



Short population report

Human leukocyte antigen class I (A, B and C) allele and haplotype variation in a South African Mixed ancestry population



Shayne Loubser, Maria Paximadis, Caroline T. Tiemessen*

Centre for HIV-1 and STIs, National Institute for Communicable Diseases, National Health Laboratory Services and Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Article history:

Received 18 January 2017

Revised 13 April 2017

Accepted 19 April 2017

Available online 21 April 2017

Keywords:

HLA alleles

HLA haplotypes

South African population

Mixed ancestry

Coloured

ABSTRACT

South Africa has a large (~53 million), ethnically diverse population (black African, Caucasian, Indian/Asian and Mixed ancestry) and a high disease burden (particularly HIV-1 and *Mycobacterium tuberculosis*). The Mixed ancestry population constitutes ~9% of the total population and was established ~365 years ago in the Western Cape region through interracial mixing of black Africans, Europeans and Asians. Admixed populations present unique opportunities to identify genetic factors involved in disease susceptibility. Since *HLA* genes are important mediators of host immunity, we investigated *HLA-A*, *-B* and *-C* allele and haplotype diversity in 50 healthy, unrelated individuals recruited from the Mixed ancestry population.

© 2017 The Authors. Published by Elsevier Inc. on behalf of American Society for Histocompatibility and Immunogenetics. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

South Africa, located at the most southerly tip of Africa, has a unique demographic history and ethnic diversity, due to extensive migration of people from the northern parts of Africa and also via oceanic trade-routes between the West and the East. The southern African region was first populated by the Khoi-San, before Bantu-speaking tribes migrated into the area ~2000 years ago. European colonisation of the Western Cape region began in the mid-1600's and included Dutch, German, English and French settlers as well as slave labour imported from Indonesia, India, Madagascar and East Africa. The South African Mixed ancestry (SAM) or "Coloured" population was established through interracial mixing of Europeans, slaves and the indigenous Khoi-San and Xhosa tribes. In 2013, the South African population was recorded at 52,982,000 individuals, divided into four major population groups (black African, Caucasian, Indian/Asian and Mixed ancestry) (statssa.gov.za/publications/P0302/P03022013.pdf). The predominantly Afrikaans (afr)-speaking SAM population (~80%), but also English (eng)-speaking, mostly reside in the Western Cape Province and constitute ~9% of the total population [1].

Several genetic studies interrogating autosomes agree with historical records of early Western Cape demographics that contributed to the genetic admixture of the SAM population [1–5]. Mitochondrial DNA and Y-chromosome analyses also confirm at

least five genetically distinct lineages involving Khoi-San, Bantu, European, Indian and Southeast Asian sources, and further reveal a gender biased admixture showing enrichment of Khoi-San maternally inherited mitochondrial DNA, while Y-chromosomes originated primarily from European and African males [3,6].

HIV-1 and *Mycobacterium tuberculosis* (*Mtb*) infection are the most common infectious diseases and leading causes of death in South Africa. In the SAM population, HIV-1 prevalence was 3.1% in 2012, well below the national prevalence of 12.2% (hsr.ac.za/en/research-outputs/view/6871), while in 2013, *Mtb* was the second leading cause of death (7.1%) (statssa.gov.za/publications/P03093/P030932013.pdf). Differential susceptibility to infectious diseases suggests a role for host immunogenetic factors. Indeed, specific *HLA* alleles associate with both *Mtb* [7] and HIV-1 infection [8].

We investigated *HLA-A*, *-B* and *-C* allele and haplotype variation in 50 healthy, unrelated individuals recruited from the SAM population. Buffy coats stored at –80 °C were available for 40 individuals from the Electricity Supply Commission (ESKOM) cohort [9], while an additional 10 volunteers were recruited from staff/students at the National Institute for Communicable Diseases, the University of the Witwatersrand (WITS) Medical School and the WITS Donald Gordon Medical Centre in Johannesburg. Ethnicity was self-determined by each participant and signed consent was also obtained. Ethics approval was granted by the Human Research Ethics Committee (Medical) of the WITS medical faculty.

Genomic DNA was extracted from buffy coats or fresh whole blood using the QiaAmp DNA blood mini kit (Qiagen, Hilden,

* Corresponding author at: National Institute for Communicable Diseases, Private Bag X4, Sandringham, 2131, South Africa.

E-mail address: carolinet@nicd.ac.za (C.T. Tiemessen).

Germany) and quantified using a nanodrop spectrophotometer (Thermo Scientific, Waltham, USA). The sequence-based typing (SBT) resolver kit (Conexio Genomics, Fremantle, Australia) was used to generate *HLA-A*, *-B* and *-C* amplicons for DNA sequencing as described by the manufacturers. PCR products were purified using Agencourt Ampure XP magnetic bead separation (Becton-Dickenson, Franklin lakes, USA). Sequence products were generated using the BigDye terminator v3.1 cycle sequencing kit (Life Technologies, Carlsbad, USA), purified using ethanol-sodium acetate precipitation and resolved on an ABI3100 PRISM Genetic Analyser (Life Technologies, Carlsbad, USA). Alleles were assigned using ASSIGN-SBT v4.7 software (Conexio Genomics, Fremantle, Australia) and v3.13.1.4 of the IPD-IMGT/HLA reference database released on July 25, 2013. No additional molecular-based methods were employed to resolve ambiguities, however, the following steps were taken to exclude unlikely allele combinations from the ambiguity list: (i) analysis using the SBT software was initially used to interrogate differences in exons 1–4 for *HLA-A/B* and exons 1–8 for *HLA-C*, (ii) since partial 5'UTR and introns 1,2 for *HLA-A/B* and partial 5'UTR, introns 1,2,4–7 and partial 3'UTR sequences were also generated, these differences could also be analysed by manual editing using the SBT software, (iii) likely allele combinations were based on “Common” and “Well-Documented” alleles listed in version 2.0.0 of the “Common and Well-Documented HLA alleles catalogue” updated in 2012 [10], (iv) in some cases, other studies investigating the same population [7] or ancestral populations such as South African black and Caucasian populations [9] were consulted and finally, (v) haplotype data available on the Allele Frequency Net Database (AFND) was also used to confirm linkage of *HLA-A*, *-B* and *-C* alleles [11]. Allele frequencies, haplotypes and deviations from Hardy-Weinberg equilibrium were calculated using the PyPop-Win32-0.7.0 software program available online (www.pyPop.org) [12].

Overall, no deviations from Hardy-Weinberg equilibrium were detected at *HLA-A*, *-B* and *-C* loci ($p = 0.1852$, $p = 0.4713$ and $p = 0.3362$, respectively). The number of unique *HLA-A*, *-B* and *-C* alleles identified was $n = 24$, $n = 43$ and $n = 25$, respectively (Table S1). *HLA-A*02:01:01:01*, *-B*07:02:01* and *-C*06:02:01:01* were the most common alleles detected at frequencies of 0.21, 0.08 and 0.12, respectively (Table S1). *HLA-A*03:01:01:01-B*07:02:01-C*07:02:01:01* was the most common haplotype detected at a frequency of 0.04, from a total of 81 unique haplotypes identified (Table S2).

The SAM population is a relatively understudied ethnic group of South Africa, and this report may contribute to a better understanding of disease susceptibility related to immune function. Data has been deposited in the AFND database under the identifier 3395.

Acknowledgements

This work is based on the research supported by Grants awards from the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council, a grantee of the Bill and Melinda Gates Foundation and the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.humimm.2017.04.006>.

References

- [1] E. de Wit, W. Delpont, C.E. Rugamika, A. Meintjes, M. Moller, P.D. van Helden, et al., Genome-wide analysis of the structure of the South African coloured population in the Western Cape, *Hum. Genet.* 128 (2010) 145.
- [2] N. Patterson, D.C. Petersen, R.E. van der Ross, H. Sudoyo, R.H. Glashoff, S. Marzuki, et al., Genetic structure of a unique admixed population: implications for medical research, *Hum. Mol. Genet.* 19 (2010) 411.
- [3] C.M. Schlebusch, M. Lombard, H. Soodyall, MitDNA control region variation affirms diversity and deep sub-structure in populations from southern Africa, *BMC Evol. Biol.* 13 (2013) 56.
- [4] M. Daya, L. van der Merwe, U. Galal, M. Moller, M. Salie, E.R. Chimusa, et al., A panel of ancestry informative markers for the complex five-way admixed South African coloured population, *PLoS One* 8 (2013) e82224.
- [5] E.R. Chimusa, M. Daya, M. Moller, R. Ramesar, B.M. Henn, P.D. van Helden, et al., Determining ancestry proportions in complex admixture scenarios in South Africa using a novel proxy ancestry selection method, *PLoS One* 8 (2013) e73971.
- [6] L. Quintana-Murci, C. Harmant, H. Quach, O. Balanovsky, V. Zaporozhchenko, C. Bormans, et al., Strong maternal Khoisan contribution to the South African coloured population: a case of gender-biased admixture, *Am. J. Hum. Genet.* 86 (2010) 611.
- [7] M. Salie, L. van der Merwe, M. Moller, M. Daya, G.D. van der Spuy, P.D. van Helden, et al., Associations between human leukocyte antigen class I variants and the *Mycobacterium tuberculosis* subtypes causing disease, *J. Infect. Dis.* 209 (2014) 216.
- [8] P.J. McLaren, M. Carrington, The impact of host genetic variation on infection with HIV-1, *Nat. Immunol.* 16 (2015) 577.
- [9] M. Paximadis, T.Y. Mathebula, N.L. Gentle, E. Vardas, M. Colvin, C.M. Gray, et al., Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population, *Hum. Immunol.* 73 (2012) 80.
- [10] S.J. Mack, P. Cano, J.A. Hollenbach, J. He, C.K. Hurley, D. Middleton, et al., Common and well-documented HLA alleles: 2012 update to the CWD catalogue, *Tissue Antigens* 81 (2013) 194.
- [11] F.F. Gonzalez-Galarza, S. Christmas, D. Middleton, A.R. Jones, Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations, *Nucleic Acids Res.* 39 (2011) D913.
- [12] A.K. Lancaster, R.M. Single, O.D. Solberg, M.P. Nelson, G. Thomson, PyPop update—a software pipeline for large-scale multilocus population genomics, *Tissue Antigens* 69 (Suppl. 1) (2007) 192.