

# History, geography and population structure influence the distribution and heritability of blood and anthropometric quantitative traits in nine Sardinian genetic isolates

LAURA PORTAS<sup>1\*</sup>†, FEDERICO MURGIA<sup>1\*</sup>, GINEVRA BIINO<sup>1,2</sup>,  
MARIA P. CONCAS<sup>1</sup>, LAURA CASULA<sup>1</sup>, STEFANIA MILIA<sup>1</sup>,  
MICHAEL B. WHALEN<sup>2</sup>, SIMONA VACCARGIU<sup>1,2</sup>, MASSIMILIANO COSSO<sup>2</sup>,  
DEBORA PARRACCIANI<sup>3</sup>, BRUNO FRONGIA<sup>1</sup> AND MARIO PIRASTU<sup>1,2</sup>

<sup>1</sup> Institute of Population Genetics, National Council of Research, Sassari, Italy

<sup>2</sup> Sharda Life Sciences, Pula (Cagliari), Italy

<sup>3</sup> Parco Genetico dell'Ogliastra, Perdasdefogu, Italy

(Received 17 March 2010 and in revised form 25 May 2010)

## Summary

Isolated founder populations which exhibit great genetic and environmental homogeneity provide an attractive setting for the study of quantitative traits (QTs). Geneticists have repeatedly turned to population isolates and the past successes have prompted increased interest among medical researchers. We studied nine small isolated villages of a secluded area of Sardinia (Ogliastra), all of them characterized by a few founders, high endogamy rates, slow population expansion and a distinct genetic makeup. Anthropometric and blood parameters, 43 QTs in all, were analysed in about 9000 voluntary subjects for whom extended genealogical information was available. We explored the distribution and examined mean differences of each trait among villages by analysis of variance (ANOVA). A heritability analysis with the variance component (VC) method was performed. Results show significant differences in the distribution of most traits between groups of villages located in two distinct geographical areas already identified by a previous population structure analysis, thus supporting the existence of differentiation among sub-populations in the same region. Heritability estimates range between 30 and 89%, demonstrating that genetic effects substantially contribute to phenotypic variation of all investigated traits and that this population provides excellent research conditions for gene-mapping projects. Results suggest that history, geographic location and population structure may have influenced the genetic and phenotypic features of these isolates. Our findings may be useful for the ongoing linkage and association studies in these isolates and suggest that a thorough characterization of population is valuable to better identify genes or variants that may be rare in the population at large and peculiar to single villages.

## 1. Introduction

Most observable differences among individuals in physiology, behaviour, morphology, disease susceptibility and reproductive fitness are measurable and present a quantitative distribution.

Quantitative genetic variation is the substrate for phenotypic evolution in natural populations and also underlies susceptibility to common complex diseases and behavioural disorders in humans, as well as responses to pharmacological therapies. In medicine, an important goal is to identify genetic risk factors for the presence of many common diseases. In some cases, however, quantitative traits (QTs) may provide more information for identifying genes than qualitative traits. They may represent intermediate traits or endophenotypes for diseases. An understanding of

† Corresponding author: Institute of Population Genetics, National Council of Research, Traversa La Crucca, 3 Regione Balduca, 07100 Sassari, Italy. Tel: (+39)079-2841301. Fax: (+39)079-2841399. e-mail: laura.portas@igp.cnr.it

\* These authors contributed equally to this work.

the genetic and environmental factors causing QTs variation is thus relevant in a number of biological contexts, including medicine, evolution and the emerging discipline of system biology.

Phenotypic variation for QTs results from the segregation of alleles at multiple QT loci with effects that are sensitive to the genetic, gender and external environments. For complex traits, the relationship between genotype and phenotype is not a simple ratio, and inferences about the net effects of all loci affecting the trait can be made by partitioning the total phenotypic variance into components attributable to additive, dominance and epistatic genetic variance, variance of genotype–environment interactions and environmental variance. These variances are population-specific, due to the dependence of the genetic terms on allele frequencies at each of the contributing loci, and to real environmental differences between populations. The relative weight of genetic factors in determining the variability of a QT is usually evaluated by means of heritability estimates. A promising approach for identifying genes affecting the complex and QTs is the study of population isolates, which allow a reduction in the complexity of the genetic models underlying the trait. Although generalizability of findings on genetic isolates to outbred populations are often uncertain, the former have the advantage of a more homogeneous genetic pool, having undergone low immigration and expansion, and of limiting potential confounders related to genetic heterogeneity. In isolated populations, the founder effect and genetic drift can cause an increase in the frequency and the attributable risk of particular alleles, maybe broadening the magnitude of effect sizes. The presence of both inbreeding and extended genomic regions in linkage disequilibrium (LD) with previously mentioned factors may generate a more homogeneous genetic background. Moreover, the members of isolated populations share a common environment and lifestyle, thus the environmental diversity is strongly reduced. In these genetically and culturally homogeneous populations, a common remote genetic pool produced a large proportion of individuals who inherited the same trait-predisposing genes.

In our study, we describe 43 QTs in the population of nine isolated villages (Baunei, Loceri, Perdasdefogu, Seui, Seulo, Talana, Urzulei, Ussassai and Triei) from the secluded Ogliastra region in Sardinia (Fig. 1). This area represents an isolate that could provide a significant advantage in studying complex traits because of its reduced genetic diversity and its environmental homogeneity. Mitochondrial analysis traced the original population back to the Neolithic era and showed that Ogliastra inhabitants rank among the most genetically homogenous European population and that they have the lowest values of mtDNA gene diversity with respect to other



Fig. 1. Geographical location of the nine isolates.

Sardinia areas (Fraumene *et al.*, 2003; Falchi *et al.*, 2006). Several Ogliastra villages slowly grew in isolation and had little or no admixture with the others, giving rise to high endogamy (from 64% in Loceri to 92% in Baunei) and inbreeding. Previous genealogical studies determined that villages used in our study have very few maternal and paternal lineages: ten mitochondrial lineages and eight Y chromosomes represent 80–85% of the entire population of each village (Angius *et al.*, 2001; Ghiani & Vona, 2002; Fraumene *et al.*, 2006 and unpublished data). Additional analyses of the population structure through high-density SNPs confirmed a great deal of genetic differentiation among sub-population isolates within Ogliastra as a consequence of distinct founder effects and genetic drift (Pistis *et al.*, 2009). All these data demonstrate that history and geography have influenced the genetic features of Ogliastra communities producing differences in LD and the population structure and allowing us to consider each village as an isolate with specific characteristics. In addition, a good parallelism with the geographical structure of the region was observed, since villages cluster according to their geographic position and relative proximity. In particular, statistically significant genetic heterogeneity was detected between two groups of villages: Seui, Seulo, Ussassai and Perdasdefogu (group 1) and Talana, Baunei, Urzulei

Table 1. Number of phenotyped individuals and family features for each village

Village	N	Males	Females	Families	Family size	Generations
Baunei	2002	833	1169	117	31.6	4.5
Loceri	744	329	415	33	47.4	5.3
Perdasdefogu	1303	563	740	66	33.6	4.4
Seui	937	399	538	44	40.9	4.4
Seulo	743	339	404	43	31.4	4.8
Talana	1070	470	600	117	15.1	3.5
Triei	854	371	483	24	60.6	5.0
Urzulei	847	347	500	45	34.1	4.6
Ussassai	376	149	227	22	35.3	4.8
Total	8877	3831	5046	511	36.7	4.6

and Triei (group 2). In contrast, Loceri can be considered as a freestanding group since it is characterized by a minor geographical isolation, due to its closeness to the seashore and the important administrative centre of Lanusei and by the lowest LD values. Furthermore, archival data, spanning many centuries, confirmed its frequent genetic exchanges with other villages (Pistis *et al.*, 2009).

Here, we present the study of distribution and heritability of 43 QTs in nine isolated villages of Ogliastra region. Studied phenotypes range from haematological and serum traits to anthropometric measures.

## 2. Materials and methods

### (i) Study design and genealogical data

Study sample comes from a large epidemiologic survey on complex diseases carried out in nine isolated villages of Ogliastra, a Central–Eastern Sardinia region that has been geographically and socially secluded for thousands of years, due to mountains and deep river valleys. The nine selected communities, with an average population of about 1500 inhabitants, share similar demographic features: limited number of founders, high endogamy, consanguinity and slow population growth. Briefly, the study design is cross-sectional and population-based. People living in the villages were invited to take part in the study by means of information campaigns and letters sent to every family. Furthermore, word of mouth, especially among family members, made a high level of recruitment possible. The participants gave a blood sample, underwent anthropometric measurements and a standardized interview collecting socio-demographic, lifestyle, medical and pharmacological history data. The research adheres to the tenets of the Declaration of Helsinki and written informed consent was obtained from all enrolled people. About 80% of the resident population participated in the study.

For each inhabitant, we collected genealogical information dating back to the 17th century and stored

them in a relational database along with clinical and epidemiological data. To construct pedigrees, we used a tool called FamPrep, implemented in the PedNavigator software (Mancosu *et al.*, 2003), which allow the reconstruction of different types of genealogies starting from a list of individuals and setting the number of generations. Pedigree structures are such that every individual in the study sample is related to every other individual through multiple lines of descent. Available data allowed us to organize 15 158 individuals, – 7916 females and 7242 males – into 511 pedigrees (on average 4.6 generations deep) with an average family size of 36.7 individuals. The overall sample includes 9024 sibling pairs, 432 half-sibling pairs, 13 373 first cousins, 19 938 parent–child pairs, 17 610 grandparent–grandchild pairs and 16 663 avuncular pairs in addition to other more distant relatives. Statistics about the family structures of each village are displayed in Table 1.

### (ii) Phenotypic data

The sample analysed in the present study consists of 8876 males (43%) and females (57%), aged over 18 years. From each participant, blood aliquots were fractionated to provide serum, plasma and white blood cells.

Blood cell counts were performed by Beckman Coulter, whereas serological parameters and the corresponding levels were determined with an automated Targa BT-3000 Chemistry Analyser. Body mass index (BMI) was calculated as weight (in kilograms) divided by height squared (in meters).

Overall, we collected data on 43 quantitative phenotypes: 3 anthropometric, 17 haematological and 23 serum traits (Table 2).

### (iii) Statistical analysis

We explored the distribution of each trait on the whole sample generating summary plots which allowed us to identify outliers falling outside the

Table 2. Sample size, overall and sex-specific traits mean values and F test P-value of sex-specific traits variance

	Overall		Males		Females		Variance test P-value
	N	Mean	N	Mean	N	Mean	
<b>Haematological traits</b>							
BA (basophils, %)	8877	0.6	3831	0.6	5046	0.5	****
EO (eosinophils, %)	8877	3.4	3831	3.6	5046	3.1	****
HCT (hematocrit, %)	8348	42.0	3586	44.4	4762	39.6	****
Hb (haemoglobin, g/dl)	8348	14.2	3586	15.1	4762	13.4	****
LY (lymphocytes, %)	8877	35.5	3831	35.4	5046	35.7	N.S.
MCH (mean RBC Hb, pg)	8348	28.9	3669	29.4	4762	28.5	****
MCHC (mean corpuscular Hb conc, g/dl)	8348	33.9	3669	34.0	4762	33.8	****
MCV (mean RBC vol, fl)	8348	85.2	3669	86.3	4762	84.1	****
MO (monocytes, %)	8876	7.8	3831	8.2	5045	7.3	****
MPV (mean PLT vol, fl)	8608	9.2	3562	9.1	5046	9.4	****
NE (neutrophils, %)	8537	52.8	3655	52.2	4882	53.3	****
PCT (PLT hematocrit, %) <sup>1</sup>	2635	0.2	1105	0.2	1530	0.2	****
PDW (PLT distribution width, fl) <sup>1</sup>	2535	16.3	1072	16.2	1463	16.5	*
PLT (platelets, 10 <sup>6</sup> ml)	8341	242.0	3585	230.2	4756	253.8	****
RBC (red blood cells, 10 <sup>6</sup> ml)	8348	5.0	3669	5.2	4762	4.8	****
RDW (red cells dispersion width, %)	8348	12.6	3669	12.5	4762	12.7	****
WBC (white blood cells, 10 <sup>3</sup> ml)	8876	7.1	3831	7.3	5045	6.9	****
<b>Serological traits</b>							
Albumin (g/dl)	8607	4.2	3562	4.3	5045	4.2	****
Alkaline phosphatase (units/l)	8608	141.9	3562	143	5046	140.4	****
ALT (alanine aminotransferase, units/l)	8608	28.6	3562	33.9	5046	23.3	****
AST (aspartate aminotransferase, units/l)	8608	22.8	3562	25.1	5046	20.4	****
BC (direct bilirubin, mg/dl)	8607	0.2	3562	0.2	5045	0.2	****
BU (indirect bilirubin, mg/dl)	8608	0.5	3514	0.6	5046	0.5	****
TBIL (total bilirubin, mg/dl)	8608	0.7	3562	0.8	5046	0.7	****
BUN (blood urea nitrogen, mg/dl)	8608	31.4	3562	33.2	5046	29.6	****
Calcium (mg/dl)	8608	9.2	3562	9.2	5046	9.2	****
Chlorine (mEq/l)	8630	101.8	3778	101.6	4852	102.1	****
Cholinesterase (units/l)	8613	7547.8	3689	7867.9	4924	7227.7	****
HDL cholesterol (mg/dl)	8756	54.6	3818	50.6	4938	58.6	****
LDL cholesterol (mg/dl)	8755	127.1	3817	127.4	4938	126.8	N.S.
Magnesium <sup>2</sup> (mg/dl)	6369	2.0	2675	2.0	3694	1.9	****
Phosphorus (mg/dl)	8429	3.2	3512	3.1	4917	3.2	****
Potassium (mEq/l)	8608	4.2	3562	4.2	5046	4.2	****
Serum creatinine (mg/dl)	8608	1.0	3562	1.1	5046	0.9	****
Serum glucose (mg/dl)	8756	96.5	3818	99.8	4938	93.2	****
Sodium (mEq/l)	8608	138.1	3562	138.2	5046	138.1	N.S.
Total cholesterol (mg/dl)	8756	203.7	3818	202.4	4938	205.0	**
Total protein (g/dl)	8613	7.3	3689	7.3	4924	7.3	N.S.
Triglycerides (mg/dl)	8756	108.6	3818	120.2	4938	96.9	****
Uric acid (mg/dl)	8608	4.6	3562	5.4	5046	3.7	****
<b>Anthropometric measures</b>							
Height (cm)	7780	158.9	3277	165.2	4503	152.7	****
BMI (kg/m <sup>2</sup> )	7780	25.9	3277	26.5	4503	25.3	****
Weight (kg)	7780	65.4	3277	72.1	4503	58.6	****

\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

<sup>1</sup> Data available for Perdasdefogu, Talana and Urzulei.

<sup>2</sup> Data not available for Triei and Loceri.

$\pm 3$  standard deviations (SDs) range of the distribution. Such outliers were then excluded from the subsequent analysis. We next computed descriptive overall statistics for all traits according to sex, and performed  $F$  test on the variance of traits in order to assess the variability pattern in males and females.

Subsequently, we examined differences in the traits' mean values among villages by analysis of variance

(ANOVA), adjusting for age and gender and using Bonferroni correction. Whenever required, we normalized traits using the appropriate transformation.

To quantify the relative genetic contribution to the variation of each trait, we carried out a heritability analysis with the standard variance component (VC) approach based on maximum likelihood methods (Hartley & Rao, 1967) as implemented in the SOLAR

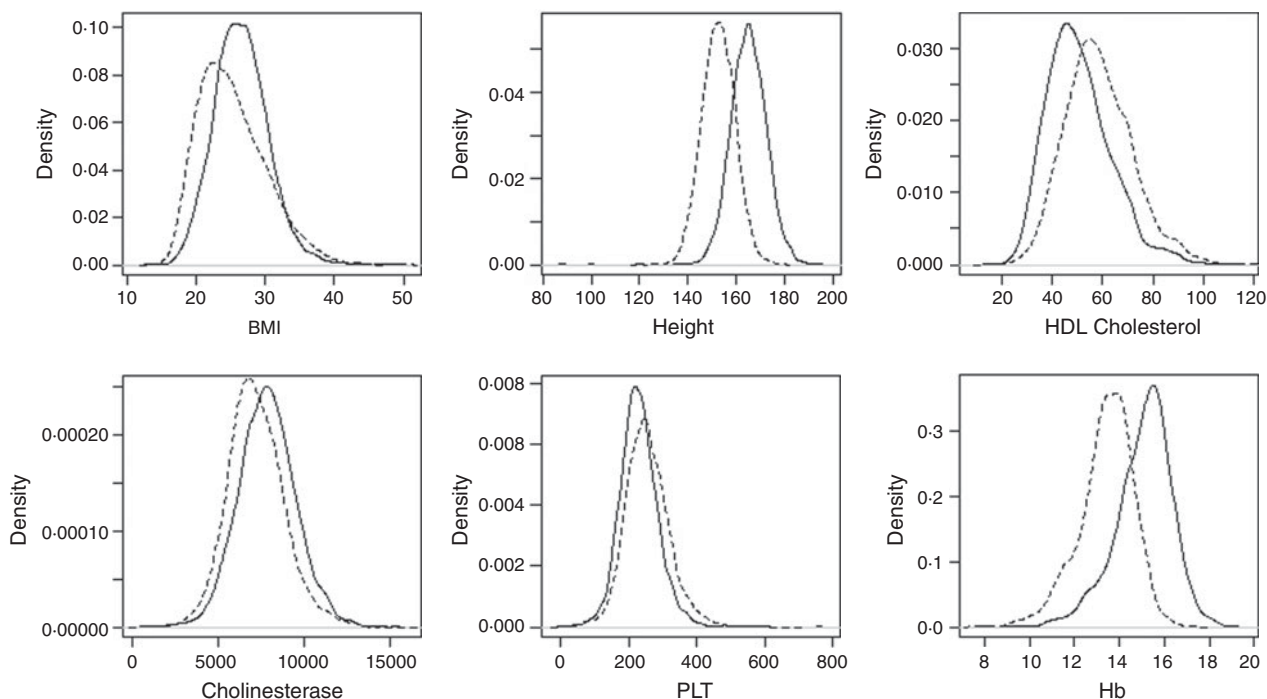


Fig. 2. Distribution of six representative traits in male and female participants. Relative densities were plotted for females (dotted lines) and males (continuous line) for 2 blood cell measures (Hb and PLT), 2 serum (LDL cholesterol and HDL cholesterol) and 2 anthropometric measures (BMI and height).

software (Almasy & Blangero, 1998). Since each village has a particular founder population and may represent a genetic sub-isolate, separate heritability estimates were obtained. We tested the influence of age and gender on each trait, and covariates with significant effect ( $P \leq 0.05$ ) were included in the analysis. We estimated a model including additive genetic ( $V_a$ ) and residual environmental ( $V_e$ ) VCs: heritability was obtained as the ratio of the additive genetic variance to the sum of the additive genetic and environmental variance.

To evaluate sex-specific effects on the variation of investigated traits, we used a standard sex-limitation modelling approach that allows testing specific patterns of interaction, such as genotype by sex ( $G \times S$ ) interactions (Martin *et al.*, 2002). The expected genetic covariance between a male and female relative pair is defined as covariance ( $G_M, G_F$ ) =  $2\phi \rho G_{(M,F)} \sigma G_M \sigma G_F$ , where  $\phi$  is the coefficient of kinship between the two individuals,  $\rho G_{(M,F)}$  the genetic correlation between the expressions of the trait in the two sexes and  $\sigma G_M$  and  $\sigma G_F$  the genetic SDs for men and women. In the absence of a  $G \times S$  interaction, the genetic correlation between relatives for a trait measured in males and females should be 1 and the genetic variances in the two groups should be equal. Conversely, if there is  $G \times S$  interaction, the genetic correlation between the sexes, is significantly  $< 1$  and/or the genetic variances are different between the sexes. To evaluate  $G \times S$  interactions influencing the traits, a likelihood ratio test (LRT) was used to

compare the three nested models, obtained constraining  $\rho G = 1$ ,  $\sigma G_M = \sigma G_F$  and  $\sigma E_M = \sigma E_F$ , respectively, to the full model, where  $\sigma E_M$  and  $\sigma E_F$  are the environmental SD in males and females.

The R statistical software (v.2.5.0) was used for all the analyses except for VC heritability analysis, which was performed using SOLAR (Sequential Oligogenic Linkage Analysis Routine, version 4.1.0).

### 3. Results

We studied 3 anthropometric, 17 haematological and 23 serum traits. For each trait, the total number of phenotyped individuals included in the analysis exceeds 8000 individuals except for a small number of traits available only in a few villages. Indeed, for platelets distribution width (PDW), the number of phenotyped individuals included in the analysis was 2535 (data available for Perdasdefogu, Talana and Urzulei), for piastrinocrite (PCT) blood samples were collected only in Perdasdefogu, Talana and Urzulei ( $n = 2635$ ) and for magnesium data were not available in Triei and Loceri ( $n = 6369$ ). Anthropometric measures were available for 7780 individuals.

Table 2 shows descriptive overall statistics of all studied traits, according to sex. For many traits there were marked differences between genders, affecting not only the trait means but also the overall pattern of variability around the mean. We detected 37 traits with a highly significant ( $P < 0.0001$ ) evidence for variance heterogeneity by sex (Fig. 2). The 37 traits

included all 3 anthropometric traits, many haematological (15 out of 17) and serological traits (19 out of 23).

ANOVA of all 43 QTs among the nine villages showed statistically significant differences for most of them (Fig. 1, Supplementary materials). Looking at these results, we realized that for many traits some villages looked more alike to each other and markedly differed from other ones. In particular, the observed differences among trait mean values were such in order to group villages in a way which resembled the clustering outlined on the basis of the population's genetic structure (Pistis *et al.*, 2009). We therefore performed ANOVA again, using population as a factor, after grouping villages as follows: Talana, Baunei, Urzulei and Triei (group 1 – northern Ogliastra), Seui, Seulo, Ussassai and Perdasdefogu (group 2 – southern Ogliastra) and Loceri (group 3).

This last analysis showed statistically significant differences between groups of villages for most of the traits (Table 3). In particular, we observed differences among all groups in: electrolytes (sodium, potassium, phosphorus, chlorine and calcium), 6 serological traits (total protein, alkaline phosphatase, cholinesterase, LDL cholesterol, total and indirect bilirubin), 4 haematological traits (RDW, RBC, MCHC and MPV) and 1 anthropometric measure (BMI). In contrast, we never observed significant differences in 2 haematological measures (HCT and neutrophils) and in 4 serological traits (ALT transaminase, triglycerides and uric acid).

Most differences (79% of the traits) were detected between group 1 and 2 for 14 haematological, 18 serological and 2 anthropometric measures. Furthermore, for 77% of traits Loceri, which represents a freestanding group, resulted as more similar to villages of group 1, whereas for the remainder 23% it resembled those of group 2.

Heritability analysis of 43 QTs by means of the VC method was separately performed by village. Only 26 out of 373 heritability estimates were not significantly different from zero (Fig. 2, Supplementary materials). Statistically significant heritability estimates varied from 12% (basophils in Perdasdefogu) to 88% (MCH in Seulo and Triei) among haematological traits, from 12% (ALT Transaminase in Baunei) to 84% (direct bilirubin in Seulo) among serological traits and, from 23% (BMI in Seulo) to 81% (height in Ussassai) among anthropometric measures. Most of the traits are rather stable and heritability estimates were similar in all the nine isolates except for basophils (Triei higher than Perdasdefogu), MCHC (Baunei, Perdasdefogu, Ussassai and Triei higher than Talana and Urzulei), RDW (Triei higher than Baunei), direct bilirubin (Seulo higher in respect with all other villages), LDL cholesterol (Perdasdefogu higher than Seui), total cholesterol (Perdasdefogu and

Urzulei higher than Loceri) and triglycerides (Perdasdefogu higher than Seui and Loceri). As we observed sporadic statistically significant differences among villages only in a few traits, we also carried out a further analysis pooling together pedigrees of all the villages. In supplementary table 1, heritability estimates obtained for each QT on the total sample are reported. After accounting for age and sex, heritability estimates ranged between 34% (basophils) and 76% (MCV and MCH) for the 17 haematological traits, between 30% (calcium and direct bilirubin) and 89% (alkaline phosphatase) for the 23 serum measures and, between 41% (BMI) and 73% (height) for the three anthropometric traits. Variance due to covariates ranged from 0.3% for PDW and 60% for height.

Since many traits were found to have significantly different variances in males and females, we also performed a sex-specific variance decomposition to explore the effects of genetic and environmental factors in males and females. Genetic and environmental SDs for men and women and genetic correlations in the 43 investigated QTs, along with sex-specific heritabilities, are shown in supplementary table 1. Genetic SDs of basophils, Hb, MCH and direct bilirubin were significantly higher in men, whereas genetic SDs of ALT transaminase, serum creatinine and serum glucose were significantly higher in women. Environmental SD of alkaline phosphatase, ALT transaminase, indirect bilirubin, LDL and total cholesterol, triglycerides and serum phosphorus were significantly higher in men, whereas BMI environmental SD was higher in women. The genetic correlations between sexes were significantly different from 1 for 3 haematological traits (basophils, eosinophils, MCHC) and 9 serological measures (ALT, indirect and total bilirubin, HDL, magnesium, phosphorus, potassium, serum creatinine, uric acid).

To better understand the nature of different heritabilities in genders for some traits, we also performed parent–offspring regression (Falconer & Mackay, 1996), which presupposes a linear relationship between the mean value of the trait in the offspring and the mean value of the trait in parents. Heritability is the regression coefficient of this linear model. Father–son heritability estimates of basophils, ALT and direct bilirubin were, respectively, about three and four times higher than mother–son ones, father–daughter heritability of direct bilirubin was about twice that of mother–daughter, whereas mother–offspring heritability of serum glucose was twice that of father–offspring (Table 4).

#### 4. Discussion

In the present study, we assessed the features and genetic contribution of 43 quantitative phenotypes in

Table 3. ANOVA of all 43 QTs among the 3 groups of villages, Bonferroni correction

	Mean trait values			Groups 1 versus 2		Groups 1 versus 3		Groups 2 versus 3	
	Group 1	Group 2	Group 3	Mean difference	P-value	Mean difference	P-value	Mean difference	P-value
<b>Haematological traits</b>									
BA (basophils, %)	0.546	0.554	0.497	0.008	NS	0.05	**	0.05	****
EO (eosinophils, %)	3.271	3.419	3.342	0.15	**	0.07	NS	0.08	NS
HCT (hematocrit, %)	42.530	42.018	40.704	0.51	NS	1.82	NS	1.31	NS
Hb (g/dl)	14.336	13.987	13.927	0.35	****	0.41	****	0.07	NS
LY (lymphocytes, %)	35.370	35.825	35.333	0.45	*	0.04	NS	0.49	NS
MCH (mean RBC Hb, pg)	28.988	28.688	29.100	0.30	***	0.11	NS	0.41	*
MCHC (mean corpuscular Hb conc, g/dl)	33.709	34.076	34.215	0.36	****	0.5	****	0.14	***
MCV (mean RBC vol, fl)	85.824	84.042	84.971	1.78	****	0.85	NS	0.93	*
MO (monocytes, %)	8.065	7.363	8.056	0.70	****	0.01	NS	0.69	****
MPV (mean PLT vol, fl)	9.039	9.351	8.875	0.31	****	0.17	****	0.48	****
NE (neutrophils, %)	52.744	52.833	52.771	0.09	NS	0.03	NS	0.06	NS
PCT (PLT hematocrit, %)	0.223	0.213	0.231	0.01	**	0.01	NS	0.02	NS
PDW (PLT distribution width, fl)	16.000	16.918	16.500	0.92	****	0.92	NS	0.42	NS
PLT (platelets, 10 <sup>6</sup> ml)	256.281	234.879	250.81	21.40	****	5.47	NS	15.93	****
RBC (10 <sup>6</sup> ml)	4.997	4.931	4.821	0.07	****	0.08	***	0.11	****
RDW (%)	12.530	12.621	12.179	0.09	*	0.34	****	0.42	****
WBC (white blood cells, 10 <sup>3</sup> ml)	6.911	7.245	7.025	0.33	****	0.11	NS	0.22	*
<b>Serological traits</b>									
Albumin (g/dl)	4.318	4.199	4.208	0.12	****	0.11	****	0.01	NS
Alkaline phosphatase (units/l)	132.811	141.091	187.425	8.28	****	54.61	****	46.33	****
ALT (alanine aminotransferase, units/l)	27.23	28.042	25.073	0.81	NS	2.16	NS	2.97	**
AST (aspartate aminotransferase, units/l)	21.922	22.585	21.614	0.66	NS	0.31	NS	0.97	NS
Direct bilirubin (mg/dl)	0.203	0.212	0.210	0.01	****	0.01	NS	0	NS
Indirect bilirubin (mg/dl)	0.540	0.468	0.387	0.07	****	0.15	****	0.08	****
Total bilirubin (mg/dl)	0.739	0.677	0.594	0.06	****	0.15	****	0.09	****
BUN (mg/dl)	32.920	29.710	32.275	3.21	****	0.65	NS	2.56	****
Calcium (mg/dl)	9.230	9.196	9.043	0.03	**	0.19	****	0.16	****
Chlorine (mEq/l)	101.804	102.411	98.084	0.61	****	3.72	****	4.33	****
Cholinesterase (units/l)	7655.382	6143.183	8157.525	1512.2	****	502.14	****	2014.34	****
HDL cholesterol (mg/dl)	57.107	53.749	54.229	3.36	****	2.88	****	0.48	NS
LDL cholesterol (mg/dl)	130.921	123.380	137.55	7.54	****	6.64	****	14.18	****
Magnesium (mg/dl)	1.981	1.919	1.950	0.06	****	0.06	NS	0.03	NS
Phosphorus (mg/dl)	3.192	3.155	2.994	0.04	**	0.2	****	0.16	****
Potassium (mEq/l)	4.147	4.191	4.295	0.04	****	0.14	****	0.1	****
Serum creatinine (mg/dl)	0.981	0.974	1.040	0.01	NS	0.06	****	0.07	****
Serum glucose (mg/dl)	98.232	93.652	97.243	4.58	****	0.99	NS	3.59	***
Sodium (mEq/l)	138.769	138.03	136.195	0.74	****	2.58	****	1.83	****
Total cholesterol (mg/dl)	209.282	198.846	212.606	72.72	****	3.32	NS	13.76	****
Total protein (g/dl)	7.349	7.230	7.293	0.12	****	0.06	**	0.06	**
Triglycerides (mg/dl)	106.077	107.552	103.544	1.47	NS	2.54	NS	4.01	NS
Uric acid (mg/dl)	4.430	4.412	4.430	0.02	NS	0	NS	0.02	NS
<b>Anthropometric measures</b>									
Height (cm)	159.048	157.298	158.467	1.75	****	0.57	NS	1.18	**
BMI (kg/m <sup>2</sup> )	25.509	25.727	26.75	0.22	NS	1.24	****	1.02	****
Weight (kg)	64.613	63.805	67.241	0.81	*	2.63	****	3.44	****

\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

Table 4. Heritability and 95% CI by means of parent–offspring regression

	Basophils $h^2$		ALT $h^2$	
	Fathers	Mothers	Fathers	Mothers
Sons	0.45 (0.38–0.53)	0.13 (0.06–0.20)	0.38 (0.32–0.44)	0.13 (0.08–0.18)
Daughters	0.38 (0.32–0.43)	0.46 (0.41–0.51)	0.19 (0.16–0.22)	0.23 (0.19–0.26)
	Serum Creatinine $h^2$		Hb $h^2$	
	Fathers	Mothers	Fathers	Mothers
Sons	0.53 (0.48–0.58)	0.55 (0.50–0.60)	0.36 (0.30–0.41)	0.46 (0.41–0.51)
Daughters	0.52 (0.48–0.56)	0.72 (0.68–0.76)	0.43 (0.38–0.48)	0.47 (0.42–0.51)
	Serum Glucose $h^2$		MCH $h^2$	
	Fathers	Mothers	Fathers	Mothers
Sons	0.14 (0.09–0.18)	0.24 (0.21–0.27)	0.74 (0.67–0.80)	0.86 (0.81–0.91)
Daughters	0.12 (0.09–0.15)	0.22 (0.19–0.25)	0.80 (0.74–0.86)	0.77 (0.72–0.82)
	Direct Bilirubin $h^2$			
	Fathers	Mothers		
Sons	0.80 (0.64–0.97)	0.22 (0.12–0.32)		
Daughters	0.29 (0.25–0.33)	0.16 (0.12–0.20)		

nine villages of the Ogliastra region in the central-eastern area of Sardinia. This region is characterized by high endogamy, low immigration, environmental homogeneity as well as genetic differentiation from the rest of the island (Cappello *et al.*, 1996). We collected information on a large number of traits spanning from anthropometric measures to hematological and serum markers of diseases.

For most of the traits, we observed evidence of heterogeneity by sex which often reflects the known sexual dimorphism of traits, for example anthropometric measures (weight, height and BMI). These results are consistent with others studies performed in isolated populations, as the Hutterites (Weiss *et al.*, 2006) or the Sardinians (Pilia *et al.*, 2006), and in the general outbred population.

While analysing trait distribution in order to detect potential differences or similarities among villages, we observed that villages clustered accordingly to already known population characteristics (Pistis *et al.*, 2009). In particular, the results showed significant differences in the distribution of most traits (about 79% of them) between the group including Talana, Triei, Urzulei and Baunei and the one with Ussassai, Seui, Seulo and Perdasdefogu. These data seem to reflect the same correspondence between LD extension and population structure characteristics with the village's geographic location, as observed in our previous studies. Northern Ogliastra differs from southern Ogliastra not only from a genetic point of view, but

also phenotypically, with the village of Loceri in an intermediate position. Loceri, as a matter of fact, is not geographically isolated as the other villages, being close to the seashore and to Lanusei, an important administrative centre in Ogliastra. We suggest that such results may be partly attributed to complex relationships reflecting the different influence of history and biodemography on the genetic features of these isolates. Distinctive founder effects and genetic drift may have given rise to a striking differentiation among sub-populations in the same region.

Heritabilities of the investigated traits, except for a few, were not significantly different among villages, and so further estimates were computed on the overall sample. Such estimates are always higher than 30%, showing that all QTs have a strong genetic component of the variance. In particular, we estimated heritability of approximately 54% for haematological and anthropometric traits and of approximately 50% for serological traits. In general, our results appear to be consistent with previous family-based studies (Lin *et al.*, 2005; Malhotra & Wolford, 2005; Marroni *et al.*, 2008) and in particular with a study conducted in another Sardinian village (Pilia *et al.*, 2006), even if our cohort is larger than those analysed in previously published studies. In contrast, several traits showed smaller heritability estimates than in previous twins studies (Evans *et al.*, 1999; Beekman *et al.*, 2002), but twin designs may lead to an overestimation of the degree of genetic determination occurring from a



greater sharing of environmental factors in monozygotic rather than in dizygotic twins. Nevertheless, twin studies explicitly correct for age affects both the means and the covariance, while family- and population-based studies are generally only correct for the direct effect of age on the mean and do not correct the covariance for the age differences between family members.

Overall, the heritability of red cells indexes was quite large, accordingly with the high prevalence of alpha and beta thalassemia in Sardinia. On the contrary, electrolytes showed small values of heritability and, considering villages separately, they were often not statistically significant or extremely variable. We believe that environment plays an important role in electrolytes concentration, while genetic factors are less influential; furthermore, this result is in agreement with other studies (Pilia *et al.*, 2006).

As far as the sporadic statistically significant differences in heritability estimates we observed among some of the villages are concerned, it must be considered that different populations might have different heritabilities even for the same trait. This is due to the fact that heritability estimates are always relative to the genetic and environmental factors in the specific population, and are not absolute measurements of the contribution of genetic and environmental factors to a phenotype. Heritability estimates reflect the amount of variation in genotypic effects compared to variation in environmental effects. As a matter of fact, heritability is larger for outbred individuals (with diverse genetic background) or for little environmental effects, and is smaller for inbred individuals or for individuals reared in very diverse environments (Falconer & Mackay, 1996). In Ogliastra isolated populations, on one hand, we may assume a relatively homogeneous environment which tends to increase heritability, but on the other hand, we know that inbreeding is relatively high which tends to decrease heritability; nonetheless there is a high genetic differentiation between villages, and so we suggest that in these villages there may be different frequencies of allelic variants involved in the expression of some QTs, which may influence heritability in both directions.

Despite the heterogeneous variability in males and females, we observed for many traits, a sex-limited expression of genetic and/or environmental factors occurred in a few of them. Sex-limitation was qualitative for traits where the genetic correlation was significantly less than 1, indicating that different genes or subset of genes contribute to their variance in men and women; although it was quantitative for traits where genetic or environmental SDs were significantly different in men and women, indicating that the same factors affect both sexes, but their impact on the phenotype is greater in one sex than the other.

The findings are relevant for the design of genome-wide association studies and for investigations on the relationship of analysed traits with common diseases.

When we performed parent-offspring regression, with the aim of investigating the nature of different heritabilities in genders, for some traits sons showed stronger heritability to their fathers than to their mothers, whereas for other traits the opposite was true. However, the estimation of heritability by regression analysis is vulnerable to errors stemming from the unequal variances in the sample populations of parents and offspring, as we observed in our sample (data not shown). No corrections were applied to the estimates to compensate for unequal variances, because the introduced errors should result in an underestimation of heritability, and the values presented here might therefore be taken as conservative estimates for the heritable values. These results need a further elaboration by means of an analysis that takes into account the genotype of subjects in order to test the influence of Y-linked or X-linked genes, paternal or maternal effects and genomic imprinting.

In future studies, we plan to refine heritability estimates for traits sensitive to major environmental factors, extending the analysis by taking into account recorded information about smoking, nutrition and alcohol consumption.

In conclusion, our findings of high heritability suggest that genetic effects substantially contribute to phenotypic variation in this population. Thus, our results support the idea that isolated populations provide excellent research conditions for gene-mapping projects, due to the reduced level of genetic heterogeneity as well as more uniform environmental background. We believe that the obtained information is potentially useful for continuing genetic research in these populations.

We thank the Ogliastra population and all the individuals who participated in this study. We are very grateful to the municipal administrators for their collaboration to the project and for economic and logistic support. This work was supported by grants from the Italian Ministry of Education, University and Research (MIUR) no. 5571/DSPAR/2002 and (FIRB) D. M. no. 718/Ric/2005.

## References

- Almasy, L. & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *American Journal of Human Genetics* **62**, 198–1211.
- Angius, A., Melis, P. M., Morelli, L., Petretto, E., Casu, G., Maestrale, G. B., Fraumene, C., Bebbere, D., Forabosco, P. & Pirastu, M. (2001). Archival demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. *Human Genetics* **109**, 198–209.
- Beekman, M., Heijmans, B. T., Martin, N. G., Pedersen, N. L., Whitfield, J. B., DeFaire, U., van Baal, G. C., Snieder, H., Vogler, G. P., Slagboom, P. E. & Boomsma, D. I. (2002). Genetic architecture of human aging. *Human Genetics* **111**, 105–115.

- D. I. (2002). Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Research* **5**, 87–97.
- Cappello, N., Rendine, S., Griffo, R., Mameli, G. E., Succa, V., Vona, G. & Piazza, A. (1996). Genetic analysis of Sardinia: I. data on 12 polymorphisms in 21 linguistic domains. *Annals of Human Genetics* **60**, 125–141.
- Evans, D. M., Frazer, I. H. & Martin, N. G. (1999). Genetic and environmental causes of variation in basal levels of blood cells. *Twin Research* **2**, 250–257.
- Falchi, A., Giovannoni, L., Calò, C. M., Piras, I. S., Moral, P., Paoli, G., Vona, G. & Varesi, L. (2006). Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. *Journal of Human Genetics* **51**, 9–14.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow, Essex, UK: Addison Wesley Longman.
- Fraumene, C., Belle, E. M., Castri, L., Sanna, S., Mancosu, G., Cosso, M., Marras, F., Barbujani, G., Pirastu, M. & Angius, A. (2006). High resolution analysis and phylogenetic network construction using complete mtDNA sequences in sardinian genetic isolates. *Molecular Biology and Evolution* **23**, 2101–2111.
- Fraumene, C., Petretto, E., Angius, A. & Pirastu, M. (2003). Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis. *Human Genetics* **114**, 1–10.
- Ghiani, M. E. & Vona, G. (2002). Y-chromosome-specific microsatellite variation in a population sample from Sardinia (Italy). *Collegium Antropologicum* **26**, 387–401.
- Hartley, H. O. & Rao, J. N. K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika* **54**, 93–108.
- Lin, H. F., Boden-Albala, B., Juo, S. H., Park, N., Rundek, T. & Sacco, R. L. (2005). Heritabilities of the metabolic syndrome and its components in the Northern Manhattan Family Study. *Diabetologia* **48**, 2006–2012.
- Malhotra, A. & Wolford, J. K. (2005). Analysis of quantitative lipid traits in the genetics of NIDDM (GENNID) study. *Diabetes* **54**, 3007–14.
- Mancosu, G., Ledda, G. & Melis, P. M. (2003). PedNavigator: a pedigree drawing servlet for large and inbred populations. *Bioinformatics* **19**, 669–670.
- Marroni, F., Grazio, D., Pattaro, C., Devoto, M. & Pramstaller, P. (2008). Estimates of genetic and environmental contribution to 43 quantitative traits support sharing of a homogeneous environment in an isolated population from South Tyrol. Italy. *Human Heredity* **65**, 175–182.
- Martin, L. J., Cole, S. A., Hixson, J. E., Mahaney, M. C., Czerwinski, S. A., Almasy, L., Blangero, J. & Comuzzie, G. (2002). Genotype by smoking interaction for leptin levels in the San Antonio Family Heart Study. *Genetic Epidemiology* **22**, 105–115.
- Pilia, G., Chen, W. M., Scuteri, A., Orrù, M., Albai, G., Dei, M., Lai, S., Usala, G., Lai, M., Loi, P., Mameli, C., Vacca, L., Deiana, M., Olla, N., Masala, M., Cao, A., Najjar, S. S., Terracciano, A., Nedorezov, T., Sharov, A., Zonderman, A. B., Abecasis, G. R., Costa, P., Lakatta, E. & Schlessinger, D. (2006). Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genetics* **2**, e132.
- Pistis, G., Piras, I., Pirastu, N., Persico, I., Sassu, A., Picciau, A., Prodi, D., Fraumene, C., Mocci, E., Manias, M. T., Atzeni, R., Cosso, M., Pirastu, M. & Angius, A. (2009). High differentiation among eight villages in a secluded area of Sardinia revealed by genome-wide high density SNPs analysis. *PLoS One* **4**, e4654.
- Weiss, L. A., Pan, L., Abney, M. & Ober, C. (2006). The sex-specific genetic architecture of quantitative traits in humans. *Nature Genetics* **38**, 218–222.