigen engagement mechanisms and affinity tuning. *Sci. Immunol.* **2023**, *8*, eadf1426. [[CrossRef](#)] [[PubMed](#)]
13. Krawczyk, M.; Fernandez-Fuentes, N.; Fidy, K.; Winiarski, T.; Pepek, M.; Graczyk-Jarzynka, A.; Davis, J.; Bousquets-Muñoz, P.; Puente, X.S.; Menendez, P.; et al. The costimulatory domain influences CD19 CAR-T cell resistance development in B-cell malignancies. *bioRxiv* **2025**. [[CrossRef](#)] [[PubMed](#)]
14. Lee, M.J.; Cichocki, F.; Miller, J.S. Chimeric antigen receptor therapies: Development, design, and implementation. *J. Allergy Clin. Immunol.* **2025**, *156*, 70–80. [[CrossRef](#)] [[PubMed](#)]
15. Seipel, K.; Abbühl, M.; Bacher, U.; Nilius, H.; Daskalakis, M.; Pabst, T. Clinical Impact of Single Nucleotide Polymorphism in CD-19 on Treatment Outcome in FMC63-CAR-T Cell Therapy. *Cancers* **2023**, *15*, 3058. [[PubMed](#)]
16. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)] [[PubMed](#)]
17. Tadmouri, G.O.; Nair, P.; Obeid, T.; Al Ali, M.T.; Al Khaja, N.; Hamamy, H.A. Consanguinity and reproductive health among Arabs. *Reprod. Health* **2009**, *6*, 17. [[CrossRef](#)] [[PubMed](#)]
18. Fakhro, K.A.; Staudt, M.R.; Ramstetter, M.D.; Robay, A.; Malek, J.A.; Badii, R.; Al-Marri, A.A.-N.; Khalil, C.A.; Al-Shakaki, A.; Chidiac, O.; et al. The Qatar genome: A population-specific tool for precision medicine in the Middle East. *Hum. Genome Var.* **2016**, *3*, 16016. [[CrossRef](#)] [[PubMed](#)]
19. Mbarek, H.; Devadoss Gandhi, G.; Selvaraj, S.; Al-Muftah, W.; Badji, R.; Al-Sarraj, Y.; Saad, C.; Darwish, D.; Alvi, M.; Fadl, T.; et al. Qatar genome: Insights on genomics from the Middle East. *Hum. Mutat.* **2022**, *43*, 499–510. [[CrossRef](#)] [[PubMed](#)]

20. Dunn, T.; Berry, G.; Emig-Agius, D.; Jiang, Y.; Lei, S.; Iyer, A.; Udar, N.; Chuang, H.-Y.; Hegarty, J.; Dickover, M.; et al. Pisces: An accurate and versatile variant caller for somatic and germline next-generation sequencing data. *Bioinformatics* **2019**, *35*, 1579–1581. [\[PubMed\]](#)
21. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Monies, D.; Abouelhoda, M.; AlSayed, M.; Alhassnan, Z.; Alotaibi, M.; Kayyali, H.; Al-Owain, M.; Shah, A.; Rahbeeni, Z.; Al-Muhaizea, M.A.; et al. The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. *Hum. Genet.* **2017**, *136*, 921–939. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Wutti-in, Y.; Sujitjoo, J.; Sawasdee, N.; Panya, A.; Kongkla, K.; Yuti, P.; Yongpitakwattana, P.; Thepmalee, C.; Junking, M.; Chieochansin, T.; et al. Development of a Novel Anti-CD19 CAR Containing a Fully Human scFv and Three Costimulatory Domains. *Front. Oncol.* **2022**, *11*, 802876. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Duzkale, H.; Shen, J.; McLaughlin, H.; Alfares, A.; Kelly, M.; Pugh, T.; Funke, B.; Rehm, H.; Lebo, M.S. A systematic approach to assessing the clinical significance of genetic variants. *Clin. Genet.* **2013**, *84*, 453–463. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Susa, K.J.; Kruse, A.C.; Blacklow, S.C. Tetraspanins: Structure, dynamics, and principles of partner-protein recognition. *Trends Cell Biol.* **2024**, *34*, 509–522. [\[CrossRef\]](#) [\[PubMed\]](#)
27. He, X.; Kläsener, K.; Iype, J.M.; Becker, M.; Maity, P.C.; Cavallari, M.; Nielsen, P.J.; Yang, J.; Reth, M. Continuous signaling of CD79b and CD19 is required for the fitness of Burkitt lymphoma B cells. *EMBO J.* **2018**, *37*, e97980. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Shoham, T.; Rajapaksa, R.; Kuo, C.C.; Haimovich, J.; Levy, S. Building of the tetraspanin web: Distinct structural domains of CD81 function in different cellular compartments. *Mol. Cell. Biol.* **2006**, *26*, 1373–1385. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Carter, R.H.; Fearon, D.T. CD19: Lowering the threshold for antigen receptor stimulation of B lymphocytes. *Science* **1992**, *256*, 105–107. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Hanbali, A.; Bahkali, F.; Syed, F.; Alahmari, A.; Alotaibi, A.; Alfayez, M.; Saad, A.; Alkhalidi, H.; Alfraih, F.; Alhayli, S.; et al. Outcomes of adolescent and young adult (AYA) patients with relapsed/refractory B-cell acute lymphoblastic leukemia treated with tisagenlecleucel, real world evidence from the middle east. *Cytotherapy* **2025**, *27*, 1308–1315. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Lee, D.W.; Santomaso, B.D.; Locke, F.L.; Ghobadi, A.; Turtle, C.J.; Brudno, J.N.; Maus, M.V.; Park, J.H.; Mead, E.; Pavletic, S.; et al. ASTCT Consensus Grading for Cytokine Release Syndrome and Neurologic Toxicity Associated with Immune Effector Cells. *Biol. Blood Marrow Transplant.* **2019**, *25*, 625–638. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Zhang, Z.; Chen, X.; Tian, Y.; Li, F.; Zhao, X.; Liu, J.; Yao, C.; Zhang, Y. Point mutation in CD19 facilitates immune escape of B cell lymphoma from CAR-T cell therapy. *J. Immunother. Cancer* **2020**, *8*, e001150. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Gacerez, A.T.; Hua, C.K.; Ackerman, M.E.; Sentman, C.L. Chimeric antigen receptors with human scFvs preferentially induce T cell anti-tumor activity against tumors with high B7H6 expression. *Cancer Immunol. Immunother.* **2018**, *67*, 749–759. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Harrer, D.C.; Li, S.-S.; Kaljanac, M.; Barden, M.; Pan, H.; Abken, H. Fine-tuning the antigen sensitivity of CAR T cells: Emerging strategies and current challenges. *Front. Immunol.* **2023**, *14*, 1321596. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Sommermeyer, D.; Hill, T.; Shamah, S.M.; Salter, A.I.; Chen, Y.; Mohler, K.M.; Riddell, S.R. Fully human CD19-specific chimeric antigen receptors for T-cell therapy. *Leukemia* **2017**, *31*, 2191–2199. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Good, Z.; Spiegel, J.Y.; Sahaf, B.; Malipatlolla, M.B.; Ehlinger, Z.J.; Kurra, S.; Desai, M.H.; Reynolds, W.D.; Lin, A.W.; Vandris, P.; et al. Post-infusion CAR TReg cells identify patients resistant to CD19-CAR therapy. *Nat. Med.* **2022**, *28*, 1860–1871. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Sarén, T.; Ramachandran, M.; Gammelgård, G.; Lövgren, T.; Mirabello, C.; Björklund, Å.K.; Wikström, K.; Hashemi, J.; Freyhult, E.; Ahlström, H.; et al. Single-Cell RNA Analysis Reveals Cell-Intrinsic Functions of CAR T Cells Correlating with Response in a Phase II Study of Lymphoma Patients. *Clin. Cancer Res.* **2023**, *29*, 4139–4152. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Bittles, A.H.; Black, M.L. Consanguinity, human evolution, and complex diseases. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1779–1786. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Sham, P.C.; Purcell, S.M. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* **2014**, *15*, 335–346. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Turner, B.E.; Steinberg, J.R.; Weeks, B.T.; Rodriguez, F.; Cullen, M.R. Race/ethnicity reporting and representation in clinical trials: A systematic review. *JAMA Netw. Open* **2024**, *7*, e2411312.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.