

Article

Variants of *IFNL4* Gene in Amazonian and Northern Brazilian Populations

Carolina Cabral Angelim ¹, Leticia Dias Martins ^{1,2}, Álesson Adam Fonseca Andrade ¹,
Fabiano Cordeiro Moreira ^{3,4}, João Farias Guerreiro ³, Paulo Pimentel de Assumpção ⁴,
Sidney Emanuel Batista dos Santos ^{1,3} and Greice de Lemos Cardoso Costa ^{1,*}

- ¹ Programa de Pós-Graduação em Agentes Infecciosos e Parasitários, Laboratório de Genética Humana e Médica, Federal University of Pará, Belém 66073-000, PA, Brazil; carolina.angelim@icb.ufpa.br (C.C.A.); lemarias@hotmail.com (L.D.M.); alesson.andrade@icb.ufpa.br (Á.A.F.A.); sidneysantos@ufpa.br (S.E.B.d.S.)
- ² Programa de Pós-Graduação em Virologia, Instituto Evandro Chagas, Ananindeua 67030-000, PA, Brazil
- ³ Programa de Pós-Graduação em Genética e Biologia Molecular, Laboratório de Genética Humana e Médica, Federal University of Pará, Belém 66073-000, PA, Brazil; fcmoreira@ufpa.br (F.C.M.); joaofg@ufpa.br (J.F.G.)
- ⁴ Núcleo de Pesquisa em Oncologia, Federal University of Pará, Belém 66073-000, PA, Brazil; assumpcap@gmail.com
- * Correspondence: greice@ufpa.br; Tel.: +55-91-32017843

Abstract: Since the discovery of the polymorphic nature of the *IFNL4* gene, its variants have been investigated and associated with several viral diseases, with an emphasis on hepatitis C. However, the impacts of these variants on mixed-race and native populations in the northern region of Brazil are scarce. We investigated three variants of the *IFNL4* gene in populations from this location, which were among the 14 most frequent variants in worldwide populations, and compared the frequencies obtained to populational data from the 1000 Genomes Project, gnomAD and ABraOM databases. Our results demonstrate that mixed-race and native populations from the northern region of Brazil present frequencies like those of European and Asian groups for the rs74597329 and rs11322783 variants, and like all populations presented for the rs4803221 variant. These data reinforce the role of world populations in shaping the genetic profile of Brazilian populations, indicate patterns of illness according to the expressed genotype, and infer an individual predisposition to certain diseases.

Keywords: *IFNL4*; Amazon population; mixed population; Brazil; genetic variants



Citation: Angelim, C.C.; Martins, L.D.; Andrade, Á.A.F.; Moreira, F.C.; Guerreiro, J.F.; de Assumpção, P.P.; Santos, S.E.B.d.; Costa, G.d.L.C.

Variants of *IFNL4* Gene in Amazonian and Northern Brazilian Populations. *Genes* **2023**, *14*, 2075. <https://doi.org/10.3390/genes14112075>

Academic Editor: Konstantin Khrapko

Received: 15 September 2023

Revised: 18 October 2023

Accepted: 25 October 2023

Published: 14 November 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of genetic variants has acquired significant importance in recent years, as it provides explanations for different clinical conditions and a profile for responses to therapeutic treatments among individuals [1]. In infectious and contagious diseases, several genes have been investigated, and especially in viral infections, the genes that encode interferons deserve to be highlighted because their translation generates proteins linked to antiviral responses [2].

Type III interferons—the family of lambda interferons (*IFNL*)—are considered specialists in antiviral action, blocking stages of the viral replicative cycle, increasing antigenic presentation, and triggering antiproliferative properties [3]. The notoriety of type III interferons began in 2003 with the description of IFN lambda 1 (*IFNL1*), lambda 2 (*IFNL2*) and lambda 3 (*IFNL3*). After 10 years of studies, a fourth type was identified, called IFN lambda 4 (*IFNL4*) [4]. This one stands out from the others due to its highly polymorphic character, followed by an apparent relationship with the sustained virological response (SVR) and the clearance of some viruses, justifying its importance.

The expression of IFN- λ 4 occurs in several cell types, but its recognition is limited to cells that express the IFNLR1 receptor [3,5]. Studies suggest that this cytokine may have evolved specifically to protect the epithelium, as its antiviral effects may have begun

before IFN- λ 1, IFN- λ 2 and IFN- λ 3 [6]. The signaling initiated by IFN- λ 4 activates the JAK-STAT pathway, leading to the expression of interferon-stimulated genes (ISGs) such as Oligoadenylate Synthetase 1 (OAS1) and Ubiquitin-specific Peptidase 18 (USP18) [1].

Studies on type III interferons emphasize how these cytokines have potential in vitro and in vivo against many viruses, such as dengue virus (DENV), cytomegalovirus (CMV), human immunodeficiency virus (HIV), and Epstein–Barr virus (EBV) [1,7]. However, the greatest potential of this gene resides in associative studies of *IFNL4* variants and the spontaneous clearance of the Hepatitis C virus (HCV), with effects on therapies against this infection [8,9].

Despite the time that has passed since the discovery and description of this gene until today, there are still few studies on the impact of its variants on world populations. Proof of this are clinical data from banks such as ClinVar—National Center for Biotechnology Information; BraVE, LovD, Beacon—Brazilian Initiative on Precision Medicine—that need to be updated [10].

The study of the distribution and prevalence of genetic variants of certain genes in mixed populations is particularly important. They can indicate patterns of illness according to the expressed genotype and infer an individual's predisposition to certain diseases [11–13]. It is important to focus on viral hepatitis which, according to the World Health Organization, varies in prevalence across continents due to the mode of transmission, existing viral genotypes, available prophylactic and therapeutic measures [14,15].

We add here, as a factor in this variability, the heterogeneous genotype distribution of genes such as *IFNL4*, especially in infections by hepatitis B (HBV) and C (HCV) viruses [16–18]. These diseases have a high distribution in African countries, East and Southeast Asia and European countries, and we observed that the populations of these countries have varying genotypes for the *IFNL4* gene [19–21].

This variability is often caused by a single allele, which can condition a better or worse resolution of viral infections and therapeutic responses [22–24]. As an example, we have the rs12979860-C/T variant in which the allelic replacement of T by C conditioned sustained virological response (SVR) in European and Asian populations, in addition to a better response to treatments with sofosbuvir [25–27].

In this sense, it is understood that investigations of genetic variants such as single nucleotide polymorphisms (SNPs) and insertion or deletion (INDEL) allow the discovery and characterization of new molecular markers, with clinical applications functioning in prognosis or even in therapeutic decisions for some population groups.

Therefore, our study aimed to characterize the molecular profile of the *IFNL4* gene in a mixed and native population from the Brazilian Amazon and compare these findings with data from the general Brazilian population described in the Online Archive of Brazilian Mutations (ABraOM), and world populations available in the 1000 Genomes Project and Genome Aggregation Database (gnomAD).

2. Materials and Methods

2.1. Identification of Variants

The first stage of this research was an exploratory search in the Ensembl (<https://www.ensembl.org/index.html>, accessed on 21 April 2023) database, which uses data from the 1000 Genomes Project, and GnomAD (<https://gnomad.broadinstitute.org/>, accessed on 22 April 2023). In these databases, we admitted the Minor Allele Frequency (MAF) on the most frequent variants (MAF > 0.05) of the *IFNL4* gene and these were investigated in world populations that could serve as a reference for Brazilian populations, especially Native Americans and mixed-race people from the Northern region of Brazil.

For comparisons between the database samples and the study-elected populations, we are only considering the exon variants of this gene since we aim to investigate exome data from mixed-race and native Amazonian populations. For specifics about the Brazilian population, the source database was ABraOM (<https://abraom.ib.usp.br>, accessed on 22 April 2023).

2.2. Data Analysis

This section of the study was carried out in two different ways, depending on the data bank considered. For Ensembl, accepting that the information available stems from 1000genomes and GnomAD, carrying a worldwide perspective, the steps were laid out as: (I) Aiming for an accurate investigation and comparison of variants that were found in Amazonian populations, the data acquisition phase was designed to include Minor Allele Frequencies (MAF) 5%; (II) using the search query “IFNL4” and selecting the matching result (gene symbol report and Hugo Gene Nomenclature Committee (HGNC) ID: 44480); (III) choosing “variant table” and adjusting the MAF upwards, so the results would depict the percentages from highest to lowest; (IV) organizing a primary table on Google Docs with columns detailing populations in order of: variant (rs), all (global percentages), AFR (African Population), AMR (American Population), EAS (East Asian Population), EUR (European Population), SAS (South Asian Population); (V) from the data generated, a second organization was made on Google Spreadsheets detailing the following: chromosome, position, reference nucleotide, variant nucleotide, amino acid and protein changes, variant kind and allelic information—number, count, frequency, homozygous or heterozygous. The study condensed these and added information from our native population as well.

When considering ABraOM, we had a more local approach since this data bank only considers data generated from Brazilian populations (Data Collection from São Paulo, SP, Brazil). For this database, the steps were: (I) opening the website, selecting SA-BE-WGS-1171 (hg38) and searching for “IFNL4”; (II) selecting the highest frequency option for the variant order display; (III) exporting the data to a Google spreadsheet considering: chromosome, position, reference and variant nucleotides, category, consequence predictor (if any, using tiers of intron, exon, substitutions or unknown) rs ID and their population frequencies (same population order as the worldwide banks, with the rs information also being crossed with NCBI references), homozygous and heterozygous numbers, allele number, count and frequency were sorted highest to lowest. All data were collected and organized between April and July 2023.

2.3. Study Populations

This study uses population data described by Cohen-Paes and collaborators in 2022 [28], which includes 63 Indigenous people, representatives of the 12 Amazonian ethnicities from the northern region of Brazil. They are from Asurini do Xingu, Arara, Araweté, Asurini do Tocantins, Awa-Guajá, Kayapó/Xikrin, Zo'é, Wajãpi, Karipuna, Phurere, Munduruku and Yudjá/Juruna tribes and their ancestry was confirmed by their exomes. Table 1 summarizes demographical information about these groups and Figure 1 presents their location in Brazil.

For our comparative analysis, the Indigenous groups were called Native American Population (NAT). Furthermore, this study includes 89 mixed-race individuals from the metropolitan region of Belém (PA), which includes the cities of Belém, Ananindeua, Marituba, Benevides, Santa Barbara, Santa Izabel, Castanhal and Barcarena; referred in this work by the acronym MISC.

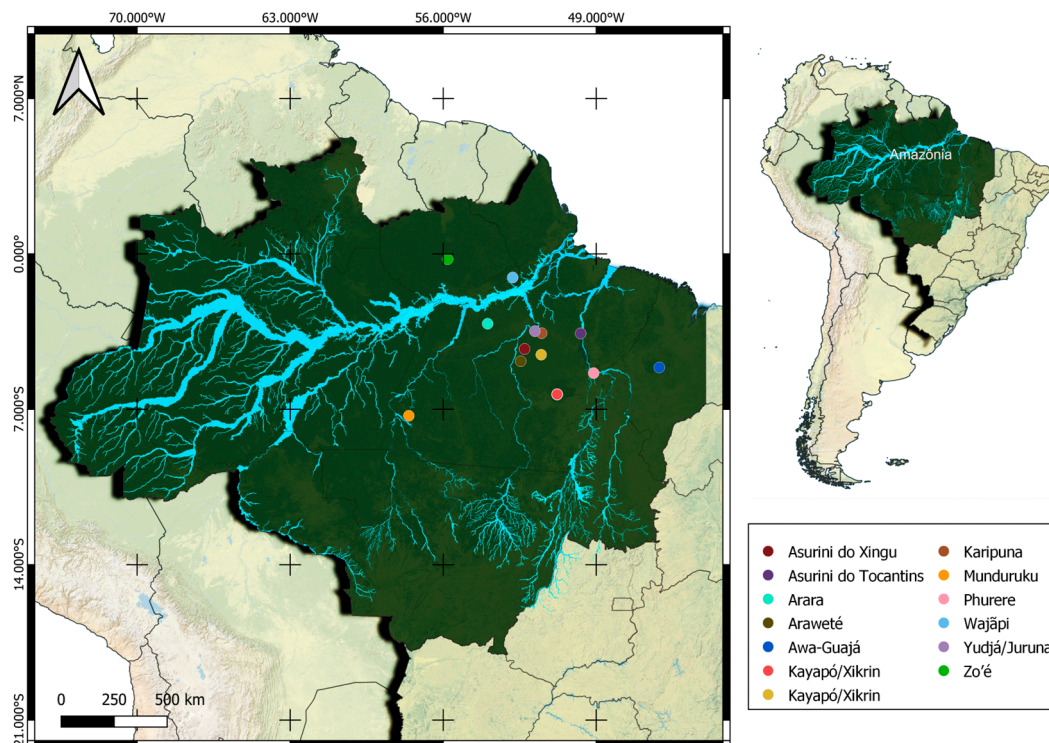
The study was approved by the National Research Ethics Committee and the Research Ethics Committee of the Tropical Medicine Center of the Federal University of Pará (CAE: 20654313.6.0000.5172). The participation of all individuals was conditioned on acceptance and signing of the informed consent form (ICF).

The data were compared to samples present in the databases 1000 Genomes (<http://www.1000genomes.org>, accessed on 19 April 2023) and GnomAD (<https://gnomad.broadinstitute.org/>, accessed on 19 April 2023), which are organized by presenting five continental populations including African (AFR), East Asian (EAS), South Asian (SAS), European (EUR), and American (AMR) individuals. Comparisons to the Brazilian population were obtained from the ABraOM database (<https://abraom.ib.usp.br/>, accessed on 19 April 2023).

Table 1. Populational groups demographics including the tribes, their location, number of habitants and data source.

Tribe	Location	State	Population (N)	Source
Asurini do Xingu	Altamira, Senador José Porfírio	Pará	182	SIASI
Arara/Arara do Iriri	Altamira, Brasil novo, Medicilância, Uruará	Pará	298	SIASI
Araweté	Altamira, São Félix do Xingu e Senador José Porfírio	Pará	467	SIASI
Asurini do Tocantins	Baião e Tucuruí	Pará	565	SIASI
Awa-Guajá	Arariboia	Maranhão	520	[29]
Kayapó/Xikrin	IL Bacajá (Altamira, Anapu, São Félix do Xingu, Senador José Porfírio)	Pará	746	FUNAI
	IL Xikrin do Cateté (Água Azu do Norte, Marabá, Parauapebas)	Pará	1183	SIASI
Zo'é	Óbidos	Pará	330	IEPE
Wajãpi	Laranjal do Jari, Mazagão, Pedra Branca do Amapari	Pará	1454	DSEI/AP
Kariyuna	TI Galibe (Oiapoque)	Amapá	151	SIASI
	TI Juminá (Oiapoque)	Amapá	291	IEPE
	Uaçá I e II (Oiapoque)	Amapá	4462	FUNAI
Phurere	Marabá	Pará		
Munduruku	Itaituba e Jacareacanga	Pará	6518	SIASI
Yudjá/Juruna	Vitória do Xingu	Pará	95	SIASI

Acronyms: IL—Indigenous land, SIASI—Indigenous Health Care Information System, FUNAI—National Foundation of Indigenous Peoples, DSEI—Special Indigenous Health Districts, IEPE—Special Indigenous Health Districts.

**Figure 1.** Location of Amazonian tribes present in this study. Image made on QGIS 3.28.0 software.

2.4. DNA Extraction and Exome Library Preparation

DNA extraction occurred using the phenol-chloroform method [30]. DNA quality was assessed using the Nanodrop-8000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA). Library preparation used the Nextera Rapid Capture Exome (Illumina[®], San Diego, CA, USA) and SureSelect Human All Exon V6 (Agilent technologies, Santa Clara, CA, USA), following the manufacturer's recommendations. Sequencing took place on the NextSeq 500[®] platform (Illumina[®], San Diego, CA, USA) with use of the NextSeq 500 High-output v2 300-cycle Kit (Illumina[®], San Diego, CA, USA).

2.5. Bioinformatics Analysis

The resulting reads in FASTQ format were analyzed (FastQC v.0.11 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 25 June 2023) and low-quality reads were eliminated (fastx_tools v.0.13 http://hannonlab.cshl.edu/fastx_toolkit/, accessed on 25 June 2023). The sequences were aligned with the reference genome (GRCh38) using the BWA v.0.7 tool (<http://bio-bwa.sourceforge.net/>, accessed on 25 June 2023).

The sequencing reads were cut to Illumina adapters and filtered using Trimmomatic v.0.36 [31] the remaining reads were aligned to the hg19 human genome reference, using BWA MEM v.0.7 [32,33]. The PCR duplicates were removed via Samblaster v.0.1 [34] and the resulting mapped reads were classified and indexed making use of Samtools v1.8 [35] and Sambamba v.0.6 [36]. Lastly, the quality of the mapped bases was assessed and recalibrated with GATK v.4.0.0 BaseRecalibrator and ApplyBQSR.

Allelic variants were noted in ViVa[®] (Viewer of Variants, Natal, RN, Brazil). Markers in the *IFNL4* gene were selected from the generated coverage (fastx_tools v.0.13-http://hannonlab.cshl.edu/fastx_toolkit (accessed on 25 June 2023)).

2.6. Statistical Analysis

The allelic frequencies were compared to verify if the mixed-race (MISC) and native (NAT) Amazonian populations present some sort of Allelic Distribution of the *IFNL4* gene variants, similar to those observed in Continental and Brazilian populational groups. For that, we used the Fisher's exact test. The results were considered statistically significant when the p value was less than 0.05 ($p \leq 0.05$). The Benjamini-Hochberg False Discovery Rate (FDR) [37] test was used as a statistic correction measure applied to all the multiple tests listed. All statistical analyses were performed using the R Studio v.3.5.1 program (R Foundation for Statistical Computing, Vienna, Austria).

The LD Link tool [38] was employed to extract data referring to variant linkage disequilibrium from the ones found both in this study and in public databases, such as the 1000 Genomes Project. The results are shown through a media estimate based on the proximity between two variants (D') in the investigated populations.

3. Results

In exploratory research, we identified 10 variants in which the altered allele was frequent in at least 10% of the populations described by the 1000 Genomes Project (MAF > 10%) and another four variants with a frequency of 5% (MAF > 5%). Among them, two variants identified in mixed-race and native populations in the northern region of Brazil had a frequency greater than 10%, and one had a frequency greater than 5%. Table 2 summarizes the characteristics of the collected variants, namely: location of the variant on the chromosome, its identification, altered alleles, variant type, and its consequence.

Table 2. Characterization of variants of the *IFNL4* gene investigated in distinct groups: African, American, East Asian, European, South Asian, Brazilian from São Paulo, mixed-race, and Native American populations from the Northern region of Brazil.

Genomic Position	SNP Identifier	Nucleotide	Variant	Consequence
19:39246873	rs12979731	T > C	SNV	3'UTR variant
19:39246936	rs370209610	T > C	SNV	3'UTR variant
19:39246970	rs11882871	G > A	SNV	3'UTR variant
19:39247226	rs12971396	C > G	SNV	3'UTR variant
19:39247247	rs137902769	A > T	SNV	3'UTR variant
19:39247389	rs73555604	C > T	SNV	missense variant
19:39247677	rs111531283	C > A	SNV	intron variant
19:39247938	rs117648444	G > A	SNV	missense variant
19:39248147	rs12979860	T > C	SNV	intron variant
19:39248489	rs4803221	C > G	SNV	missense variant
19:39248514	rs11322783	T > TT	INDEL	frameshift variant
19:39248515	rs74597329	G > T	SNV	missense variant
19:39248713	rs4803222	G > C	SNV	5'UTR variant
19:39737578	rs570739705	C > G	SNV	3'UTR variant

Exome analysis of samples from mixed-race and native Amazonians showed three of these variants being 2 SNVs (single nucleotide variant) and 1 INDEL (Insertion-Deletion). The SNV variants were rs74597329-G/T, rs4803221-C/G and are classified as moderate impact and the INDEL type variant –rs11322783-ΔT/TT–of high impact as it is a frameshift mutation.

Table 3 below presents the allelic frequencies of the *IFNL4* gene variants in each of the five population groups described by GnomAD v3.1.2—Africa, Americas, East Asia, Europe, and South Asia, while Table 4 presents the frequencies observed in the 1000 Genomes Project in the same groups—the frequencies are presented in decimal number format, in both tables. They also contain the frequencies of São Paulo population obtained from the ABraOM database, serving as a basis for comparison of the frequencies observed in mixed-race and Native American populations in the Northern region of Brazil.

Table 3. Frequencies of *IFNL4* gene variants investigated in African (AFR), American (AMR), East Asian (EAS), European (EUR), South Asian (SAS) from GnomAD, Brazilian population from Online Archive of Brazilian Mutations (ABraOm), mixed-race (MISC) and Native American populations from the Northern region of Brazil (NAT).

SNP ID	Alleles	AFR	AMR	EAS	EUR	SAS	ABraOm	MISC	NAT
rs74597329 *	T	0.364	0.621	0.932	0.683	0.766	0.624	0.744	0.838
	G	0.636	0.379	0.068	0.317	0.234	0.376	0.256	0.162
rs4803221 *	C	0.774	0.721	0.936	0.797	0.845	0.78	0.832	0.937
	G	0.226	0.279	0.064	0.203	0.155	0.22	0.169	0.064
rs73555604	C	0.801	0.977	0.999	0.98	0.984	0.952		
	T	0.199	0.023	0.001	0.02	0.016	0.048		
rs117648444	G	0.932	0.937	0.995	0.906	0.95	0.921		
	A	0.068	0.063	0.005	0.094	0.05	0.079		
rs11322783 *	TT	0.364	0.621	0.932	0.683	0.766	0.624	0.744	0.838
	T-	0.636	0.379	0.068	0.317	0.234	0.376	0.256	0.162

Shaded lines represent the frequencies of major allele and non-shaded represent minor allele. * Represents the SNVs evaluated in native and mixed populations.

Table 4. Frequencies of variants of the *IFNL4* gene investigated in African (AFR), American (AMR), East Asian (EAS), European (EUR), South Asian (SAS) populations from the 1000 Genomes Project, Brazilian population from Online Archive of Brazilian Mutations (ABraOm), mixed-race (MISC) and Native Americans from the Northern region of Brazil (NAT).

SNP ID	Alleles	AFR	AMR	EAS	EUR	SAS	ABraOm	MISC	NAT
rs74597329 *	T	0.293	0.597	0.92	0.688	0.758	0.624	0.744	0.838
	G	0.707	0.403	0.08	0.312	0.242	0.376	0.256	0.162
rs4803221 *	C	0.837	0.7	0.925	0.828	0.834	0.78	0.832	0.937
	G	0.163	0.3	0.075	0.172	0.166	0.22	0.169	0.064
rs73555604	C	0.737	0.976	1	0.983	0.979	0.952		
	T	0.263	0.024		0.017	0.021	0.048		
rs117648444	G	0.925	0.935	0.995	0.882	0.956	0.921		
	A	0.075	0.065	0.005	0.118	0.044	0.079		
rs11322783 *	T	0.707	0.403	0.081	0.311	0.239	0.376	0.256	0.162
	TT	0.293	0.597	0.919	0.689	0.761	0.624	0.744	0.838

Shaded lines represent the frequencies of major allele and non-shaded represent minor allele. * Represents the SNVs evaluated in native and mixed populations.

We observed subtle differences between data from GnomAD and the 1000 Genomes Project (Tables 5 and 6). Pairwise comparison with data from the 1000 Genomes Project showed that the native group (NAT) differs from AFR (p -value: <0.0001), AMR (p -value: 0.0097), and São Paulo—AbraOm (p -value: 0.0097) in rs74597329 and rs11322783 variants. In the rs4803221 variant, NAT was also different from AFR (p -value: 0.0164) and AMR (p -value: 0.0031) groups. The mixed group (MISC) presented differences when compared to AFR (p -value: <0.0001) and EAS (p -value: 0.0031) for rs74597329 and rs11322783 variants. The mixed group (MISC) presented differences when compared to AFR (p -value: <0.0001) and EAS (p -value: 0.0031) for rs74597329 and rs11322783 variants.

Table 5. Pairwise comparison (p -value) of allelic frequencies in Native, Mixed-population and with each one of the five continental populations from the 1000 Genomes Project database for *IFNL4* gene variants investigated.

	rs74597329	rs4803221	rs11322783
AFR vs. MISC	<0.0001	0.4549	<0.0001
AMR vs. MISC	0.1395	0.1395	0.1395
EAS vs. MISC	0.0031	0.0585	0.0031
EUR vs. MISC	0.4976	0.7734	0.4976
SAS vs. MISC	0.7734	1	0.7734
ABraOm vs. MISC	0.1395	0.5294	0.1395
MISC vs. NAT	0.2281	0.1395	0.2281
AFR vs. NAT	<0.0001	0.0164	<0.0001
AMR vs. NAT	0.0097	0.0031	0.0097
EAS vs. NAT	0.1659	1	0.1659
EUR vs. NAT	0.0638	0.0585	0.0638
SAS vs. NAT	0.4131	0.1395	0.4131
ABraOm vs. NAT	0.0097	0.025	0.0097

Shaded cells represent significant differences between two populations (p -value < 0.05).

Table 6 refers to pairwise comparisons between our local data and data from GnomAD. It also shows differences between NAT, AFR, AMR and Southern Brazilian population in rs74597329 and rs11322783 variants. AMR and AFR groups also differ from NAT for rs4803221 variant (p -value: 0.016 and 0.003). Moreover, examining the mixed population, we saw that either AFR (p -value < 0.0001) and EAS (p -value = 0.0003) differs from it for rs74597329 and rs11322783 variants. The differences can also be noticed in the plot of Supplementary Materials S1. The analysis of connection imbalance between the variants observed in the present study shows a strong imbalance of these variants ($r^2 = 0.990$; $D' = 1$) in all populations already observed and, therefore, they can be analyzed as a haplotype [38].

Table 6. Pairwise comparison (*p*-value) of allelic frequencies in Native, Mixed-population and with each one of the five continental populations from GnomAD database for *IFNL4* gene variants investigated.

	rs74597329	rs4803221	rs11322783
AFR vs. MISC	<0.0001	0.4579	<0.0001
AMR vs. MISC	0.1417	0.1417	0.1417
EAS vs. MISC	0.0003	0.057	0.0003
EUR vs. MISC	0.4999	0.7745	0.4999
SAS vs. MISC	0.7745	1	0.7745
ABraOm vs. MISC	0.1417	0.5313	0.1417
MISC vs. NAT	0.2303	0.1417	0.2303
AFR vs. NAT	<0.0001	0.016	<0.0001
AMR vs. NAT	0.0095	0.003	0.0095
EAS vs. NAT	0.1679	1	0.1679
EUR vs. NAT	0.0622	0.057	0.0622
SAS vs. NAT	0.4162	0.1417	0.4162
ABraOm vs. NAT	0.0095	0.0244	0.0095

Shaded cells represent significant differences between two populations (*p*-value < 0.05).

4. Discussion

Modern-day Latin American people are the result of a complex process of mixing different genetic profiles [39]. The studies on genetics are still scarce, especially on Amazonian natives, which makes it difficult to know the specifics of their singular profile in comparison to other groups. Many authors highlight that allelic variance in genes present amongst different ethnic groups may show associations with significant consequences for human illnesses [13,40–42].

This can be clearly exemplified by the *IFNL4* gene, whose variants may regulate IFN- λ 4 expression and functionality. Interferon lambda proteins can activate the JAK-STAT pathway, leading to a positive expression of genes with antiviral properties aiming to eliminate pathogens, such as hepatitis C virus. This gene presents interesting characteristics due to a polymorphism that regulates IFN- λ 4 expression; it is composed of the rs74597329 (IFNL4-G/T) and rs11322783 (IFNL4- Δ /T) variants, which are in complete connection imbalance and, due to this, are analyzed together [5,10,28].

Those variants behave as a dinucleotide variation, named rs368234815 (Δ G/TT). In practical terms, this means that individuals will always have the IFNL4- Δ G (rs11322783 Δ /rs74597329G) and IFNL4-TT (rs11322783T/rs74597329T) haplotypes [5,28]. IFNL4- Δ G haplotype allows the formation and expression of fully functional IFN- λ 4 (p179) and the IFNL4-TT haplotype alters the reading frame of the protein, creating inactive forms (p124 and p143) [10]. A paradoxical fact is discussed in the Fang [6] and Vergara [43] studies, where they show that individuals not expressing IFN- λ 4 (haplotype IFNL4-TT) have higher clearance rates and better results from treatments against HCV infection.

Our study shows that the IFNL4-TT haplotype is predominant in Amazonian natives as well as in mixed and Asian populations, reaching values between 74% and 93%. The mixed group presents a profile similar to American (AMR), European (EUR), Southern Asian (SAS) and São Paulo (ABraOm) groups. It was an expected aspect since the mixed population of the study was formed by founding groups such as Europeans and Asians [44].

The differences observed when comparing the frequencies of African people to our populations of study happen because the AFR group has a differing prevalence of the Δ G haplotype when compared to all the other populations. Some authors discuss how this strong selective pressure for the IFNL4-TT haplotype may have arisen before the migration of human populations out of Africa. Due to other factors still unknown, IFNL4-TT became stronger in European and Asian groups but not in the original population. In other words, the downregulation of this protein is beneficial in many populations. [6,45]

A proposal to justify different results in American populations is given by [46]. He proposes that the genetic variations of the native populations of North, Central and South

America are different due to the colonialism process and that the South American natives are even more genetically isolated from other American populations. Therefore, frequencies observed in Brazilian population groups are the result of different formation processes among our people [47].

The variant rs4803221 has controversial results. While [48,49] show G, the minor allele, as a poor predictor of viral clearance in individuals undergoing therapy with Pegylated Interferon- α (pegIFN α), and DAAs [43] discuss that this allelic variance did not impact the elimination or remaining of viral infection in Africans.

In our study, the rs4803221 variant did not show any statistically differing results on the comparison between the NAT and MISC samples and the database-checked populations. Low frequencies of minor allele (G) were observed on the NAT and EAS samples, which is expected when we evaluate the formation process of American populations, which considers American ancestry to have come from the migration of Asians by the Bering Strait [50]. However, the frequencies of the minor allele of this variant in the MISC sample are similar to those of the groups that formed it: AFR, AMR and EUR.

The SNPs rs73555604 and rs117648444 collected in the exome of the *IFNL4* gene are shown in worldwide populations (GnomAd and 1000 Genomes) and in southeastern Brazil (ABraOm), but they are not present in the samples of our study. They are missense-type SNVs, and their absence in the sample of mixed-race and native Amazonians is explained by their minor frequencies of MAF in most continental populations, except for the AFR population, in which the frequencies of the lowest allele are slightly larger.

A better understanding of the genetic profile of our population is important to validate the existing hypotheses of distinct patterns in disease responses, infections, and treatments in different people. The importance of characterizing the genetics of native populations and people who have suffered intense processes of ethnic mixing is justified by essentially helping to understand illness processes that are, sometimes, unclear.

Other studies follow ours to elucidate (1) the distribution of these variants in patients with hepatitis C and (2) the direct consequences of these polymorphisms on both the HCV clearance and the efficiency of drug treatment in populations from the north of Brazil. This information will contribute knowledge and base the development of public policies aiming for health improvement in many groups across the country.

5. Conclusions

The results indicate that the variants rs74597329, rs11322783 and rs4803221 of the *IFNL4* gene evaluated in mixed-race and native populations from the northern region of Brazil have frequencies resembling those in world populations, especially Asian and European ones. In practical terms, these results suggest that world populations played a significant role in the formation of the Brazilian population. Therefore, this study aid further future public policies associated with the treatment of viral infections in these populational groups.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes14112075/s1>.

Author Contributions: Conceptualization, S.E.B.d.S. and G.d.L.C.C.; Data curation, J.F.G.; Formal analysis, P.P.d.A., S.E.B.d.S., F.C.M. and G.d.L.C.C.; Investigation, C.C.A., L.D.M. and Á.A.F.A.; Methodology, F.C.M., G.d.L.C.C., J.F.G., P.P.d.A. and S.E.B.d.S.; Supervision, G.d.L.C.C.; Validation, J.F.G. and P.P.d.A.; Visualization, L.D.M., Á.A.F.A. and J.F.G.; Writing—original draft, C.C.A.; Writing—review and editing, C.C.A., L.D.M. and G.d.L.C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is part of the Rede de Pesquisa em Genômica Populacional Humana (Biocomputacional Protocol No. 3381/2013/CAPES), FAPESPA (Amazon Foundation for the Support of Studies and Research) and Pro-rector of Research and Post-graduation (Proresp) of the Federal University of Para supported this work. The authors C.C.A, L.D.M and Á.A.F.A were also contemplated with governmental scholarships according to the ID numbers 130241/2023-7 from CNPq (Council for

Scientific and Technological Development), 88887.842950/2023-00 and 88887.940110/2023-00 from CAPES (Coordination for the Improvement of Higher Education Personnel), respectively.

Institutional Review Board Statement: The study was approved by the National Research Ethics Committee and the Research Ethics Committee of the Tropical Medicine Center of the Federal University of Pará (CAE: 20654313.6.0000.5172) on 18 December 2018.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article and supplementary materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Syedbasha, M.; Egli, A. Interferon Lambda: Modulating Immunity in Infectious Diseases. *Front. Immunol.* **2017**, *8*, 119.
2. Hemann, E.A.; Gale, M.; Savan, R. Interferon Lambda Genetics and Biology in Regulation of Viral Control. *Front. Immunol.* **2017**, *8*, 1707. [CrossRef]
3. Ye, L.; Schnepf, D.; Staeheli, P. Interferon- λ orchestrates innate and adaptive mucosal immune responses. *Nat. Rev. Immunol.* **2019**, *19*, 614–625. [CrossRef]
4. Bruening, J.; Weigel, B.; Gerold, G. The Role of Type III Interferons in Hepatitis C Virus Infection and Therapy. *J. Immunol. Res.* **2017**, *2017*, 7232361. [CrossRef]
5. O'Brien, T.R.; Prokunina-Olsson, L.; Donnelly, R.P. IFN- λ 4: The Paradoxical New Member of the Interferon Lambda Family. *J. Interferon Cytokine Res.* **2014**, *34*, 829–838. [CrossRef]
6. Fang, M.Z.; Jackson, S.S.; O'Brien, T.R. IFNL4: Notable variants and associated phenotypes. *Gene* **2020**, *730*, 144289. [CrossRef]
7. Rugwizangoga, B.; Andersson, M.E.; Kabayiza, J.C.; Nilsson, M.S.; Ármannsdóttir, B.; Aurelius, J.; Nilsson, S.; Hellstrand, K.; Lindh, M.; Martner, A. IFNL4 Genotypes Predict Clearance of RNA Viruses in Rwandan Children with Upper Respiratory Tract Infections. *Front. Cell. Infect. Microbiol.* **2019**, *9*, 340. [CrossRef]
8. Bhushan, A.; Chinnaswamy, S. Identifying causal variants at the interferon lambda locus in case-control studies: Utilizing non-synonymous variant rs117648444 to probe the role of IFN- λ 4. *Gene* **2018**, *664*, 168–180. [CrossRef]
9. Prokunina-Olsson, L. Genetics of the Human Interferon Lambda Region. *J. Interferon Cytokine Res.* **2019**, *39*, 599–608. [CrossRef]
10. National Center for Biotechnology Information. ClinVar; [VCV000225949.4]. Available online: <https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000225949.4> (accessed on 3 May 2023).
11. Pinto, P.; Salgado, C.; Santos, N.P.C.; Santos, S.; Ribeiro-dos-Santos, Â. Influence of Genetic Ancestry on INDEL Markers of NFK β 1, CASP8, PAR1, IL4 and CYP19A1 Genes in Leprosy Patients. *PLOS Neglected Trop. Dis.* **2015**, *9*, e0004050. [CrossRef]
12. Leal, D.F.d.V.B.; da Silva, M.N.S.; Fernandes, D.C.R.d.O.; Rodrigues, J.C.G.; Barros, M.C.d.C.; Pinto, P.D.d.o.C.; Pastana, L.F.; Silva, C.A.; Fernandes, M.R.; Assumpção, P.P.; et al. Amerindian genetic ancestry as a risk factor for tuberculosis in an amazonian population. *PLoS ONE* **2020**, *15*, e0236033.
13. Carvalho, D.C.; Wanderley, A.V.; dos Santos, A.M.R.; Moreira, F.C.; de Sá, R.B.A.; Fernandes, M.R.; Modesto, A.A.; de Souza, T.P.; Cohen-Paes, A.; Leitão, L.P.; et al. Characterization of pharmacogenetic markers related to Acute Lymphoblastic Leukemia toxicity in Amazonian native Americans population. *Sci. Rep.* **2020**, *10*, 10292. [CrossRef] [PubMed]
14. World Health Organization (WHO). Hepatitis. Health Topics. Available online: https://www.who.int/health-topics/hepatitis#tab=tab_1 (accessed on 4 May 2023).
15. Cuautle-Rodríguez, P.; Llerena, A.; Molina-Guarneros, J. Present status, and perspective of pharmacogenetics in Mexico. *Drug Metab. Drug Interact.* **2014**, *29*, 37–45.
16. Natri, A.C.S.S.; Malta, F.M.; Diniz, M.A.; Yoshino, A.; Abe-Sandes, K.; dos Santos, S.E.B.; de Castro Lyra, A.; Carrilho, F.J.; Pinho, J.R. Association of IFNL3 and IFNL4 polymorphisms with hepatitis C virus infection in a population from southeastern Brazil. *Arch. Virol.* **2016**, *161*, 1477–1484. [CrossRef]
17. Ma, N.; Zhang, X.; Yang, L.; Zhou, J.; Liu, W.; Gao, X.; Yu, F.; Zheng, W.; Ding, S.; Gao, P.; et al. Role of Functional IFNL4, IFNL1, IFNA, IFNAR2 Polymorphisms in Hepatitis B virus-related liver disease in Han Chinese population. *J. Viral Hepat.* **2017**, *25*, 306–313. [CrossRef]
18. Gonzalez-Aldaco, K.; Rebello Pinho, J.R.; Roman, S.; Gleyzer, K.; Fierro, N.A.; Oyakawa, L.; Ramos-Lopez, O.; Ferraz Santana, R.A.; Sitnik, R.; Panduro, A.; et al. Association with Spontaneous Hepatitis C Viral Clearance and Genetic Differentiation of IL28B/IFNL4 Haplotypes in Populations from Mexico. *PLoS ONE* **2016**, *11*, e0146258.
19. Jefferies, M.; Rauff, B.; Rashid, H.; Lam, T.; Rafiq, S. Update on global epidemiology of viral hepatitis and preventive strategies. *World J. Clin. Cases* **2018**, *6*, 589–599. Available online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6232563/> (accessed on 4 April 2023).
20. Razavi, H. Global Epidemiology of Viral Hepatitis. *Gastroenterol. Clin. N. Am.* **2020**, *49*, 179–189.
21. Sant'Anna, T.B.; Araujo, N.M. Hepatitis B Virus Genotype D: An Overview of Molecular Epidemiology, Evolutionary History, and Clinical Characteristics. *Microorganisms* **2023**, *11*, 1101.

22. Pedernana, V.; Irving, W.L.; Barnes, E.; McLauchlan, J.; Spencer, C.C.A. Impact of IFNL4 Genetic Variants on Sustained Virologic Response and Viremia in Hepatitis C Virus Genotype 3 Patients. *J. Interferon Cytokine Res.* **2019**, *39*, 642–649. [[CrossRef](#)]
23. Galmozzi, E.; Facchetti, F.; Grossi, G.; Loglio, A.; Viganò, M.; Lunghi, G.; Colombo, M.; Lampertico, P. IFNL4 rs368234815 and rs117648444 variants predict off-treatment HBsAg seroclearance in IFN-treated HBeAg-negative chronic hepatitis B patients. *Liver Int.* **2017**, *38*, 417–423. [[CrossRef](#)] [[PubMed](#)]
24. De Re, V.; Tornesello, M.L.; De Zorzi, M.; Caggiari, L.; Pezzuto, F.; Leone, P.; Racanelli, V.; Lauletta, G.; Zanussi, S.; Repetto, O.; et al. PDCD1 and IFNL4 genetic variants and risk of developing hepatitis C virus-related diseases. *Liver Int.* **2020**, *41*, 133–149.
25. Luo, Y.; Jin, C.; Ling, Z.; Mou, X.; Zhang, Q.; Xiang, C. Association study of IL28B: rs12979860 and rs8099917 polymorphisms with SVR in patients infected with chronic HCV genotype 1 to PEG-INF/RBV therapy using systematic meta-analysis. *Gene* **2013**, *513*, 292–296. [[CrossRef](#)]
26. Ragheb, M.M.; Nemr, N.A.; Kishk, R.M.; Mandour, M.F.; Abdou, M.M.; Matsuura, K.; Watanabe, T.; Tanaka, Y. Strong prediction of virological response to combination therapy by *IL28B* gene variants rs12979860 and rs8099917 in chronic hepatitis C genotype 4. *Liver Int.* **2014**, *34*, 890–895. [[CrossRef](#)]
27. Prokunina-Olsson, L.; Muchmore, B.; Tang, W.; Pfeiffer, R.M.; Park, H.; Dickensheets, H.; Hergott, D.; Porter-Gill, P.; Mumy, A.; Kohaar, I.; et al. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat. Genet.* **2013**, *45*, 164–171. [[CrossRef](#)]
28. Cohen-Paes, A.d.N.; de Carvalho, D.C.; Pastana, L.F.; Dobbin, E.A.F.; Moreira, F.C.; de Souza, T.P.; Fernandes, M.R.; Leal, D.F.; de Sá, R.B.; de Alcântara, A.L.; et al. Characterization of PCLO Gene in Amazonian Native American Populations. *Genes* **2022**, *13*, 499. [[CrossRef](#)]
29. Garcia, U. *Crônicas de caça e Criação*; Hedra: São Paulo, Brazil, 2022.
30. Sambrook, J.; Green, M.R. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2012.
31. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
32. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
33. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
34. Faust, G.G.; Hall, I.M. SAMBLASTER: Fast duplicate marking and structural variant read extraction. *Bioinformatics* **2014**, *30*, 2503–2505. [[CrossRef](#)]
35. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078. [[CrossRef](#)] [[PubMed](#)]
36. Tarasov, A.; Vilella, A.J.; Cuppen, E.; Nijman, I.J.; Prins, P. Sambamba: Fast processing of NGS alignment formats. *Bioinformatics* **2015**, *31*, 2032–2034. [[CrossRef](#)]
37. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [[CrossRef](#)]
38. Machiela, M.J.; Chanock, S.J. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants: Fig. 1. *Bioinformatics* **2015**, *31*, 3555–3557. [[CrossRef](#)]
39. Salzano, F.M. Genetic diversity of South American human populations at the DNA and protein levels. *J. Exp. Zool.* **1998**, *282*, 157–163. [[CrossRef](#)]
40. Rodrigues, J.C.G.; Fernandes, M.R.; Guerreiro, J.F.; da Silva, A.L.D.C.; Ribeiro-Dos-Santos, Â.; Santos, S.; Santos, N.P. Polymorphisms of ADME-related genes and their implications for drug safety and efficacy in Amazonian Amerindians. *Sci. Rep.* **2019**, *9*, 7201. [[CrossRef](#)]
41. Rodrigues, J.C.G.; de Souza, T.P.; Pastana, L.F.; dos Santos, A.M.R.; Fernandes, M.R.; Pinto, P.; Wanderley, A.V.; de Souza, S.J.; Kroll, J.E.; Pereira, A.L.; et al. Identification of NUDT15 gene variants in Amazonian Amerindians and admixed individuals from northern Brazil. *PLoS ONE* **2020**, *15*, e0231651. [[CrossRef](#)]
42. Dobbin, E.A.F.; Medeiros, J.A.G.; Costa, M.S.C.R.; Rodrigues, J.C.G.; Guerreiro, J.F.; Kroll, J.E.; Souza, S.J.; de Assumpção, P.P.; Ribeiro-dos-Santos, Â.; Santos, S.E.; et al. Identification of Variants (rs11571707, rs144848, and rs11571769) in the BRCA2 Gene Associated with Hereditary Breast Cancer in Indigenous Populations of the Brazilian Amazon. *Genes* **2021**, *12*, 142. [[CrossRef](#)] [[PubMed](#)]
43. Vergara, C.; Thio, C.L.; Johnson, E.; Kral, A.H.; O'Brien, T.R.; Goedert, J.J.; Mangia, A.; Piazzolla, V.; Mehta, S.H.; Kirk, G.D.; et al. Multi-Ancestry Genome-Wide Association Study of Spontaneous Clearance of Hepatitis C Virus. *Gastroenterology* **2019**, *156*, 1496–1507. [[CrossRef](#)] [[PubMed](#)]
44. Santos, N.P.; Ribeiro-Rodrigues, E.M.; Ribeiro-Dos-Santos, K.; Pereira, R.; Gusmão, L.; Amorim, A.; Guerreiro, J.F.; Zago, M.A.; Matte, C.; Hutz, M.H.; et al. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum. Mutat.* **2010**, *31*, 184–190. [[CrossRef](#)]
45. Key, F.M.; Peter, B.; Dennis, M.Y.; Huerta-Sánchez, E.; Tang, W.; Prokunina-Olsson, L.; Nielsen, R.; Andrés, A.M. Selection on a Variant Associated with Improved Viral Clearance Drives Local, Adaptive Pseudogenization of Interferon Lambda 4 (IFNL4). *PLoS Genet.* **2014**, *10*, e1004681. [[CrossRef](#)] [[PubMed](#)]

46. Wang, S.; Lewis, C.M.; Jakobsson, M.; Ramachandran, S.; Ray, N.; Bedoya, G.; Rojas, W.; Parra, M.V.; A Molina, J.; Gallo, C.; et al. Genetic Variation and Population Structure in Native Americans. *PLoS Genet.* **2007**, *3*, e185. [[PubMed](#)]
47. Amador, M.A.; Cavalcante, G.C.; Santos, N.P.; Gusmão, L.; Guerreiro, J.F.; Ribeiro-dos-Santos, Â.; Santos, S. Distribution of allelic and genotypic frequencies of IL1A, IL4, NFKB1 and PAR1 variants in Native American, African, European and Brazilian populations. *BMC Res. Notes* **2016**, *9*, 101. [[CrossRef](#)] [[PubMed](#)]
48. Smith, K.R.; Suppiah, V.; O'Connor, K.; Berg, T.; Weltman, M.; Abate, M.L.; Spengler, U.; Bassendine, M.; Matthews, G.; Irving, W.L.; et al. Identification of improved IL28B SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort. *Genome Med.* **2011**, *3*, 57. [[CrossRef](#)] [[PubMed](#)]
49. Booth, D.R.; Ahlenstiel, G.; George, J. Pharmacogenomics of hepatitis C infections: Personalizing therapy. *Genome Med.* **2012**, *4*, 99. [[CrossRef](#)]
50. Skoglund, P.; Reich, D. A genomic view of the peopling of the Americas. *Curr. Opin. Genet. Dev.* **2016**, *41*, 27–35. [[CrossRef](#)]

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.