Integrating US Census Demographic Data with HLA Frequencies to Accelerate Targeted Cell Therapy Patient Enrollment

ASHI's 50th Annual Meeting
October 21 - 25, 2024

Poster #203

Christian Roy [1], Tomasz Sewastianik [1], Ileana Saenz [1], Gregory J. Opiteck [1], Sean Stagg [2], Martin Maiers [2], Dirk Nagorsen [1]
[1] Affini-T Therapeutics; [2] NMDP

Introduction

Recent TCR-based therapies approvals (e.g. Adaptimmune's Tecelra for HLA*02:01 + MAGEA4 in Synovial Sarcoma) are catalyzing more clinical development in the space (TCR-T's, TCEs, vaccines). These approaches require the determination of a specific HLA allele in each patient. Patient identification, communication, and ultimately clinical trial enrollment can all be assisted through consideration of the US demographic diversity and its strong association with HLA allelic frequencies.

Methods

The most comprehensive public United States HLA frequency dataset (Gragert et al. 2013) was integrated with regional demographics obtained from the US 2020 Census. Once allelic frequencies were adjusted for US demographics on a national, state, county or other geographically resolved basis (e.g. H3 Hexagons), genotypic frequencies were calculated and displayed on a map of the continental US.

Flowchart 1: details of population-corrected genotypic frequency process

US HLA Frequencies
- US-wide HLA Frequencies
- Source: Gragert2013 (aka: NMDP) Allelic frequencies by race code
- Represents a country-wide summary and does not represent regional population structure

Regional Demographics
- US Population by Race within [region]
- Source: US Census via *censusapi* (Recht2022)
- Multi-race responders accounted for
- [regions] can be state, county, tract

Adjusted Frequencies
- Gragert2013 * [region] demographics
- Sum allelic frequencies (AF) across races within [region]
- Convert to genotypic frequency (GF)
- $GF = 1 - (1 - AF)^2$

Country	Ethnic Code	Single Race Population	Multi-race Population	Total Population	Percentage of Total	Allele	Calculated GF	Population-Adjusted GF
US	AFA	39,940,338	2,064,019	42,004,357	13%	A*11:01	3%	0%
US	API	20,240,737	1,820,295	22,061,032	7%	A*11:01	32%	2%
US	CAU	191,697,647	5,944,911	197,642,558	60%	A*11:01	12%	7%
US	HIS	62,080,044	NA	62,080,044	19%	A*11:01	9%	2%
US	NAM	2,251,699	2,131,361	4,383,060	1%	A*11:01	10%	0%
US	UNKOWN	1,689,833	1,419,206	3,109,039	1%	A*11:01	NA	NA
Total				331,280,090	99%		66%	11%

Table 1: Example values for calculation of US-wide population-adjusted genotypic frequency. Single race population counts are combined with multi-race apportioned counts by region (in this case the entire US), by race code. The percent of self-assigned race is then multiplied by the allelic frequency of Gragert2013 for the matching race code. Prior to genotypic frequency calculation, allelic frequencies are summed across race codes.

Results

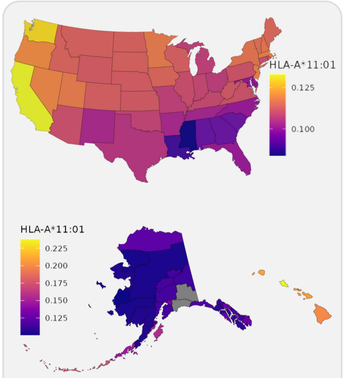


Figure 1: Population-adjusted HLA-A*11:01 frequencies by Continental US State (top), Alaska and Hawaii by County (bottom).
Note: scale difference between top and bottom.

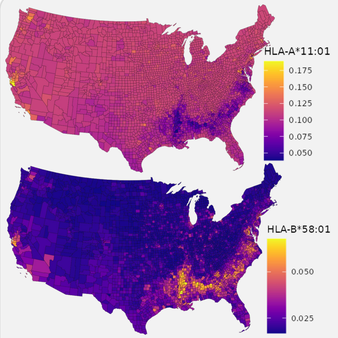


Figure 2: Population-adjusted genotypic frequencies for HLA-A*11:01 (top) and HLA-B*58:01 (bottom) by US county. HLA-B*58:01 is linked to severe adverse reactions from allopurinol, a common treatment for gout, and is monitored in certain high-risk populations (APA and AFA).

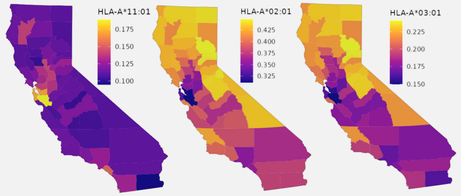
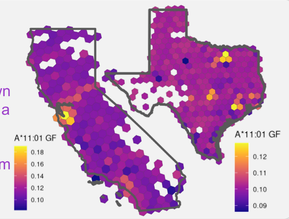


Figure 3: Population-adjusted genotypic frequencies for HLA-A*11:01 (left); HLA-A*02:01 (middle); HLA-A*03:01 (right) by California county.

Figure 4: Population-adjusted genotypic frequencies for HLA-A*11:01 for California (left) and Texas (right) H3:h4 hexagons. White areas are shown where a hexagon does not have a geometry center overlapping a census tract.
Use of H3 allows for more uniform geographic-area consideration compared to census areas.



CA County	HLA-A*11:01 Census-Corrected Genotypic Frequency
Santa Clara	18.90
San Francisco	17.80
San Mateo	17.24
Alameda	17.00
Orange	15.19
Sutter	14.32
Contra Costa	14.09
Trinity	14.02
Sacramento	14.00

Table 2: Population-adjusted Genotypic frequencies for the top 10 CA counties by HLA-A*11:01.

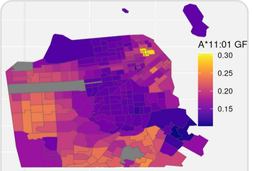


Figure 5: San Francisco County, with census tracts colored by population-adjusted HLA-A*11:01 genotypic frequencies.

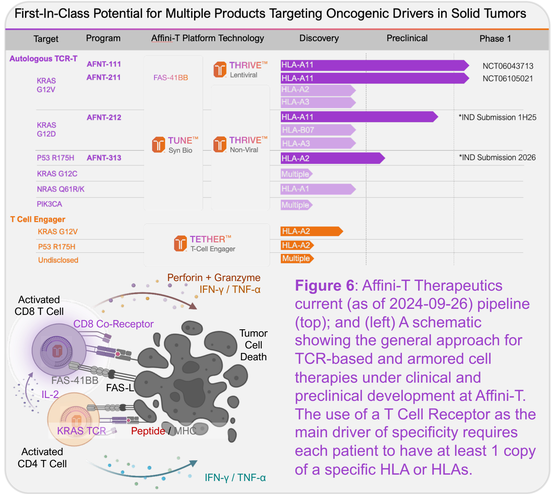
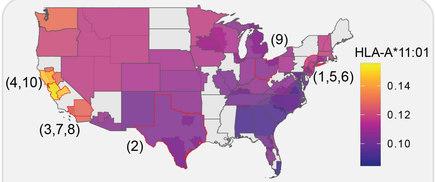


Figure 6: Affini-T Therapeutics current (as of 2024-09-26) pipeline (top); and (left) A schematic showing the general approach for TCR-based and armored cell therapies under clinical and preclinical development at Affini-T. The use of a T Cell Receptor as the main driver of specificity requires each patient to have at least 1 copy of a specific HLA or HLAs.



Cancer Center Catchment Area (DelNero2022)	2020 Census Pop	A*11:01 GF	Estimated A11:01 + population
MSKCC; NY (1)	30,168,760	12%	3,497,825
MD Anderson; TX (2)	28,864,201	11%	3,048,978
City of Hope; CA (3)	17,532,615	13%	2,240,127
UCSF; Helen Diller; CA (4)	13,417,610	15%	1,974,542
Columbia; Herbert Irving; NY (5)	10,762,612	12%	1,304,497
NYU; Langone Health; NY (6)	10,226,716	12%	1,277,349
UCLA; Jonsson; CA (7)	9,847,003	13%	1,254,394
USC; Norris; CA (8)	9,847,003	13%	1,254,394
Ohio State; OH (9)	11,652,945	11%	1,247,337
Stanford; CA (10)	7,817,822	16%	1,217,373

Figure 7: HLA-A*11:01 Population-adjusted genotypic frequencies by NCI Catchment areas (delNero2022). Locations of the top 10 catchment centers are labeled according to their rank (top, 1 = most patients), based on number of individuals living within a catchment that are estimated to have at least one copy of A*11:01 (table, bottom).

Validation

There are significant (p value < 2.2e-16) and positive ($\rho > 0.5$) correlations between US-census based regional estimates (county and hexagon (Sahr2003)) and geo-located donor data from the NMDP, suggesting that the approach accurately re-creates regional diversity from orthogonal (Census data) and summarized (published allelic frequencies) information.

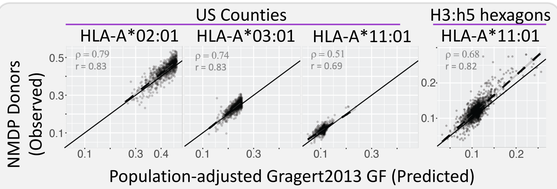


Figure 8: Correlation scatter plots comparing population-adjusted genotypic frequencies for HLA-A*11:01, A*02:01, and A*03:01 to geographically-matched frequency data from NMDP. Importantly, only counties or hexagons with >= 500 people were included

Conclusions

- TCR/HLA-mediated therapies would benefit from allele frequency mapping to enhance patient identification and enrollment.
- This method for mapping geographical HLA allele frequencies is accurate at state, county, and 253m² (H3:h5) hexagon levels.
- Validation with NMDP data allows geographic predictions from de-identified data and use of Census data allows for periodic updates based on demographic changes over time.
- HLA-A*11:01 and other alleles show significant geographical variation both between and within states.
- This approach can be applied to NCI catchment areas to optimize patient treatment through targeted enrollment at centers where allele frequencies are highest.
- Utility of this approach spans beyond immune-oncology patient identification and can flag areas with increased risk to drug reactions .

References

Bureau, U.S. Census. n.d. "Hispanic or Latino, and Not Hispanic or Latino by Race." https://data.census.gov/table/DECENNIALCD1182020.P97?g=Race.

Gragert, Loren, Abeer Madbouly, John Freeman, and Martin Maiers. 2013. "Six-Locus High Resolution HLA Haplotype Frequencies Derived from Mixed-Resolution DNA Typing for the Entire US Donor Registry." *Human Immunology* 74 (10): 1313–20. https://doi.org/10.1016/j.humimm.2013.06.025.

DelNero, Peter F., Ian D. Buller, Rena R. Jones, Zaria Tatalovich, Robin C. Vanderpool, Henry P. Ciolino, and Robert T. Croyle. 2022. "A National Map of NCI-Designated Cancer Center Catchment Areas on the 50th Anniversary of the Cancer Centers Program." *Cancer Epidemiology, Biomarkers & Prevention* 31 (5): 965–71. https://doi.org/10.1158/1055-9965.EPI-21-1230.

Sahr, Kevin, Denis White, and A. Jon Kimerling. 2003. "Geodesic Discrete Global Grid Systems." *Cartography and Geographic Information Science* 30 (2): 121–34. https://doi.org/10.1559/152304003100011090. and https://h3geo.org/

Recht, Hannah. 2022. "Censusapi: Retrieve Data from the Census Apis." Manual. https://www.hrecht.com/censusapi/.