Prevalence of the Main Human Genetic Variants Related to Resistance to Malaria in a Population of the Colombian Pacific Coast

Diana Carolina Ortega
Universidad del Valle
Maria Paula Arango
Universidad del Valle
Sergio Cañón
Universidad del Valle
Heiber Cárdenas
Universidad del Valle
Ranulfo González
Universidad del Valle
Guillermo Barreto (guillermo.barreto@correounivalle.edu.co)
Universidad del Valle

Research Article

Keywords: Buenaventura, communities, malaria, age groups, protection genotypes

Posted Date: December 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3693432/v1

Additional Declarations: No competing interests reported.

Abstract

Background. Malaria is an endemic disease present in many areas of Colombia. In the city of Buenaventura (Colombian Pacific), there is a high incidence of malaria cases, with a high number of deaths due to this disease. Some genetic variants provide protection against malaria, as is the case for individuals heterozygous for some haemoglobin variants (HbS, HbC, and β-thalassemias), individuals homozygous and heterozygous for the A- variant of the G6PD gene, and individuals homozygous for the FYB^{ES} allele of the Duffy gene. The objective of this research was to establish the prevalence of these variants through molecular characterization in a representative sample of the population of the urban area of Buenaventura.

Methods. A total of 819 individuals selected randomly from each of the 12 communities of the city were included. The analysis at the molecular level was carried out using PCR-RFLP and allele-specific PCR. The contained data were subjected to descriptive, independence and regression analyses. These evaluations were carried out using Arlequin 3.5, SPSS 20 and R 3.4.1.

Results. Frequencies of 3.1%, 2.2%, 72.2%, 2.1%, 2.8%, and 11% were found for the resistance alleles HbS, HbC, Duffy, β-Thalassemia-29, β-thalassemia-88 and G6PD, respectively. For the Duffy gene, there was a higher frequency of the resistance genotype in the entire population, as well as a higher occurrence of the Duffy resistance genotype combination with the G6PD and HbS/C resistance genotypes. In addition, compared with other age groups, adolescents and young adults (13 to 26 years) presented the highest proportion of resistance genotypes. Likewise, compared with other communities, the communities of the insular zone of Buenaventura (1, 2, 4, 5) had the highest proportion of resistance genotypes. These data are important to take into account by health and prevention entities in the city because they reveal age groups and communities more susceptible to infection by malaria in the city of Buenaventura and groups prone to developing and/or propagating genes that can increase the prevalence of haemoglobinopathies in the population in the long term.

Introduction

Malaria is an infectious disease caused by an intracellular parasite of the genus *Plasmodium*, whose main vector is mosquitoes of the genus Anopheles [1], and is one of the most serious public health problems worldwide [2]. The Amazon region of Peru and the Pacific coast in Colombia are two of the areas most affected by malaria in the Americas [3], and Buenaventura, located on the Pacific coast, is a city with one of the highest number of register cases of malaria in the region [4].

In areas where malaria is endemic, high frequencies of genetic mutations that cause haemoglobinopathies or defects in erythrocytes have been found, and a protective effect of these mutations has been observed in the clinical development of malaria; studies show that the burden is reduced in affected erythrocytes, with the latter being phagocytosed to a greater extent than healthy erythrocytes, favouring parasite removal [5, 6, 7, 8, 9, 10, 11, 12, 13]. Among the main genotypes that confer resistance to or protection from malaria are haemoglobin variants, such as HbS (heterozygous genotype AS), HbC (genotype AC and CC) and heterozygotes for α and β thalassemias; the heterozygous, homozygous and hemizygous genotype of the A-

allele of the glucose 6-phosphate dehydrogenase (G6PD) gene; and the homozygous genotype of the FYB^{ES} allele of the Duffy gene [6, 7, 8, 14, 15].

In Colombia, haemoglobinopathies and erythrocyte defects prevail, with frequencies ranging from 2.4 to 70% being reported in different regions of the country. The results of population studies of some of these resistance variants have also been reported [16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33], but thus far, no study has been carried out that establishes, in a population endemic to malaria, such as Buenaventura, the prevalence of all previously mentioned resistance variants, the genotypic combinations that are found and in which proportion and which population distribution variables (such as sex, age and community) influence the genotypes of these resistance variants.

The main objective of this study was to establish, through molecular characterization, the prevalence of the variants of resistance to malaria (HbS, HbC, Duffy, β-thalassemias and G6PD) in the city of Buenaventura, an endemic region for malaria in Colombia, and to visualize the genotypic combinations of these variants, thus determining the degree of protection of individuals by age group and geographical area of the city (community).

Materials And Methods

Population and type of study

This was a cross-sectional study with stratified random sampling, where each “stratum” corresponded to each of the 12 communities of the city of Buenaventura, municipality of Valle del Cauca (Fig. 1).

The population sample consisted of 819 individuals of both sexes (214 minors and 605 adults) covering all age ranges. Foreigners, visitors or inhabitants who could not certify their origin and/or birth in the city were not included. Participants related to each other in the same generation were excluded to guarantee the nonrepetition of samples by kinship.

The health status of the participants was variable, and most of them were apparently healthy. Blood samples were stored at -4 °C in the LGMH of Universidad del Valle (Cali), where they were processed. This study was approved by the Human Ethics Review Committee (CIREH) of Universidad del Valle (Act No. 199-020).

Molecular diagnostics for study variants

DNA was extracted from blood samples using a "salting out" technique as previously described [34].

Duffy blood group

For the analysis of the Duffy gene, amplification of the promoter and coding regions (GATA and DARC, respectively) was performed as previously described by Ortega and collaborators [26]. For the GATA region, the 392-bp fragment obtained after amplification was treated with three units of the restriction enzyme *StyI*. The resulting fragments were 205, 110 and 65 bp, corresponding to the negative Duffy genotype (Fy-), and

205, 110, and 80 bp, corresponding to the positive Duffy genotype (Fy+). For the DARC region, the 159 bp fragment originating from the amplification was treated with one unit of the restriction enzyme *BanI*. The products of this digestion were 159 bp for Fy (a- b+), 86 and 73 bp for Fy (a+ b-) and 159, 86 and 73 bp for Fy (a+ b+). The fragments were visualized on 8% polyacrylamide gels.

Haemoglobin S (HbS)

The analysis of this variant was carried out as previously described by Ortega et al. [26]. The 571-bp amplified fragment was treated with 2 units of the restriction enzyme *DdeI.* The resulting DNA fragments were separated by polyacrylamide gel electrophoresis and examined by silver nitrate staining. For homozygous (SS) individuals, there was a single band of 308 bp; for heterozygous (AS) individuals, there were three bands of 308, 201 and 107 bp; and for homozygous (AA) individuals, there were 2 bands of 201 and 107 bp.

Haemoglobin C (HbC)

The analysis of this variant was carried out by allele-specific PCR using the primers proposed by Weatherall & Clegg [35]. A 206-bp fragment corresponding to the variant and an 860-bp fragment corresponding to the control were amplified. The absence of the variant was indicated by a 207-bp band. Fragments were visualized by 8% polyacrylamide gel electrophoresis.

β-Thalassemias

The -88 (CàT) and -29 (AàG) variants were detected using the allele-specific PCR technique proposed by Weatherall & Clegg [35]. For the -88 variant, a 684-bp fragment indicated the presence of the variant, and a 683-bp band indicated the absence of the variant. For the -29 variant, amplification of a 625-bp fragment indicated the presence of the variant, and a 624-bp fragment indicated the absence of the variant. Each reaction contained a set of control primers to amplify an 860-bp fragment. All fragments were visualized on 8% polyacrylamide gels.

Glucose 6-phosphate dehydrogenase deficiency (DG6PD)

Two regions corresponding to the A + (376A> G) and A- (202G> A) variants were identified by the allele-specific PCR technique following the method described by Liese et al. [36]. Primers 202G and 202A were used to identify variant A-, and primers 376A and 376G were used to identify variant A+. The amplified fragments were visualized on 8% polyacrylamide gels. The expected sizes of the fragments were 452 bp and 266 bp for the A + and A- variants, respectively. Each reaction contained a set of control primers to amplify a 333-bp fragment.

Data analysis

Descriptive statistics

The contingency tables for the different variables were constructed using the statistical program R version 3.4.1 [37]. The study variables were "Sex", "Age group", "Communities", " HbS/C variant", "Duffy variant", " β-

Tal-29 variant", " β-Tal-88 variant", "G6PD variant", and a variable derived from the latter called "Protection" (Table 1). “Protection” comprised categories corresponding to whether an individual had one, at least 2, all or none of the resistance genotypes of the variants studied. Age was categorized into three groups, i.e., 8 months to 12 years, 13 to 26 years and 27 to 93 years, following the life cycle guide proposed by the Ministry of Health [38].

Table 1. Variables of study.

Variable	Categories or values	Type of variable	Level of measurement	Measure method
Sex	1=female; 2= male	Qualitative	Nominal	Survey
Age group	1= 8 months to 12 years	Qualitative	Ordinal	Survey
	2=13 to 26 years			
	3=27 to 93 years			
Communities	1= Community 1	Qualitative	Nominal	Survery
	2= Community 2			
	3= Community 3			
	(…)			
	12= Community 12			
HbS y HbC (HbS/C) variant	1= AA; 2= CC*; 3= SS 4= AC*; 5= AS*; 6= SC*	Qualitative	Nominal	PCR-RFLP and Allele-specific PCR, followed by band counting
Duffy variant	1=FYB^{ES}/FYB^{ES}*; 2=FYA/FYA 3=FYB/FYB; 4=FYA/FYB 5=FYA/FYB^{ES}; 6=FYB/FYB^{ES}	Qualitative	Nominal	PCR-RFLP followed by band counting
βTal -29 variant	1= AA; 2= AG*; 3= GG	Qualitative	Nominal	Allele-specific PCR followed by band counting
βTal -88 variant	1= CC; 2= CT*; 3= TT	Qualitative	Nominal	Allele-specific PCR followed by band counting
G6PD variant	1= BB; 2= A+A+; 3= A-A-* 4= BA+; 5= BA-; 6= A+A-* 7=B; 8=A+; 9=A-*	Qualitative	Nominal	PCR-RFLP and Allele-specific PCR, followed by band counting
Protection	1=None (No protective genotypes) 2=One (having one protective genotype from a single variant) 3=At least 2 (having a combination of two or more protective genotypes from 2 or more variants)	Qualitative	Ordinal	Band counting for each individual considering the results obtained for all variants

*Genotypes of protection (resistance to malaria)

Calculation of allele and genotype frequencies and population differentiation

The allele and genotype frequencies were calculated for all study variants. The exact population differentiation test of Raymond and Rousset [39] was also carried out to evaluate subgroups of individuals by age and to compare the distribution of genotype frequencies with those reported for other populations in Colombia, which was complemented with the χ^2 homogeneity test. For the HbS and HbC variants, comparisons were made with HbS and HbC variant results reported in two studies conducted in Buenaventura (Valle) [23, 18], for two populations from San Andrés and Providencia [19, 31], for one population from Cali (Valle) [33], for two populations from Cartagena (Bolívar) [17, 24] and for one population from each of the following departments: Putumayo, Nariño, Guajira, Chocó and Valle [31]. For the Duffy variant, comparisons were made with Duffy variant results reported for the rural population of Buenaventura (Valle del Cauca) and for the populations of Tumaco (Nariño), Tierra Alta (Córdoba) [25] and Italy, located in the municipality of San José del Palmar (Chocó) [21]. For the G6PD variant, comparisons were made with G6PD variant results for the populations of Buenaventura (Valle), Quibdó (Chocó), Tierra Alta (Córdoba) and Tumaco (Nariño) [22]. For the -29 and -88 variants of β-thalassemia, it was not possible to make comparisons because no studies were found that reported relevant data for these variants in the Colombian population. All the tests referred to in this section were carried out with the Arlequin 3.5.2.2 statistical package [40] and R software version 3.4.1 [37]. For all the analyses, a value of 0.05 was used as the maximum accepted type I error.

Independence test, correspondence analysis and regression.

Chi-square tests and Fisher's exact tests [41] were used to examine the independence between the variables. The effect of the different independent variables (“sex”, “age”, “community”) on the response variables (“HbS/C variant”, “Duffy variant”, “β-Tal-29 variant"," β-Tal-variant”, “G6PD variant”, and “Protection”) was performed by means of a deviancy analysis associated with the multinomial or multiordinal logistic regression model. Once a significant effect of at least one of the aforementioned independent variables was evidenced, Fisher's multiple comparison test with Bonferroni correction was applied to better understand that significance. All the aforementioned tests were carried out with the statistical software R version 3.4.1 [37]

Results

Variant differentiation by age range and allele and genotype frequencies

The sample consisted of a total of 819 individuals belonging to the 12 communities of the urban area of Buenaventura; 214 were minors and 605 were adults (mean = 35 years, SD = 20 years). For some variants, detection was not possible in the 819 individuals; therefore, the total for some variants may vary slightly. Similarly, there were some records without information on age (called "SG"); the results for these individuals were taken into account when calculating the general frequencies for the population. When performing the population differentiation analysis for the genotypes of the variants in individuals with different ages (8 months to 12 years, 13 to 26 years and 27 to 93 years), there were significant differences among the HbS/C

variants (p = 0.0154), β-Tal-29 variant (p < 0.0001), and β-Tal-88 variant (p < 0.0001); there was no significant difference for the G6PD and Duffy variants (p = 0.1171 and p = 0.2295, respectively). These age groups were taken into account in the different analyses, especially in the calculation of fit, selection coefficients and average deviation.

Haemoglobin S and C variants

For these variants, 814 individuals were successfully diagnosed. The most frequent genotype was AA (wildtype homozygous), with 730 individuals (89.7%), represented mainly by age group 1 (8 months to 13 years), followed by the heterozygous carrier genotypes (or resistance genotypes) AS and AC, with frequencies of 5.8% and 4.2%, respectively, mostly evidenced in age groups 2 (13 to 26 years) and 3 (26 to 93 years). The least represented or unrepresented genotypes were the heterozygous genotypes of both variants, SC (0.2%), followed by homozygous SS (0.1%) and homozygous CC (0%), with the latter mostly present in age group 2. Finally, the following allele frequencies were found: 94.7% for allele A, 3.1% for allele S and 2.2% for allele C (Table 2).

Table 2. Allelic and genotypic frequencies for the HbS and C variants for the total individuals and each age group.

Age group	Genotype						Total
	AA	CC	SS	AC	AS	SC	
G1							152
	144*	0	0	3	5	0	
	94.7%**	0.0%	0.0%	2.0%	3.3%	0.0%	
G2							145
	123	0	1	4	16	1	
	84.8%	0.0%	0.7%	2.8%	11.0%	0.7%	
G3							494
	442	0	0	25	26	1	
	89.5%	0.0%	0.0%	5.1%	5.3%	0.2%	
SG							23
	21	0	0	2	0	0	
	91.3%	0.0%	0.0%	8.7%	0.0%	0.0%	
Total							814
	730	0	1	34	47	2	
	89.7%	0.0%	0.1%	4.2%	5.8%	0.2%	
	Allelic frequencies						
		A=	94.7%				
		C=	2.2%				
		S=	3.1%				

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records.

Duffy variant

For this gene, 819 individuals were diagnosed. Four hundred fifty (55%) had the homozygous genotype, i.e., the null or negative Duffy genotype (FYB^{ES}*FYB^{ES}), which was the most frequent genotype. This genotype was most represented in age group 2. That genotype was followed in frequency by the genotypes carrying the null allele, i.e., FYA*FYB^{ES} (18.6%) and FYB*FYB^{ES} (14.9%), with the highest proportions in age group 1. The FYA*FYB genotype was present in 6.2% of the individuals, with most in age group 3). The less frequent genotypes were FYA*FYA and FYB*FYB, present in 3% and 2.5% of the sample, respectively, mostly

evidenced in age group 3 (27 to 93 years). The FYB ^{ES} allele frequency was 72.2%, and the FYA and FYB allele frequencies were 15.2% and 12.6%, respectively (Table 3).

Table 3. Allelic and genotypic frequencies for the Duffy variant for the total individuals and each age group.

Age group	Genotype						Total
	FYB^{ES}*FYB^{ES}	FYA*FYA	FYB*FYB	FYA*FYB	FYA*FYB^{ES}	FYB*FYB^{ES}	
G1							154
	76*	3	3	7	35	30	
	49.4%**	1.9%	1.9%	4.5%	22.7%	19.5%	
G2							145
	85	6	2	4	28	20	
	58.6%	4.1%	1.4%	2.8%	19.3%	13.8%	
G3							497
	277	14	14	37	84	71	
	55.7%	2.8%	2.8%	7.4%	16.9%	14.3%	
SG							23
	12	1	1	3	5	1	
	52.2%	4.3%	4.3%	13.0%	21.7%	4.3%	
Total							
	450	24	20	51	152	122	819
	54.9%	2.9%	2.4%	6.2%	18.6%	14.9%	
	Allelic frequencies	FYA=	15.2%				
		FYB=	12.6%				
		FYB^{ES}=	72.2%				

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records.

β-thalassemia-29 and -88 variants

For these variants, 816 individuals were diagnosed. For the -29 variant, among the 816 individuals, 789 had the homozygous AA or normal haemoglobin genotypes (96.0%), 32 had the AG heterozygous genotype (3.9%) (or resistance genotype), and one had the homozygous GG or double variant homozygous genotype (0.1%). The allele frequencies were 97.9% for A and 2.1% for G. The AA genotype was most represented in age group 3, with 98.8%; the AG genotype was most represented in age group 2, with 9.7%; and the GG genotype was most represented in age group 1, with 0.7% (Table 4).

Table 4. Allelic and genotypic frequencies for the *β-thalassemia -29* variant for the total individuals and each age group

Age group	Genotype			Total
	AA	AG	GG	
G1				152
	139*	12	1	
	91.4%**	7.9%	0.7%	
G2				145
	131	14	0	
	90.3%	9.7%	0%	
G3				496
	490	6	0	
	98.8%	1.2%	0%	
SG				23
	23	0	0	
	100%	0%	0%	
Total		32	1	816
	783			
	95.9%	3.9%	0.1%	
	Allelic frequencies	A=	97.9%	
		G=	2.1%	

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records.

For variant -88, among the 816 individuals, 776 had the normal homozygous CC genotype (95.0%), 35 had the heterozygous TC genotype (4.29%), and 5 had the double variant or homozygous TT genotype (0.61%); the allele frequencies were 97.2% for C and 2.8% for T (Table 2). The CC genotype was most represented in age group 3, with 98.6%, and the heterozygous CT and homozygous TT genotypes were most represented in age group 2, with 11% and 2.1%, respectively (Table 5).

Table 5. Allelic and genotypic frequencies for the *β-thalassemia -88* variant for the total individuals and each age group

Age group	Genotype			
	CC	**CT**	**TT**	**Total**
G1				152
	139*	12	1	
	91.4%**	7.9%	0.7%	
G2				145
	126	16	3	
	86.9%	11.0%	2.1%	
G3				496
	489	6	1	
	98.6%	1.2%	0.2%	
SG				23
	22	1	0	
	95.7%	4.3%	0.0%	
Total				816
	776	35	5	
	95.1%	4.3%	0.6%	
	Allelic frequencies	**C=**	97.2%	
		T=	2.8%	

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records.

G6PD A+ and A- variants

The total sample was 817 individuals from the urban area of Buenaventura. The overall frequencies were 72.8% for the B allele, 16.2% for the African A+ allele, and 11.0% for the A- allele (between hemizygous men and heterozygous-homozygous women). Because G6PD is a sex-linked gene, the genotype frequencies obtained are reported by sex.

For the female population (n = 616), the following genotype frequencies were found: 54.5% for the BB genotype (wild genotype), most represented in age group 3, and 22.1%, 13.5% and 6% for the heterozygous BA+, BA-, and A+ and A- genotypes, respectively (the last two resistance genotypes), most represented in age groups 2 and 3. The homozygous genotypes AA+, with 2.6%, and AA-, with 1.3%, were most represented in age groups 3 and 1, respectively. Thus, allele frequencies of 72.3%, 16.6%, and 11.0% were observed for alleles B, A + and A-, respectively (Table 6).

Table 6. Allelic and genotypic frequencies for the G6PD variant in females for the total individuals and each age group

Age group	Female genotypes						Total
	BB	A+A+	A-A-	BA+	BA-	A+A-	
G1							85
	48*	0	2	20	12	3	
	56.5%**	0.0%	2.4%	23.5%	14.1%	3.5%	
G2							110
	52	2	2	34	12	8	
	47.3%	1.8%	1.8%	30.9%	10.9%	7.3%	
G3							404
	228	14	4	75	58	25	
	56.4%	3.5%	1.0%	18.6%	14.4%	6.2%	
SG							17
	8	0	0	7	1	1	
	47.1%	0.0%	0.0%	41.2%	5.9%	5.9%	
Total							616
	336	16	8	136	83	37	
	54.5%	2.6%	1.3%	22.1%	13.5%	6.0%	
	Allelic frequencies	B =	72.3%				
		A+=	16.6%				
		A- =	11.0%				

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records.

For the 201 men analysed, the following genotype frequencies (and therefore allele frequencies) were obtained: 74.1% for the B genotype (wild genotype), 14.9% for the A+ genotype and 10.9% for the A- genotype (resistant genotype). These alleles were most represented in age group 3 (YB genotype) and age group 2 (YA + and YA- genotypes) (Table 7).

Table 7. Allelic and genotypic frequencies for the G6PD variant in males for the total individuals and each age group

Age group	Male genotypes			Total
	B	A+	A-	
G1	51	9	9	69
	73.9%	13.0%	13.0%	
G2	23	7	5	35
	65.7%	20.0%	14.3%	
G3	71	13	7	91
	78.0%	14.3%	7.7%	
SG	4	1	1	6
	66.7%	16.7%	16.7%	
Total	149	30	22	201
	74.1%	14.9%	10.9%	

*Absolute frequency; **Relative frequency or prevalence. G1= 8 months to 12 years; G2 = 13 to 26 years; G3 = 27 to 93 years; SG = individuals without age records

Prevalence of variants by community

Table 8 shows in detail the genotype frequencies found for each variant in each community of the city of Buenaventura. The AA, AS and AC genotypes of the HbS/C variants were the most predominant in all the communities. In all the communities, the most prevalent genotype was wildtype AA, followed by heterozygous variants related to resistance to malaria: in some communities, heterozygous AS was more predominant (communities 1, 3, 7, 8, 9, 10, 11, 12), and in another, heterozygous AC was more predominant (communities 2, 4, 5, 6), with community 1 having the highest frequency of the AS genotype and community 5 having the highest frequency of the AC genotype. For the Duffy variant, the most prevalent genotypes were the FYB^{ES}*FYB^{ES} genotype (resistance genotypes) and the heterozygous FYA*FYB^{ES} FYB*FYB^{ES} genotypes. In all the communities, the genotype with the highest prevalence was the null homozygous FYB^{ES}*FYB^{ES}. For the β-thalassemia-29 and -88 variants, in all the communities, the most frequent genotypes were wild-type homozygous AA and CC, respectively, followed by heterozygous AG and CT (resistance genotypes), except for communities 7, 8, and 9, which did not have any individuals with the heterozygous CT genotype for -88, and communities 8, 9 and 12, which did not have any individuals with the heterozygous AG genotype for -29. Finally, for the G6PD variant in women, the most prevalent genotype in all communities was the wild genotype BB followed by the heterozygous BA+ and BA- genotypes (the latter

related to resistance). For men, the most frequent genotype was the wild BY genotype, followed by the A+ Y and AY genotypes, with the latter being related to resistance.

(Place table 8 here)

Independence test, multiple regression and prevalence

The results of Fisher's independence tests (p <0.005) were significant for "Community" and "Age group" (variables described in Table 1) with respect to all response variables. The exceptions were the "Duffy variant", which was not related to the variable "Age group" but with the variable "community" and the "G6PD variant", indicating that there was a relationship between "Duffy variant" and "Age group" but not between "Duffy variant" and "Community" (Table 9). Regarding the multinomial regression (and polynomial for the response variable "Protection"), the variable "Age group" influenced HbS/C, β-thalassemias -29 and -88 and G6PD (Table 9), and "community" influenced the Duffy variants, β-thalassemias -29 and -88, G6PD and "Protection". For both tests, no significance was found for the variables with respect to "Sex", except for "G6PD variant" in Fisher's test of independence (an expected result because G6PD is sex-linked).

Table 9. *Independence test and multinomial and ordinal regression for the study variables.*

	Variables						
		Duffy variant	HbS/C variant	βTal. -29 variant	βTal. -88 variant	G6PD variant	Protección
Variables (Independence test)	Sex	0,8734	0,9725	0,2629	0,0455	0,0005*	0,5508
	Age group	0,2584	0,0165*	0,0005*	0,0005*	0,0005*	0,0111*
	Community	0,0005*	0,0001*	0,0005*	0,0005*	0,1999	0,0005*
Factores (Regression)	Sex	0,8843	0,8353	0,3884	0,1617	0,4693	0,2758
	+ Age group	0,1957	0,0297*	<0,0000*	<0,0001*	0,0112*	0,0029*
	+ Community	<0,0001*	0,1224	0,0001*	<0,0001*	0,0002*	<0,0001*
*Significan p-values (p<0.05)							

Considering the results of Fisher's independence tests and multiple regression analyses, the multiple comparison test was conducted for “categorized age” and “community” with respect to “protection”. For age, there were significant differences between the age group of 13 to 26 years and the age group of 8 months to 13 years and between the age group of 13 to 26 years and the age group of 27 to 93 years. On the other hand, communities 1, 3, 4 and 5 were significantly different from communities 6, 7, 8, 9, 10, 11 and 12 ($p \leq 0.0361$)

To delve deeper into the results of these tests, the prevalences of “categorized age” and “community” were established in relation to the variable “protection” (Fig. 2 and 3). For the age groups, the age group from 13 to 26 years had a higher percentage of individuals in the category “at least two” (33.1%) than did the other two age groups. This highest level of protection indicates that these individuals present combinations of two or more protective genotypes of 2 or more of the resistance variants investigated in this study. The frequency of “at least two” was second highest in the age group from 27 to 93 years, followed by the age group from 8 months to 12 years (Fig. 2). The “one” category (that is, presenting only one resistance genotype of only one variant) was the most prevalent in all age groups, being higher in the age group from 27 to 93 years and the age group from 8 months to 12 years. Finally, the category “none” (not having any resistance genotype) was also most represented in the age groups of 8 months to 12 years and 27 to 93 years (Fig. 2).

Regarding the communities (Fig. 3), 1, 3, 4, and 5 had higher prevalences of individuals with “at least 2” (from 25.6% to 38.3%), with less representation of “at least 2” in communities 2, 6, 7, 8, 9 and 10. The category “one” was also the most prevalent in all communities, with quite similar percentages among them, with 2, 4 and 9 having the highest percentages. Finally, individuals with no protective variant were most represented in communities 6 to 12 (30.6% to 44%), with very little representation in communities 1 to 5 (Fig. 3).

Regarding the “protection” variable, 276 individuals were included in the “none” category; that is, for any variable, they did not have a resistance genotype. The category “one” included 360 individuals, with the Duffy variant (genotype $FYB^{ES*}FYB^{ES}$) being the most represented, with 78.3%, followed by G6PD, with 10.8%, with the BA- and A+A- genotypes being more frequent, and HbS/C, with 6.4%, with the AS and AC genotypes being more frequent (Fig. 4). For "at least 2" for the “protection” variable, among the 183 individuals who were included in that category, 160 (87.4%) presented a combination of only two protection genotypes for two resistance variants, 21 (11.9%) presented a combination of three protection genotypes for three resistance variants, and only 1 (0.0%) presented a combination of four protection genotypes. No individual presented a combination of 5 protection genotypes from the 5 variants investigated (Table 1).

For "at least two" for the protection variable, the most prevalent double combination included the resistance genotypes of the Duffy + G6PD variants (53.8%), with the genotype combination $FYB^{ES*}FYB^{ES}$/BA- being the most frequent, followed by $FYB^{ES*}FYB^{ES}$/A+A-. The HbS/C + Duffy pair (25.6%) was the second most frequent, represented mostly by the AS/$FYB^{ES*}FYB^{ES}$ genotype, followed by the AC/$FYB^{ES*}FYB^{ES}$ combination. No individual had a combination of resistance genotypes of the HbS/C variants and β-Tal-29 (Fig. 5).

Regarding triple combinations (three resistance genotypes of three different variants), the HbS/C + Duffy + G6PD combination was the most prevalent (36.4%), with the genotypic combination AS/FYB^{ES}*FYB^{ES}/BA- being more frequent (Fig. 6). That combination was followed in frequency by the triple combination Duffy + β-Tal-29 + β-Tal-88 with 18.2% (FYB^{ES}*FYB^{ES}/AG/CT genotypes). There were no individuals with the HbS/C + β-Tal-29 + β-Tal-88, HbS/C + β-Tal-29 + G6PD or HbS/C + β-Tal-88 + G6PD combinations. Finally, only one individual had a combination of 4 resistance genotypes: HbS/C + Duffy + β-Tal-88 + G6PD (AS/FYB^{ES}*FYB^{ES}/CT/BA- genotypes).

Comparison with other populations of Colombia

The analysis of population differences and homogeneity in the Buenaventura population, with respect to the results reported for the variants in populations in Colombia, yielded the following results. Regarding the results for the three alleles of HbS and HbC obtained herein, there were significant differences ($p \leq 0.0064$) in the distribution of genotype frequencies in the populations of Cali [33], Cartagena [17], Putumayo, Nariño, Guajira, San Andrés, Chocó and Valle [31]; however, there were no significant differences ($p \geq 0.0935$) in the distribution of genotype frequencies in the populations of San Andrés [19], Cartagena [24] and Buenaventura [23]. Regarding the results for the Duffy variant, there were significant differences ($p < 0.0001$) in the distribution of genotype frequencies between the population of Buenaventura in this study and the populations of Italy, Chocó [21], Tierra Alta, and the indigenous population of Buenaventura [25], with no significant differences in the frequency distribution between the population in this study and the population of Tumaco [25]. Finally, for the G6PD variant, there were significant differences ($p \leq 0.0465$) in the distribution of genotype frequencies between the population in this study and the populations of Quibdó, Tierra Alta and Tumaco, with no differences with respect to the population of Buenaventura [22].

Discussion

The allele and genotype frequencies found for the variants investigated in this study coincide with those reported for the black population in the world, especially for some populations from Africa and the Middle East [42]. In this study, the allele frequencies were 3.1% for HbS, similar to that reported for the populations of Kenya, Niger, Ghana, Mozambique, Saudi Arabia, Iran and India [42, 43, 44, 45]; 2.2% for HbC, close to that for populations in Mauritania, Niger, Guinea and Ghana [42, 46]; 72.2% for the FYB^{ES} allele, similar to that for populations from Namibia, Niger, Algeria, Mauritania and Sudan [42, 47, 48]; 11% for the A- allele of the G6PD variant, as in the populations of Senegal, Angola, Tanzania, Nigeria, Saudi Arabia, Iran, India and Pakistan [42, 49, 50]; 2.1% for the G allele of the β-Tal-29 variant, similar to those found in populations from Tunisia, Algeria and Morocco [35, 51]; and 2.8% for the β-Tal-88 T allele, close to that reported in Middle Eastern countries such as Iran, Pakistan and Lebanon [52, 53], and although this mutation is of African origin and it is estimated that the maximum frequency in these countries is approximately 21% [35, 52], allele frequencies are not reported for specific African populations but rather in general in databases or for Afro-descendant individuals living in other geographical areas [35, 52, 54]. Regarding studies carried out in Colombian populations, for the three alleles of the HbS and HbC variants, frequencies ranged from 0.1 to 4.6% for the S allele and from 0.2 to 3.0% for the C allele [23, 18, 19, 31, 33, 17, 24, 31]. For those studies, the Buenaventura population was significantly different from the populations of Cali, Putumayo, Nariño,

Guajira, Chocó, Valle, one of the two populations from Cartagena and one of the two populations from San Andrés. Analysing each allele separately, the C allele is different between the population of Buenaventura and the other populations, in which it was found in lower frequencies, with the exception of the population of Cali, where the frequency of both the S and C alleles is low and both are significantly different. For the previously mentioned populations, the differences found could be explained by the Afro-descendant component of the sample because although these populations have a significant percentage of Afro descendants, the percentage does not exceed that reported for the Buenaventura population (~ 90%) [55]. In addition, for the population in this study, all participants identified as Afro, and in the samples of the other studies, there was a mixture of ethnicities. Currently, this is the only study in which analyses were carried out directly by PCR using DNA for these variants; in other studies, capillary electrophoresis, isoelectric focusing and HPLC techniques were used, methods that can generate discrepancies with respect to molecular diagnoses using DNA, which is usually a definitive diagnostic technique for confirming gene variants [56]. For the Duffy gene, there were discrepancies between the population in this study and the populations of Italy in Chocó, Tierra Alta and the indigenous population of Buenaventura. In those populations, there was a lower frequency of the null Duffy allele (FYB^{ES}) and a higher frequency of the FYA and FYB alleles with respect to those estimated for the Buenaventura population in our study; this difference is attributed to the high Amerindian component of the samples in the other studies [25]. This is also an indication that the genetic component of the rural Buenaventura population cannot be equated with that of the urban population, and therefore, the most appropriate measures to consider when conducting diagnostic studies is to take into account the population structure. For the G6PD variant, it was not possible to compare allele and genotype frequencies in this study population with those in other populations because this is the first study in Colombia that makes a distinction between the three alleles of this variant (B, A+ and A-) and the 6 possible genotypes. In the other studies, although some include molecular diagnoses, they do not report allele and genotype frequencies but rather measure G6PD enzyme activity. For example, Herrera-Valencia and collaborators [57] mixed all the carrier genotypes of each allele by group; by having the same grouping in our study, it was possible to compare our population with their population, finding significant differences between the population of Buenaventura in this study and the populations of Quibdó, Tierra Alta and Tumaco but not between the population of Buenaventura in this study and the population of Buenaventura in their report.

As observed in the results of the independence and regression analyses, the genotypic patterns of the resistance variants, as well as the “protection” variable, were influenced in some way by the age of the individuals as well as by their geographical location within the city of Buenaventura ("communities"). This finding is similar to that reported for the same population in previous studies, only taking into account the resistance variants of HbS and Duffy [26, 28]; it was surprising that the phenomenon continued to occur after taking into account more variants of resistance to malaria. Regarding age, the 13- to 26-year-old group had clear differentiation in genotypic patterns with respect to the other two age groups (8 months to 12 years and 27 to 93 years); this group had a higher proportion of resistance genotypes of the different variants investigated, both individually (Tables 2-7) and as a whole (Fig. 3), as well as a lower proportion of genotypes that do not confer resistance. These results indicate that compared to other age groups, adolescents and young adults from Buenaventura have a genotypic composition that could confer greater

resistance to or protection from malaria infection; importantly, however, children under 12 years of age and older adults may be more susceptible to malaria infection. These data agree with results reported in the literature, i.e., individuals under 5 years of age (or in some, under 14 years of age), pregnant women and older adults are more susceptible to malaria infection [58, 59, 60]. In Buenaventura, specifically, the highest numbers of reported weekly and yearly cases of malaria infection occur in individuals between the ages of 10 and 29 years, decreases considerably with aging [61]. This observed pattern has been related to the work activities of these groups because young people and adults are more exposed to infection by the parasite due to their occupations, which often involve travel to rural areas and, therefore, increased exposure to mosquito bites. In contrast, young children and older adults, due to their daily activities, tend to be less exposed to mosquito bites and have lower infection rates [62, 63]. Currently, the highest mortality rates from malaria in Valle del Cauca have been reported in people over 60 years of age and under 5 years of age [64], as has been reported in various regions of the world. Therefore, it is worth noting once again that although adolescents and young adults seem to have the highest rates of malaria infection, they have the lowest mortality rates from malaria, which could be due in large part to the fact that they have the highest frequencies of the genotypes of variants that generate resistance to or protection from malaria [26]; in contrast, those younger than 5 years and older than 60 years could be more affected by having lower frequencies of these resistance genotypes. Although as an individual ages he or she acquires immunity due to constant infections [63], as ageing continues, the proportion of naïve and memory T cells in the immune system tends to gradually decrease due to the accumulation of exposure to various antigens, in addition to a reduction in thymus activity, which can limit the production of such cells [65].

These findings once again support the fact that genotypes that confer malaria resistance do not protect an individual from developing disease but rather decrease the parasitic loads of *Plasmodium* and decrease in the clinical severity of the disease (such as cerebral malaria and malaria with severe anaemia), which generally lead to death; additionally, in some cases, a protective effect has been observed in uncomplicated malaria [5, 6]. For the variants investigated in this study, parasite loads were lower in individuals with resistance genotypes than in individuals with normal genotypes, mainly for *Plasmodium falciparum* (for the 4 variants mentioned) and *P. vivax* (for Duffy, G6PD and β thalassemias) [7, 8, 9, 10, 11, 12, 13]. Regarding the reduction in cerebral malaria and severe anaemia, a greater protective effect against these clinical manifestations has been found for the HbS, HbC and G6PD variants (except for severe anaemia, for which it has been observed that the G6PD variants worsen symptoms) [5, 66].

Regarding the geographic zones of the urban area of Buenaventura, called “communities”, it was found that factors such as age have an effect on the genotypic distribution of the resistance variants and of the “protection” variable: communities 1, 3, 4, and 5 had higher frequencies of the protective genotypes of the variants as well as higher frequencies of “at least two” protective variants; communities 6 to 12 had lower proportions of “at least two” and higher frequencies of “none” (Fig. 3). This clear differentiation is supported by the analysis of multiple comparisons, showing significant differences between the two groups of communities. The previous finding agrees, strikingly, with the geographical distribution of the communities in the urban area of Buenaventura because communities 1, 2, 3, and 4 are in the insular area of the city, communities 6 to 12 are in the continental zone, and community 5 serves as the bridge between the communities of the insular zone and the continental zone. For many years, the communities of the insular

zone, the oldest in the city, have had the lowest rates of malaria infection. In contrast, those in the continental zone have had the highest number of infections (especially communities 10 and 12) [60, 66]. The latter has been attributed to the greater population density of the vector in these communities, the higher proportion of forest area and the substantial number of individuals with unsatisfied basic needs [67, 68]. To the above, we can add what was found in this study: the communities of the continental zone were those with the lowest frequencies of resistance genotypes; therefore, their inhabitants are the most susceptible to developing severe forms of malaria. On the other hand, the populations in the communities of the insular zone, because they are older, may have had more time to adapt to malaria through their "inventory" of resistance genotypes. However, although having resistance genotypes and higher allelic frequencies of these variants than those for other non-Afro-descendant populations of Colombia, the population of Buenaventura is more robust when facing malaria, making these individuals more susceptible to developing haemoglobinopathies because these alleles are carried by heterozygous individuals who do not develop severe forms of such diseases (such as sickle cell anaemia or thalassemia major) and therefore will be more likely to pass these genes to their offspring and thus perpetuate cases of haemoglobinopathies in the population in the future.

Finally, having found a greater representation of individuals with the Duffy resistance variant both in the total number of individuals (Table 3) and in the single variant and double and triple combinations (Figs. 4-6) indicates that the FYB^{ES} allele or Duffy null allele (which provides resistance), unlike the other variants in this study, is the only variant that is not associated with any haemoglobinopathy or erythrocyte defect; in fact, it is the only resistance variant with a frequency of 100% in some African populations [42]. Similarly, the A- variant of the G6PD gene, which is next in frequency (Table 6, Fig. 4) and was also the most prevalent in the double and triple combinations together with Duffy and HbS/C (Fig. 5 and 6), is a variant that, although it reduces enzyme activity, does not reduce activity by 100%: hemizygous men and homozygous women have approximately 12 to 20% enzymatic activity [11, 69], and this allele can be found at maximum frequencies of 32.5 to 35% in some African and Middle Eastern populations [42]. The third in frequency for the total population and the second unique and combined variants were HbS and HbC (Table 2, Fig. 4), which have been recorded at maximum frequencies of 18.4% in some African populations [42]. These frequencies are lower than those of the aforementioned variants and are expected for haemoglobin S due to its lethal or almost lethal character but not expected for HbC because, on the contrary, homozygous CC presents a higher protective effect than does heterozygous AS (approximately 90% for CC and 70% - 85% for AS) [12], without a lethal effect. However, due to the allele frequency found in this study, CC homozygotes in the population are unlikely, with a higher likelihood of AC heterozygotes, a genotype that has a lower protective effect than AS (47%) [13]. Finally, for β-thalassemias in this study, -29 and -88, their frequencies were the lowest and similar to those for haemoglobin C (Tables 4 and 5 and Fig. 4) but much lower than those reported for Africa (60% and 21% for -29 and -88, respectively) [35, 54] and were the least frequently found in double and triple combinations. This is perhaps because their resistance genotypes coincided (which is only one, the heterozygote), compared to the other variants that have more than one resistance genotype, and their frequencies were higher in this population. In addition, although in the African population these are variants that generate thalassemia minor symptoms, in other populations it has been found that these variants can generate more severe symptoms of the disease (thalassemia major) [70]. Therefore, it is not

truly known if these variants could be causing deleterious effects in the population of Buenaventura and if this could be the reason for finding a decreased frequency of these variants, not to mention the ethnic mixture, which may be influencing the diversity of allelic frequencies for this and all variants.

Conclusions

In the urban area of Buenaventura, the group of adolescents and young adults (13 to 26 years) has higher frequencies of resistance genotypes that could confer greater resistance to malaria, as do individuals from the communities of the insular area (1, 3, 4 and 5). Likewise, individuals younger than 12 years and older than 26 years could have a greater susceptibility to malaria infection due to the low presence of resistance genotypes, as could individuals residing in communities 6, 7, 10 and 12 of the continental zone of the city. Finally, the distribution of allele and genotype frequencies for the resistance variants investigated in this study in the urban area of Buenaventura cannot be equated with those of other areas of Colombia, not even in the rural area of Buenaventura (this for the Duffy gene). This is possibly due to the different proportions of individuals of Afro descent that make up the regions of the country.

Declarations

Acknowledgements:

The authors express their gratitude to the community leaders of the city of Buenaventura and all study participants. Thanks to the Universidad del Valle for funding the project and to the Human Molecular Genetics Laboratory for providing space and equipment. Special appreciation to Wilmar Torres for his statistical guidance.

Ethics Approval and Consent to Participate:

This study received approval from the Human Ethics Review Committee (CIREH) of Universidad del Valle (Act No. 199-020). All participants read and signed the informed consent.

Author Contributions:

D.C.O. conceived the study. D.C.O., G.B., H.C., and R.G. designed the research protocol and participated in resource acquisition. G.B. led the project. D.C.O., M.P.A., and S.C. were involved in sample collection, laboratory work, and data analysis. D.C.O. drafted the manuscript, and G.B., H.C., and R.G. reviewed, corrected, and contributed to its intellectual content. All authors read and approved the final manuscript.

Funding:

This work was funded by the internal call 141-2022 (CI 71340) and the internal call for support to Ph.D. students, 2021 (CI 71311), from Universidad del Valle.

Consent for Publication:

Not applicable.

Competing Interests:

The authors declare no conflicts of interest.

References

1. Antinori S, Galimberti L, Milazzo L, Corbellino M. Biology of human malaria plasmodia including *Plasmodium knowlesi*. Mediterr J Hematol and Infect Dis. 2012; 4(1): e2012013.
2. CDC-Centers for Disease Control, Prevention. CDC - Malaria - Malaria Worldwide - Impact of Malaria. 2009. https://www.cdc.gov/malaria/malaria_worldwide/ impact.html. Accessed 13 Oct 2023.
3. OPS.org. Aumentan los casos de malaria en las Américas. 2018. https://www.paho.org/col/index.php?option=com_content&view=article&id=2892:aumentan-los-casos-de-malaria-en-las-americas&Itemid=487. Accessed 15 Oct 2023.
4. INS. Malaria Colombia 2020. Inst Nac Salud. 2020. https://www.ins.gov.co/buscador_eventos/informesdeevento/MALARIA_2020.pdf.
5. Taylor SM, Parobek CM, Fairhurst RM. Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. Lancet. Infect Dis. 2012; 12(6): 457–68. https://doi.org/10.1016/S1473-3099(12)70055-5.
6. Roberts DJ, Williams TN. Haemoglobinopathies and resistance to malaria. Redox Rep. 2003; 8(5):304–310.
7. Amoah LE, Asare KK, Dickson D, Abankwa J, Busayo A, Bredu D, et al. Genotypic glucose-6-phosphate dehydrogenase (G6PD) deficiency protects against *Plasmodium falciparum* infection in individuals living in Ghana. PloS one. 2021; 16(9): e0257562.
8. Brown CA, Pappoe-Ashong PJ, Duah N, Ghansah A, Asmah H, Afari E, et al. High frequency of the Duffy-negative genotype and absence of *Plasmodium vivax* infections in Ghana. Malar J. 2021; 20(1), 99.
9. Abou-Ali RK, Dhyani A, Terço AL, Toro DM, Gomes KS, Tezza LC, et al. Impact of Duffy polymorphisms on parasite density in Brazilian Amazonian patients infected by *Plasmodium vivax*. Malar J. 2019; 18(1):289.
10. Kuesap J, Chaijaroenkul W, Rungsihirunrat K, Pongjantharasatien K, Na-Bangchang K. Coexistence of Malaria and Thalassemia in Malaria Endemic Areas of Thailand. The Korean J Parasitol. 2015; 53(3); 265–270.
11. Uyoga S, Ndila CM, Macharia AW, Nyutu G, Shah S, Peshu N, et al. Glucose-6-phosphate dehydrogenase deficiency and the risk of malaria and other diseases in children in Kenya: a case-control and a cohort study. Lancet. Haematol. 2015; 2(10): e437–44.
12. Mangano VD, Kabore Y, Bougouma EC, Verra F, Sepulveda N, Bisseye C, et al. Novel Insights Into the Protective Role of Hemoglobin S and C Against *Plasmodium falciparum* Parasitemia. J Infect Dis. 2015; 212(4): 626–634.
13. Mockenhaupt FP, Ehrhardt S, Cramer JP, Otchwemah RN, Anemana SD, Goltz K, et al. Hemoglobin C and resistance to severe malaria in Ghanaian children. J Infect Dis. 2004; 190(5): 1006–1009.

14. Kariuki SN, Williams TN. Human genetics and malaria resistance. Hum Genet. 2020; 139(6-7), 801–811. doi.org/10.1007/s00439-020-02142-6.
15. Travassos MA, Coulibaly D, Laurens MB, Dembélé A, Tolo Y, Koné AK, et al. Hemoglobin C Trait Provides Protection From Clinical Falciparum Malaria in Malian Children. J Infec Dis. 2015; 212(11): 1778–86.
16. Aguirre M, Medina D, Araujo MV, Campo MA, Castro A, Fernández-Trujillo L, et al. Importancia de la detección temprana de hemoglobinopatias en la población pediátrica en países en desarrollo. Revista chilena de pediatría. 2020; 91(4):568–572.
17. Alvear CC, Barboza M, Viola M, Moneriz C, Araque LM. Pilot study of hemoglobinopathies in newborns of the Rafael Calvo maternity clinic of Cartagena, Colombia. Colombia médica. 2012; 43(3):196–199.
18. Bernal M, Collazos A, Bonilla RD, Tascón EP. Determination of the prevalence of hemoglobin S, C, D, and G in neonates from Buenaventura, Colombia. Colombia Médica. 2010; 41(2):141–147.
19. Bernal MD, Giraldo A, Bermúdez AJ, Moreno E. Estudio de la frecuencia de hemoglobinopatías en las islas de San Andrés y Providencia, Colombia. Biomédica, 1995; 15(1):5–9.
20. Carmona-Fonseca J, Álvarez G, Ríos A, Vásquez MF. Deficiencia de glucosa 6-fostato deshidrogenasa en hombres sanos y en pacientes maláricos; Turbo (Antioquia, Colombia). Rev Bras Epidemiol. 2008; 11(2), 252–265.
21. González L, Vega J, Ramirez JL, Bedoya G, Carmona-Fonseca J, Maestre A. Relationship between genotypes of the Duffy blood groups and malarial infection in different ethnic groups of Choco, Colombia. Colomb Med. 2012; 43(3), 189–195.
22. Valencia SH, Ocampo ID, Arce-Plata MI, Recht J, Arévalo-Herrera M. Glucose-6-phosphate dehydrogenase deficiency prevalence and genetic variants in malaria endemic areas of Colombia. Malar J. 2016; 15(1), 291.
23. Moyano M, Méndez F. Defectos eritrocíticos y densidad de la parasitemia en pacientes con malaria por *Plasmodium falciparum* en Buenaventura, Colombia. Rev Panam Salud Publica. 2005; 18(1), 25–32.
24. Silva JR, Malambo D, Silva DF, Fals Borda E, Fals O, Rey, J. Tamizaje de hemoglobinopatías en una muestra de la población infantil de Cartagena. Pediatría (Santiago).1998; 33: 86–9.
25. Vallejo AF, Chaparro PE, Benavides Y, Álvarez Á, Quintero JP, Padilla J, et al. High prevalence of sub-microscopic infections in Colombia. Malar J. 2015; 14(1): 201.
26. Ortega DC, Cardenas H, Barreto G. Genetic variants of Duffy and hemoglobin S genes in an Afrodescendant population from Colombia. Hum Biol. 2018; 90(4): 271–280.
27. Ortega DC, Cardenas H, Barreto G. Joint selection for two malaria resistance mutations in a south-west Colombian population. Infec Genet Evol. 2020; 80(104188): 104188. doi.org/10.1016/j.meegid.2020.104188.
28. Ortega DC, Fong C, Cardenas H, Barreto G. Evidence of over-dominance for sickle cell trait in a population sample from Buenaventura, Colombia. Int J Genet Mol Biol. 2015; 7(1): 1–7.
29. Quintero-Santacruz M, Flórez Elvira L, Mejía Hurtado AF, Macia Mejía C. Estimated prevalence of the Duffy null phenotype Fy (a-b-) among black blood-donors in Southwestern Colombia. Transfus Apher Sci. 2020; 59(6): 102884.

30. Restrepo AM, Gutierrez E. The frequency of glucose-6-phosphate dehydrogenase deficiency in Colombia. Am J Hum Genet. 1968; 20(1): 82–5.
31. Rosero MJ, Bermúdez AJ. Análisis de hemoglobinopatías en regiones afrocolombianas usando muestras de sangre seca de cordón umbilical. Acta Med Colomb. 2012; 37(3): 117-26.
32. Romero-Sánchez C, Gómez Gutiérrez A, Duarte Y, Amazo C, Manosalva C, Chila M L, et al. Variantes de hemoglobina en una población con impresión diagnóstica positiva para hemoglobinopatías en Colombia. Rev Med Chi. 2015; 143(10): 1260–8.
33. Satizabal JM, Neuta P, Torres J, Somoyar P. Tamizaje de hemoglobinopatías en neonatos de Cali, Colombia. Rev Gastro. 2013; 15*(2)*: 4-7.
34. Miller SA,Dykes DD, Polesky HF. Asimple salting out procedure for extracting DNA from human nucleated cells. Nucle Acid Res. 1988; 16(3):1215.
35. Weatherall DJ, Clegg JB. The Thalassaemia Syndromes. Wiley. 2001; doi: 10.1002/9780470696705.
36. Liese AM, Siddiqi MQ, Spolacris Z. Rapid detection of glucose-6-phosphate dehydrogenase type A-(202A/376) deficiency by allele-specific polymerase chain reaction. Am J Hematol. 2000; 63(3):159-60.
37. R Core Team. R: A Language and Environment for Statistical Computing. 2017. https://www.R-project.org/. Accessed 13 Oct 2023.
38. Comisión de la Verdad: Los pueblos indígenas de Buenaventura y su aporte a la verdad. https://web.comisiondelaverdad.co/actualidad/noticias/los-pueblos-indigenas-de-buenaventura-y-su-aporte-a-la-verdad (2019). Accessed 19 Nov 2023
39. Raymond M, Rousset F. An exact test for population differentiation. Evolution; International Journal of Organic Evolution. 1995; 49(6): 1280–1283.
40. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010; 10(3), 564–567. doi.org/10.1111/j.1755-0998.2010.02847.x.
41. Fisher RA. The Logic of Inductive Inference. J R Stat Soc. 1935; 98(1):39–82.
42. Malaria Atlas Project. https://data.malariaatlas.org/maps. Accessed 15 Oct 2023.
43. Pule GD, Chimusa ER, Mnika K, Mhandire K, Kampira E, Dandara C, et al. Beta-globin gene haplotypes and selected Malaria-associated variants among black Southern African populations. Glob Health Epidemiol Genom. 2017; 2(e17).
44. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Williams TN, et al. Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. Nat Commun. 2010; 1(1): 104. doi.org/10.1038/ncomms1104.
45. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Dewi M, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet; 2013; 381(9861): 142–151.
46. Piel FB, Howes RE, Patil AP, Nyangiri OA, Gething PW, Bhatt S, et al. The distribution of haemoglobin C and its prevalence in newborns in Africa. Sci Rep. 2013; 3(1). doi.org/10.1038/srep01671.

47. Albsheer MM, Pestana K, Ahmed S, Elfaki M, Gamil E, Ahmed SM, et al. Distribution of Duffy Phenotypes among *Plasmodium vivax* Infections in Sudan. Genes. 2019; 10(6):437.
48. Howes RE, Patil AP, Piel FB, Nyangiri OA, Kabaria CW, Gething PW, et al. (2011). The global distribution of the Duffy blood group. Nat Commun. 2011; 2(1): 266.
49. Howes RE, Piel FB, Patil AP, Nyangiri OA, Gething PW, Dewi M. G6PD deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based map. PLoS Med. 2012; 9(11): e1001339.
50. Howes RE, Battle KE, Satyagraha AW, Baird JK, Hay SI. G6PD deficiency: global distribution, genetic variants and primaquine therapy. Advances in Parasitology. 2013; 81: 133–201.
51. Haj Khelil A, Denden S, Leban N, Daimi H, Lakhdhar R, Lefranc G, et al. Hemoglobinopathies in North Africa: a review. Hemoglobin. 2010; 34(1): 1–23.
52. Giardine BJ, Viennas E, Pavlidis C, Moradkhani K, Joly P, Bartsakoulia M, et al. Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. Nucleic acids research. 2014. 42(DI): D1063–D1069.
53. Miri-Moghaddam E, Bahrami S, Naderi M, Bazi A, Karimipoor M. Molecular Characterization of β-Thalassemia Intermedia in Southeast Iran. Hemoglobin. 2016; 40(3): 173–178.
54. Old JM. Screening and genetic diagnosis of haemoglobinopathies. Scand J Clin Lab Invest. 2007; 67(1): 71–86.
55. Ministerio de Salud y Protección Social Colombia. https://www.minsalud.gov.co/proteccionsocial/Paginas/cicloVida.aspx. Accessed 19 Nov 2023.
56. Sorroche P, Sáez, MS. La electroforesis capilar en el estudio de las Hemoglobinopatías y Talasemias. Hematología. 2014; 18(*3)*: 272–6.
57. Herrera-Valencia S, Ocampo ID, Arce-Plata MI, Recht J, Arévalo-Herrera M. Glucose-6-phosphate dehydrogenase deficiency prevalence and genetic variants in malaria endemic areas of Colombia. Malar J. 2016; 15(1), 291.
58. Jeevatharan H, Wickremasinghe R. Susceptibility to malaria during the prevention of re-establishment phase in Sri Lanka. Malar J. 2022. doi.org/10.1186/s12936-022-04127-4.
59. World Health Organization. Malaria. 2023. https://www.who.int/news-room/fact-sheets/detail/malaria. Accessed 15 Oct 2023.
60. Brooker SJ, Clarke S, Fernando D, Gitonga CW, Nankabirwa J, Schellenberg D, et al. Malaria in Middle Childhood and Adolescence. In Disease Control Priorities, Third Edition (Volume 8): Child and Adolescent Health and Development. The World Bank; 2017. p. 183–198.
61. SIGIVILA. Informe equipo de vigilancia en salud pública. Secretaria de salud de Buenaventura. https://www.buenaventura.gov.co/articulos/informe-equipo-de-vigilancia-en-salud-publica. Accesessed 15 Oct 2023.
62. Galindo-Buitrago JI, Hernández Rodríguez RA, Jiménez Barbosa WG. Comportamiento epidemiológico de la malaria en la costa pacífica del departamento de Nariño, Colombia, 2003-2017. Rev Cienc Salud. 2020; 18(3).

63. Mendez F, Carrasquilla G, Muñoz A. Risk factors associated with malaria infection in an urban setting. Trans R Soc Trop Med Hyg. 2000; 94(4);367-71.
64. Osorio L, Fernández Ja, Murillo O, Escobar H, Bustamante P, Agudelo H, et al. Caracterización de la mortalidad por malaria en el Valle del Cauca, 2005-2006. Biomedica. 2010; 29(4), 582-590.
65. Baird JK. Age dependent characteristics of protection v. susceptibility to *Plasmodium falciparum*. Ann Trop Med Parasitol. 1988; *92*(4):367–390.
66. Clarke GM, Rockett K, Kivinen K, Hubbart C, Jeffreys AE, Rowlands K, et al. Characterisation of the opposing effects of G6PD deficiency on cerebral malaria and severe malarial anaemia. Elife. 2017; doi.org/10.7554/eLife.15085.
67. Méndez F, Carrasquilla G. Epidemiología de la malaria en el área urbana de Buenaventura: análisis de la ocurrencia en el período 1987-1993. Colombia Médica. 1995; 26.77–85.
68. DANE. La información del DANE en la toma de decisiones de los municipios del país. 2020. https://www.dane.gov.co/files/investigaciones/planes-desarrollo-territorial/100320-Info-Alcaldia-Buenaventura.pdf. Accessed 15 Oct 2023.
69. Shah, SS, Macharia A, Makale J, Uyoga S, Kivinen K, Craik R., et al. Genetic determinants of glucose-6-phosphate dehydrogenase activity in Kenya. BMC Med Genet. 2014; 15(1).
70. Huang SZ, Wong C, Antonarakis SE, Ro-Lein T, Lo WHY, Kazazian HHJ. The same TATA box β-thalassemia mutation in Chinese and US blacks: another example of independent origins of mutation. Hum Genet. 1986; 74(2): 162-4.

Tables

Table 8 is available in the Supplementary Files section.

Figures

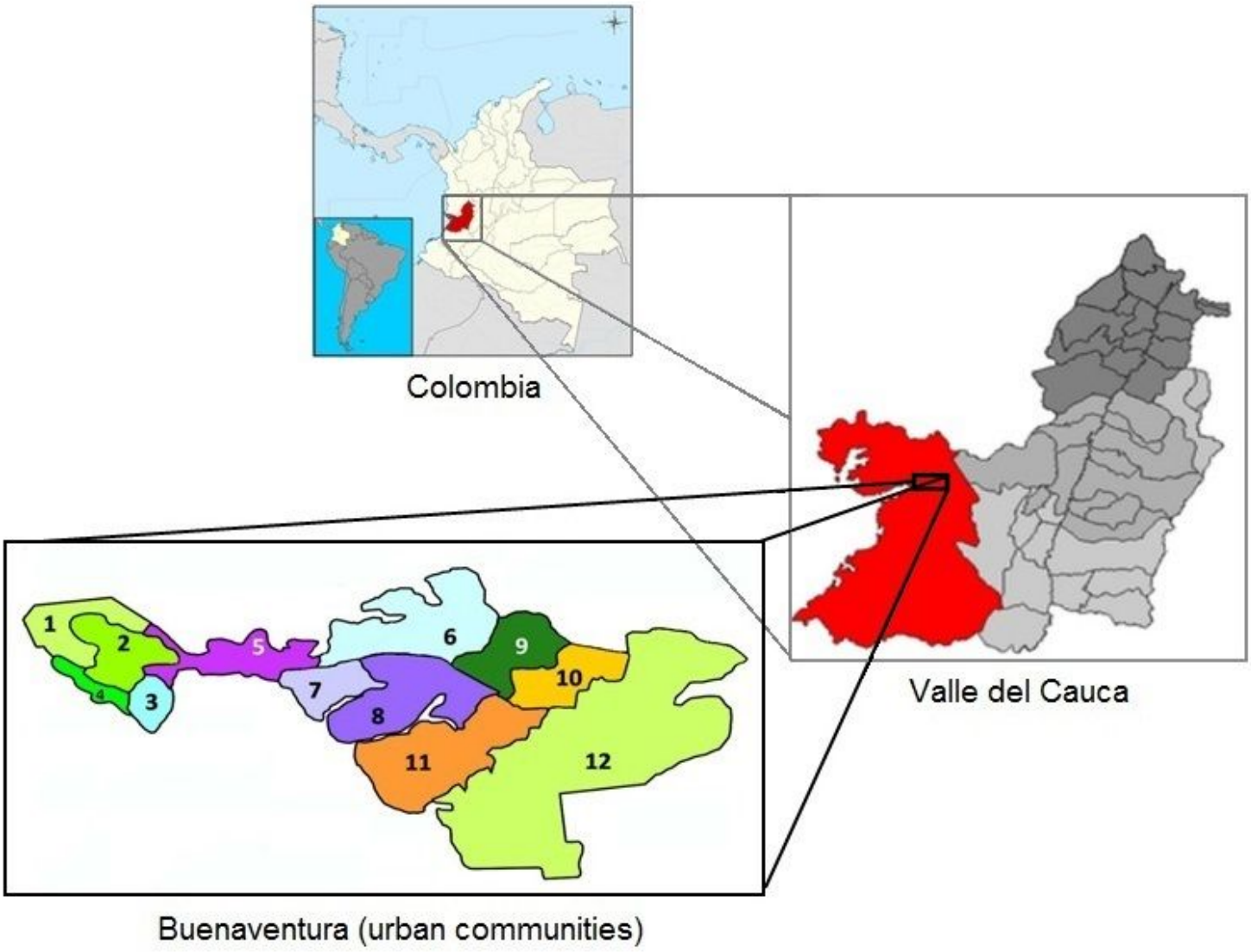

Figure 1

Map of the communities in the city of Buenaventura, located in the Department of Valle del Cauca. The department is situated in Colombia, a country located in South America. From Ortega et al. [28]. Licensed under CC BY 4.0: https://creativecommons.

org/licenses/by/4.0/. Modified from original. Available at https://academicjournals.
org/journal/IJGMB/article-abstract/8E4460450921

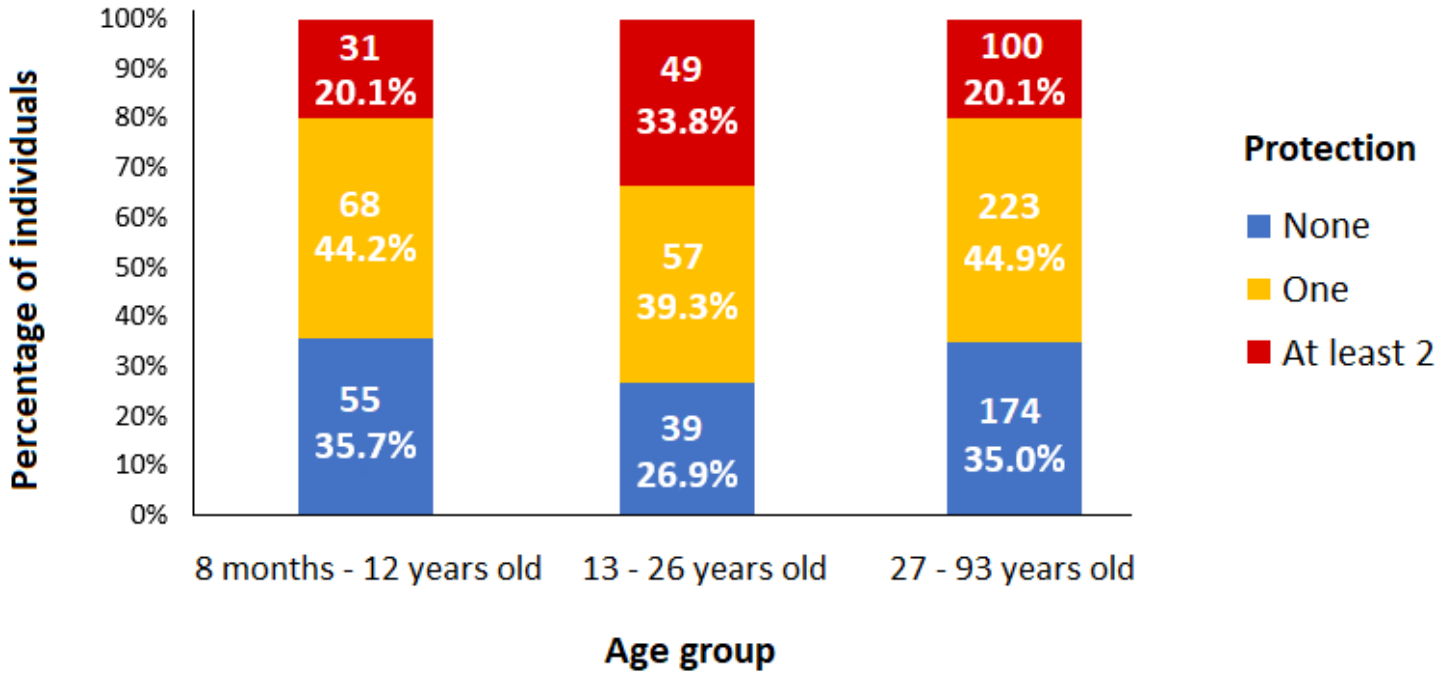

Figure 2

Bar chart illustrating age groups in relation to the 3 levels of the “protection” variable.

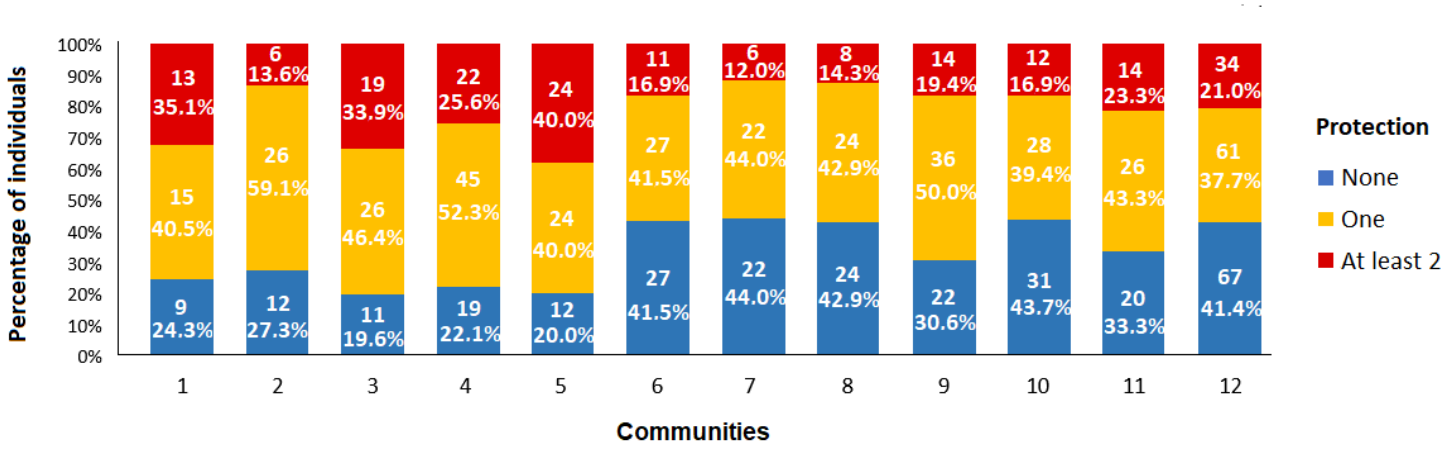

Figure 3

Bar chart illustrating the 12 communities of the city of Buenaventura in relation to the 3 levels of the “protection” variable.

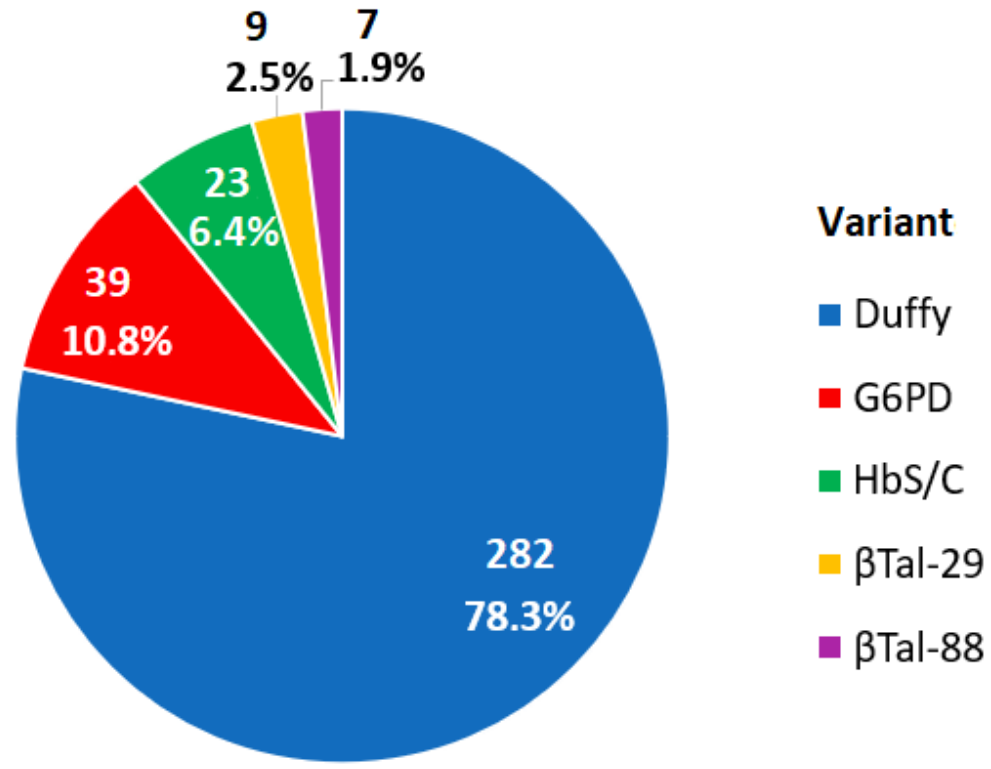

Figure 4

Pie chart depicting frequencies of resistance variants within the "one" category of the "protection" variable.

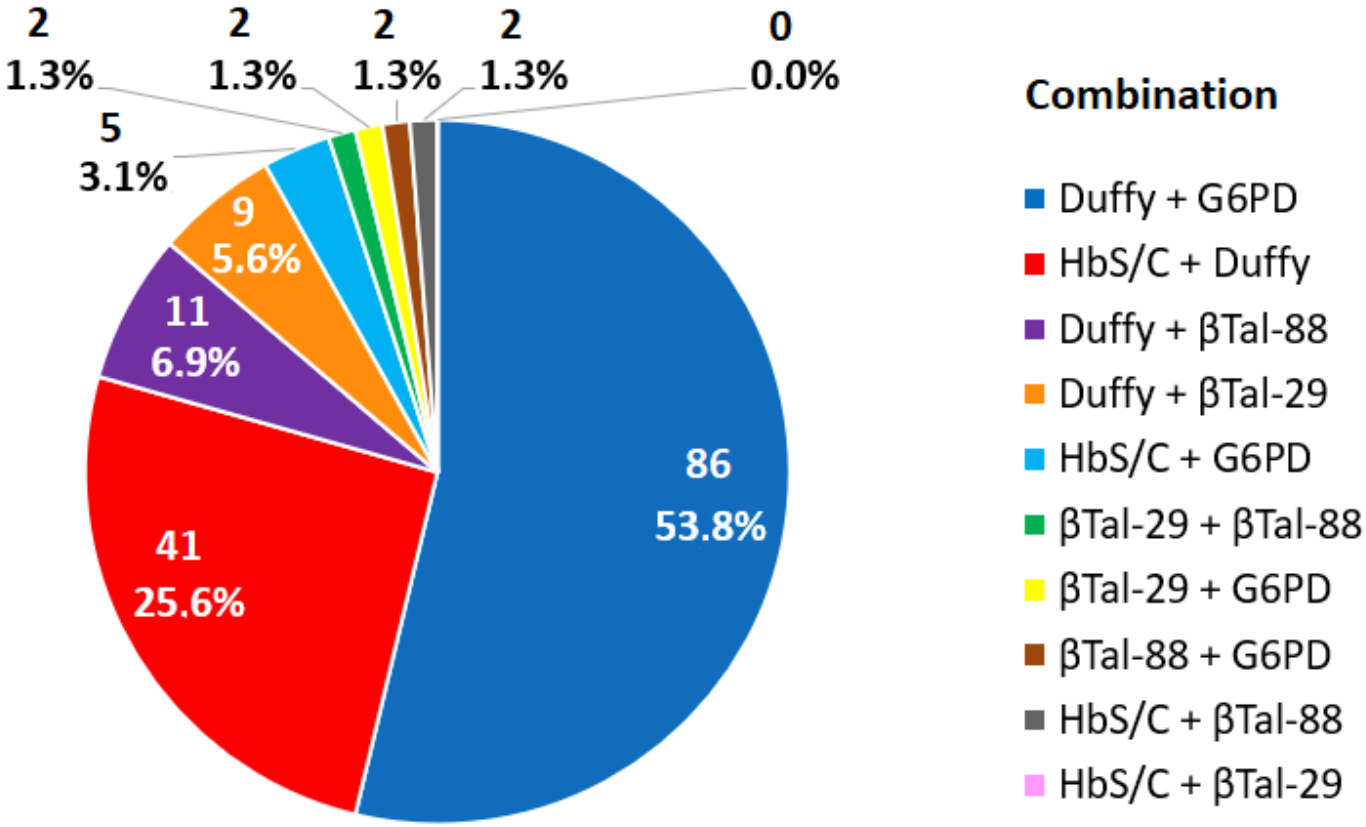

Figure 5

Pie chart for double variant combinations within the "at least two" category of the "protection" variable.

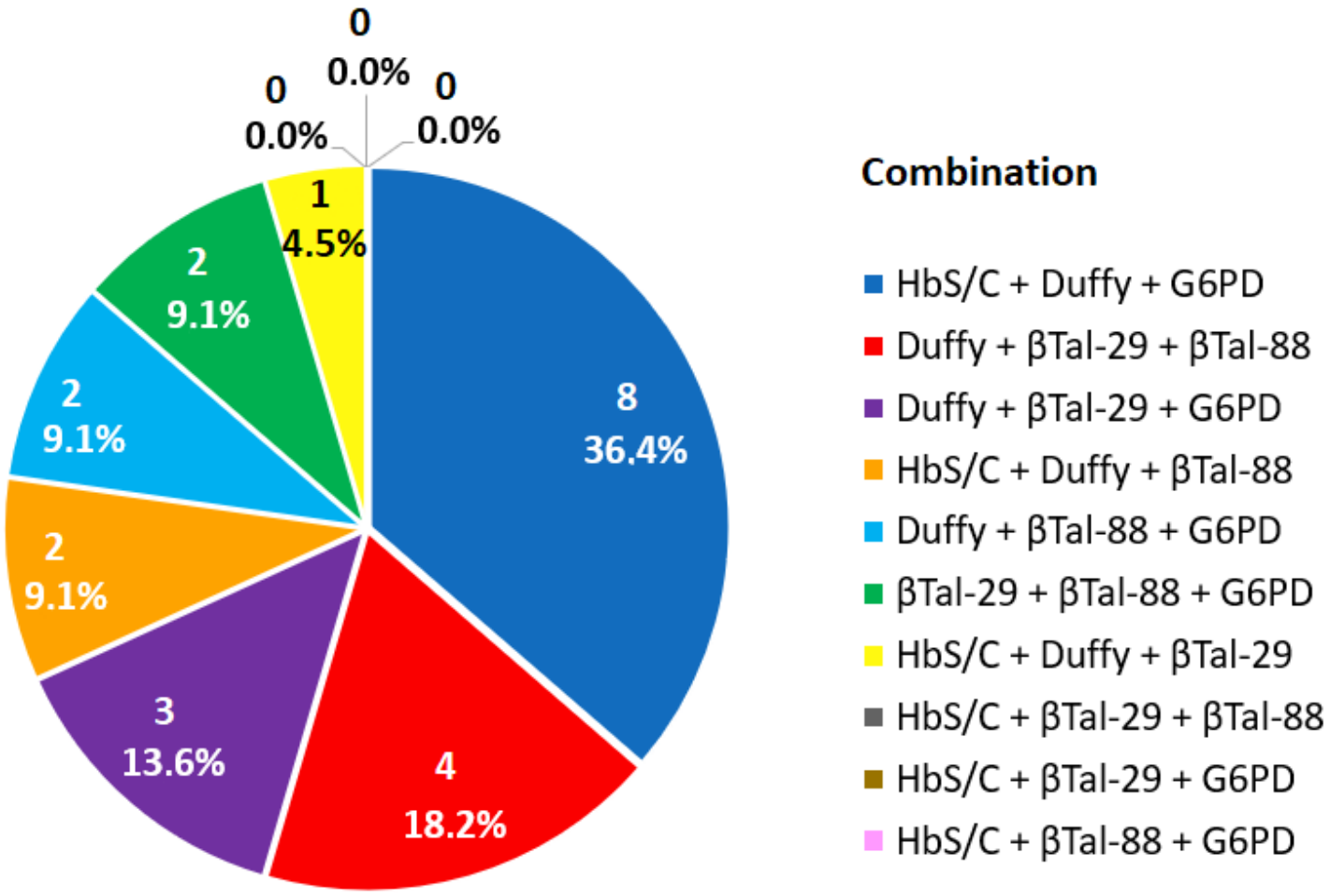

Figure 6

Pie chart for triple variant combinations within the "at least two" category of the "protection" variable.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table8.docx