


## RESEARCH ARTICLE

# Y-chromosome variability and genetic history of Commons from Northern Italy

Stefania Sarno<sup>1</sup> | Rajiv Boscolo Agostini<sup>2</sup> | Sara De Fanti<sup>1,3</sup> | Gianmarco Ferri<sup>4</sup> |  
Silvia Ghirotto<sup>2</sup> | Giorgia Modenini<sup>1</sup> | Davide Pettener<sup>1</sup> | Alessio Boattini<sup>1</sup> 

<sup>1</sup>Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy

<sup>2</sup>Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy

<sup>3</sup>Interdepartmental Centre Alma Mater Research Institute on Global Challenges and Climate Change, University of Bologna, Bologna, Italy

<sup>4</sup>Department of Diagnostic and Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena, Italy

## Correspondence

Alessio Boattini, Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.  
Email: alessio.boattini2@unibo.it

## Abstract

**Objectives:** Genetic drift and admixture are driving forces in human evolution, but their concerted impact to population evolution in historical times and at a micro-geographic scale is poorly assessed. In this study we test a demographic model encompassing both admixture and drift to the case of social-cultural isolates such as the so-called “Commons.”

**Materials and methods:** Commons are peculiar institutions of medieval origins whose key feature is the tight relationship between population and territory, mediated by the collective property of shared resources. Here, we analyze the Y-chromosomal genetic structure of four Commons (for a total of 366 samples) from the Central and Eastern Padana plain in Northern Italy.

**Results:** Our results reveal that all these groups exhibit patterns of significant diversity reduction, peripheral/outlier position within the Italian/European genetic space and high frequency of Common-specific haplogroups. By explicitly testing different drift-admixture models, we show that a drift-only model is more probable for Central Padana Commons, while additional admixture (~20%) from external population around the same time of their foundation cannot be excluded for the Eastern ones.

**Discussion:** Building on these results, we suggest central Middle Ages as the most probable age of foundation for three of the considered Commons, the remaining one pointing to late antiquity. We conclude that an admixture-drift model is particularly useful for interpreting the genetic structure and recent demographic history of small-scale populations in which social-cultural features play a significant role.

## KEYWORDS

ABC modeling, demographic history, social-cultural isolates, Y-chromosome

## 1 | INTRODUCTION

Among the evolutionary forces shaping the genetic variability of modern human populations, drift and admixture are of the uttermost importance. In general, both of them can act at the same time on a

Stefania Sarno and Rajiv Boscolo Agostini contributed equally to the study.  
Davide Pettener and Alessio Boattini contributed equally to the study.

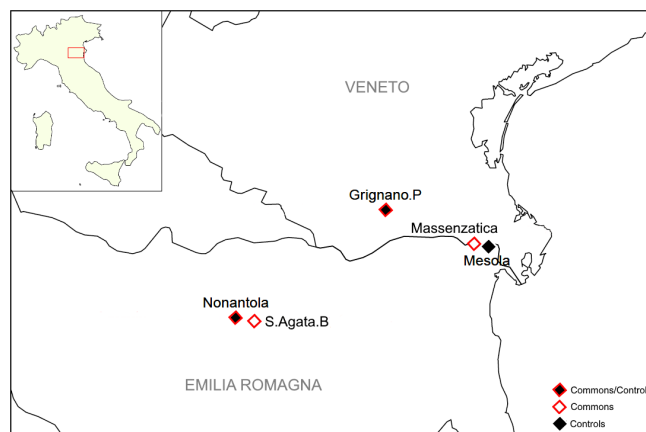
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *American Journal of Physical Anthropology* published by Wiley Periodicals LLC.

given population, albeit with shifts from one to another depending on the historic/demographic vicissitudes experienced by different human groups. In particular, genetic drift is usually associated with isolated populations, that is groups with a moderate demographic size and low exchange with other populations. However, migration and admixture can play a significant role even in isolates, thus contributing to the peculiar genetic structure that these populations frequently exhibit. Accordingly, in this study with ‘drift’ we mean the effects of regional/local isolation, while ‘admixture’ more precisely refers to long-range migration events. Well-known examples of this combination are ethnic-linguistic minorities such as the thoroughly investigated Arbereshe and Greek-speaking groups from Southern Italy (Boattini et al., 2011; Destro Bisol et al., 2008; Sarno et al., 2016; Sarno et al., 2017; Sarno et al., 2021; Sineo et al., 2014). The genomic history of these populations may be summarized according to the following scheme: (a) a founding migration event; (b) admixture—to a lesser or higher extent—with local groups; (c) isolation and drift, usually caused by a combination of cultural and geographical factors. Paradoxically, isolation and drift contributed to the conservation, in these groups, of clear genomic traces of their migration history. Furthermore, it has been shown that these populations may function as privileged observatories for detecting and reconstructing specific demographic events, whose genetic traces were diluted or disappeared in more general contexts (Boattini et al., 2015; Gokcumen et al., 2011; Sarno et al., 2016).

Ethnic-linguistic minorities—where the exogenous origin of (at least part of) the group is usually well-documented—are not the only cases to which an admixture-drift model could be applied at a micro-geographic scale. In this study, we investigate the genetic variability of a set of populations which share some aspects with isolates but whose origin is less clear or not known at all. For reasons of simplicity, we will refer to these populations with the term of ‘Commons,’ which actually designs a peculiar way of sharing and devolving collective resources. According to the Digital Library of the Commons (<http://dlc.dlib.indiana.edu/dlc/>), these are juridical institutions based on ‘shared resources in which each stakeholder has an equal interest.’ Italy hosts a wide archipelago of Commons, each of them with its own features and history, which makes them a particularly interesting set of populations to explore. Our first research in that vein regarded the *Partecipanza* of S. Giovanni in Persiceto, a Common located in the Padana Plain whose members form the patrilineal descent of a group of well-identified founder families. Indeed, that study suggested that: (a) an admixture event, probably involving a group of Germanic origins, was at the origin of the Common and (b) subsequent isolation helped to preserve trace of such event in the Y-chromosomal structure of the *Partecipanza* (Boattini et al., 2015).

Here we focus on four more Commons from the Padana Plain in Northern Italy (Figure 1). Two of them (Nonantola and S. Agata Bolognese) are located in the Central Padana Plain (from now on Central Commons), in the same area of the previously mentioned S. Giovanni in Persiceto, with which they also share the denomination of *Partecipanza*. The remaining two (Grignano Polesine and Massenzatica) are at the Easternmost fringes of the Padana Plain, not far from the Po delta (from now on Eastern Commons).



**FIGURE 1** Geographic map showing the location of the analyzed Commons and Controls

All these Commons share a number of features, most notably all of them evolved from typical medieval contracts called emphyteutic leases, according to which the holder had the perpetual right to the enjoyment of a given property in exchange of some conditions, most notably the increase of the value of the land (*ad meliorandum* clause), the payment of a small rent and the obligation of stable residence. In our case, emphyteutic concessions were established between important abbeyes (Nonantola for Central Commons and Pomposa for Eastern Commons) with vast but deserted properties and local communities as a whole. In exchange of that, the locals contributed significantly to land reclamation and peopling of the territory. At a later time most of these Commons ‘closed’ their ranks and reserved the right to share the collective lands to a restricted group of ‘founder’ families (Ariotti et al., 1990; Costato, 1968; Cori, 2011; Fregni et al., 1992). Indeed, three of the four Commons considered here (namely Nonantola, S. Agata B., and Grignano P.) still limit the access to the shared resources to the legitimate, paternal descent of a given set of families/surnames. In addition, and in all the cases, stakeholders must maintain their place of residence within the legal boundaries of the Common. The combined effect of these rules is a long-term relationship between population and land, modeled upon the shared property of the Common.

With this study, and in consideration of the patrilineal nature of most of these Commons, we specifically explore their paternal (Y-chromosome) genetic heritage with the aim of detecting effects of the peculiar social-cultural features related to their presence in the corresponding populations. In particular, we will (a) evaluate if a drift-only model can fully explain their genetic structure or if a model encompassing both admixture and drift provides a better fit and (b) reconstruct their genetic histories and their implication for the wider scenario of Northern Italy.

## 2 | MATERIALS AND METHODS

### 2.1 | The populations

Among the considered Commons (Table 1, Figure 1), all Central Commons are *Partecipanze*. Currently, six *Partecipanze* are still present and

**TABLE 1** Sample information and details about the considered Commons

Location	Type	Criteria	Area (ha)	N. Stakeholders	N. Samples	
					Common	Control
Nonantola	Partecipanza	Pat, Res	765	2814 (2009)	54	55
S. Agata Bolognese	Partecipanza	Pat, Res	553	237 (2011)	49	
Grignano Polesine	Comune	Pat, Res	130	327 (2012)	66	49
Massenzatica	Formerly Comune, now Consorzio	Res	353	640 (2020)	51	42

Note: Pat: legitimate patrilineal descent of a given set of founder families/surnames. Res: residence within the legal boundaries of the Common.

active in the study area: Nonantola, S. Agata Bolognese, San Giovanni in Persiceto, Cento, Pieve di Cento, and Villa Fontana. Among them, the here studied Nonantola and S. Agata Bolognese occupy the westernmost position. Their assets extend for 765 and 553 ha, respectively, while the number of stakeholders is 2814 (2009) and 237 (2011) (Ariotti et al., 1990; Fregni, 1992; <http://www.partecipanzanonantola.it>; <https://www.partecipanza.org/>). The history of these Commons is quite similar, though the case of Nonantola is best documented. Traditionally, the origins of the Partecipanza of Nonantola are traced back to 1058, when emphyteutic grants were stipulated between the local abbey and the community, but only in 1584 the Common closed its ranks to external families (“*chiusura dei ruoli*”). This was only the beginning of a bitter fight between two factions, one of them formed by peasants (“*bocca viva*”), and the other by rich landowners, many of them foreigners (“*bocca morta*”). Such issue was settled only in 1915, when the descendants of the peasants liquidated the landowners group. In the meantime, in 1856 the Partecipanza separated its administration from that of the municipality and in 1961 finally severed its 1000-year old ties with the Abbey of Nonantola (Venturoli, 2004). S. Agata B. had a less complicated history, being able of redeeming its assets from the Abbey already in 1577 and separating from the municipality in 1797 (Ariotti et al., 1990; Fregni, 1992). Both these Commons, as all other *Partecipanze*, are characterized by a peculiar procedure called *cavazione*, which is a periodic re-shuffling of shared lands among heads of the participating households. Currently in Nonantola this happens every 18 years and in S. Agata Bolognese every 9 years (Ariotti et al., 1990).

The Eastern Commons here considered are Grignano Polesine and Massenzatica. The Common of Grignano P. is currently named *Antichi Beni Originari* and it is referred to by locals as *Comune*. Its shared assets are extended for 130 ha and the number of stakeholders was 327 in 2012 (<http://www.antichibenioriginari-grignano.it/>). The organization and functioning of the *Comune* of Grignano P. are very similar to *Partecipanze*, including a periodic re-shuffling of the shared lands that in this case is performed every 5 years. The *Comune* originated from an emphyteutic grant from the Abbey of Pomposa probably during the central middle ages. This Common closed its ranks to external families already at the end of 15th century (1494) and redeemed its assets from the Abbey in 1968 (Costato, 1968). As for Massenzatica, the current name of the Common is “*Consorzio Uomini di Massenzatica*.” The origin of this Common is usually associated to an 1182 emphyteutic grant from the abovementioned Abbey of Pomposa. Differently from the *Partecipanze* (Nonantola, S. Agata B.) and the *Comune* (Grignano P.), Massenzatica never

established a closure of the ranks to foreigners/immigrants neither procedures of re-shuffling among the stakeholders. The right to benefit of the shared goods is here allowed to all male heads of the household permanently residing in the villages of Massenzatica, Monticelli and Italba. The extension of these goods is 353 ha and the number of stakeholders is presently 640. The Common of Massenzatica reached its present form in 1894 with the creation of the “*Consorzio*” and the resolution of all the easements that, from the times of Pomposa, still weighed on the land (Cori, 2011; <http://www.uominidimassenzatica.it/>).

From the social-cultural point of view, there were no significant differences between members of Commons and their neighbors: they spoke the same dialect, they shared the same religion and cultural identity. Furthermore, up to 19th century usually Commons were not separated from their municipalities, meaning that local civic administration coincided with that of the Common. Of course, this fact could entail some advantages for Common members, who in addition could count on the availability of unalienable land portions in an eminently agricultural society and on the collaboration/mutual assistance among themselves (Fregni, 1992).

As these advantages disappeared or became less important since the economy of the region shifted towards industry and tertiary, members of Commons however conserved the greatest regard for their ancient institutions and their traditions, which in turn had and still have a great impact in shaping the territory on which the community is settled. Starting from the late 18th century and especially after the Napoleonic period, Commons experienced long and bitter fights with public authorities, which manifested increasing hostility toward the peculiar way of possessing that these institutions embodied. As a result, some Commons disappeared during this period turning into “normal” private properties. It can be said that the still-existing institutions are only the remnants of a wider phenomenon (Alfani & Rao, 2011; Mantovani, 2017).

Finally, Commons were (and are) not “closed” communities to the “outside world” also from the demographic point of view, as suggested by the relatively low number of endogamic marriages (see also results below), save the partial exception of Nonantola in which an increase of endogamy rates was indeed measured after the “*chiusura dei ruoli*” in 1584 (Alfani, 2015).

## 2.2 | DNA samples

For each location, we sampled both individuals from the Common and individuals residing in the same place, hence sharing the same

environmental and cultural features, but not belonging to the Common itself. We refer to these last samples as “Controls.” For the case of Massenzatica, Controls were sampled in the locality of Mesola, which is the administrative center of the municipality that includes the Common. The Partecipanze of Nonantola and S. Agata Bolognese, given their geographical proximity, share the same Control group, which was sampled in Nonantola. In total, we collected 366 samples, 220 of them belonging to Commons and 146 to Controls (Table 1).

All these samples were collected according to the standard “grandparents” criterion (i.e., at least three generations of ancestry in the area of the considered Common) and by excluding related individuals. As for the Control samples, all the selected individuals share surnames that were identified as autochthonous for the considered area by Boattini et al. (2012). Data from 63 individuals have been previously published in Boattini et al. (2019). The collection of biological samples was performed during various sessions from 2015 to 2017. For all subjects, a written informed consent was obtained and the Bioethic Committee of the University of Bologna (Italy) approved all procedures. The confidentiality of personal information for each participant to the study was assured and the research was performed in accordance with relevant guidelines and regulations for studies involving human subjects stated by the WMA Declaration of Helsinki.

Whole genome DNA was extracted from buccal swabs by using a salting out protocol modified from Miller (Miller et al., 1988) and quantified with the Qubit® dsDNA BR Assay Kit (Life Technologies).

## 2.3 | Y-chromosome genotyping

All samples were amplified for the 23 Y-STRs loci included in the PowerPlex® Y23 System (Promega) following manufacturer recommended protocols. PCR products were sized on an ABI PRISM 310 Genetic Analyzer and alleles were called with GeneMapper ID software (Thermo Fisher Scientific) according to the manufacturer's instructions.

Next, all individuals were additionally genotyped for 42 Y-SNP loci using multiplex SNaPshot mini-sequencing assays (Thermo Fisher Scientific), as described in Sarno et al. (2014). The SNP genotyping was carried out by means of PCR Multiplex amplification, followed by Minisequencing reaction based on dideoxy Single Base Extension (SBE), which was performed with the SNaPshot multiplex kit (Applied Biosystems). SBE products were finally analyzed through capillary electrophoresis on an ABI Prism 310 Genetic Analyser.

Y-STRs data for comparison populations were extracted from the literature for both Italy (Boattini et al., 2013) and Europe (Purps et al., 2014), respectively. In addition, comparison data for within-haplogroup comparisons from 16 Euro-Mediterranean populations (329 samples) were also considered and retrieved from Hallast et al. (2015).

## 2.4 | Pedigree reconstruction

The *Partecipanze* (Nonantola and S. Agata B.) and the *Comune* (Grignano P.) kept records of all the households participating to the sharing of leased assets. These records were regularly updated in

occasion of the abovementioned periodic re-shuffling of the lands among the householders. Such detailed historic-demographic information allowed the reconstruction of paternal pedigrees up to the late 16th century for Grignano P. and up to 19th century for the *Partecipanza* of S. Agata B. Unfortunately, records from the *Partecipanza* of Nonantola could not be accessed after the earthquake of 2012. When more than one individual was found to share a recent paternal ancestor, they were grouped into a single pedigree.

## 2.5 | Statistical methods

### 2.5.1 | Generation time

We calculated average generation times based on reconstructed paternal pedigrees, using those individuals for which birth information was available. Rates were calculated by dividing the total number of years for the number of generations encompassed in all the considered pedigrees (after excluding the most remote ancestors, due to uncertainty of their birth date). Confidence intervals (95%) were obtained bootstrapping along different branches of pedigrees (1000 replications).

### 2.5.2 | Social endogamy

By inspecting all marriages included in reconstructed pedigrees (S. Agata B., Grignano P.), we considered as endogamic those marriages in which both partners bear surnames of the Common. Marriages for which the surname of the bride was not available were excluded from calculations. Final rates were calculated by dividing the number of socially endogamic marriages for the total number of marriages. Confidence intervals were computed based on a binomial distribution (where the size parameter is equal to the total number of marriages and the probability parameter is given by the observed endogamy rate) using the R function *qbinom* (R Core Team, 2017).

### 2.5.3 | Diversity indexes

Standard within-population diversity parameters (Gene Diversity, Mean Number of Pairwise Differences, Nucleotide Diversity) for Y-chromosome haplogroups and STR haplotypes were estimated with Arlequin software 3.5.1.2 (Excoffier et al., 2007). Comparisons of single haplogroup (hg) frequencies between Commons and Controls were performed with Fisher tests and p-values were corrected using the Bonferroni criterion. The overall differentiation in haplogroup composition between Common-Control pairs was assessed through AMOVA analysis based on pairwise *F<sub>st</sub>* statistics as implemented in the above mentioned Arlequin software.

### 2.5.4 | Multivariate analyses between populations

In order to check the position of our populations within the Italian and European Y-chromosomal genetic landscape, we performed a

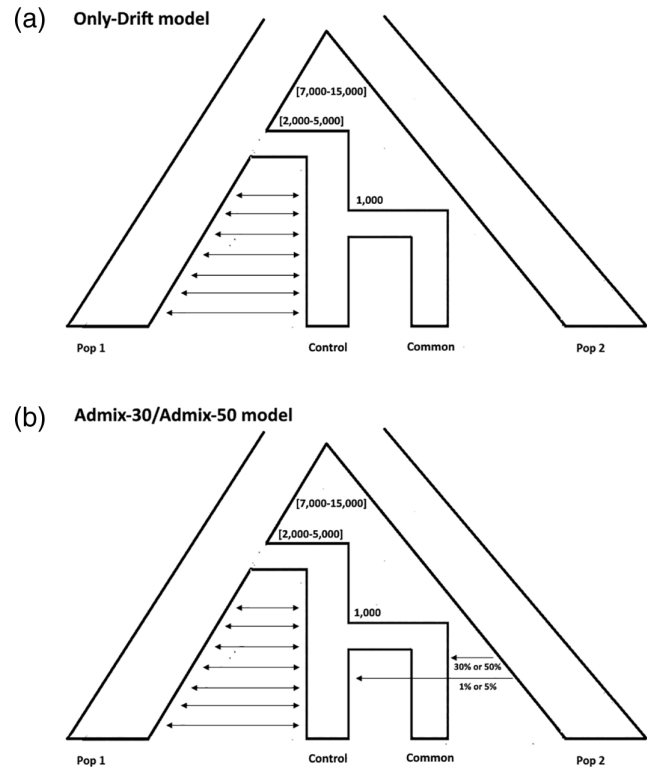
nonmetric multidimensional scaling (MDS). Because different studies used different levels of hg resolution, the analyses were based on the 23 Y-STR haplotypes that were available for all the considered reference populations. Calculations were performed using the Rst genetic distances computed by the Arlequin software (Excoffier et al., 2007) and the function *isoMDS* implemented in the R software MASS package (R Core Team, 2017; Venables & Ripley, 2002). The first and the second dimension were represented in a scatterplot, along with the corresponding stress value.

## 2.5.5 | Within-haplogroup comparisons

In order to explore the genetic variability within haplogroups exhibiting significant frequency differences between Commons and Controls, we used Discriminant Analysis of Principal Components (DAPC, Jombart et al., 2010) based on Y-STR data as in Boattini et al. (2013, 2015). This analysis is aimed to (a) identify well-resolved groups of haplotypes within haplogroups; (b) highlight possible affinities/similarities with reference haplotypes from Italy and Europe; and (c) constitute a starting point for time estimates. All the analyses were performed within the R software package *adegenet* (Jombart, 2008).

## 2.5.6 | Time estimates

Time estimates focused on the haplotype clusters identified by DAPC within the most frequent haplogroups and were limited to clusters including at least nine individuals and separate estimates were also performed for clusters with at least nine individuals within a specific Common. Y-STR mutation rates adopted in the procedure were taken from Ballantyne et al. (2010). Because population events involving Commons are relatively recent, the biasing effect of Y-STRs saturation through time is negligible (Boattini et al., 2016; Boattini et al., 2019) and all Y-STRs (minus *DYS385a/b*) were therefore used for calculations. In addition, since estimates may be very sensitive to the presence of outliers, we adopted the outlier detection and exclusion procedure described in Boattini et al. (2013). Time estimates were calculated using two different approaches: (a) the *SD* estimator (Sengupta et al., 2006); (b) the Bayesian method implemented in the *Batwing* software (Wilson et al., 2003). As for the latter, we adopted a standard coalescent model with constant (effective) population size (*N*). Since we are considering not the entire population, but rather clusters of individuals within single lineages (as in Balaresque et al., 2010 and more recently Platt et al., 2017; Huang et al., 2018), we assigned a uniform distribution on the interval (10,1000) to the *N* prior. Accordingly, we set STR-specific priors of mutation rates using gamma distributions in the form  $\text{gamma}(nmut,ngen)$ , where *nmut* is the number of observed mutations and *ngen* is the total number of meioses (as in Ballantyne et al., 2010). The number of times parameters were updated between samples (*Nbetsamp*) was 10, and the number of times trees were changed before updating parameters (*treebetN*) was 20. The number of samples between writing the outfile (*pigap*)



**FIGURE 2** Demographic histories of the considered models. “Pop 1” represents a general Italian population; “Pop 2” represents a non-Italian group, source of potential external admixture/introgression

was 1,500,000. A total of 3.5 million of MCMC runs with different random seeds were run for each haplotype cluster, and the first 1 million iterations were discarded as burn-in. Time estimates were calculated from the resulting outfile using the product of the posteriors estimated for the population size *N* and the total height of the tree *T*. For both methods, a generation time of 33.72 years (as estimated by our analysis, see Results) was specifically used for converting all time estimates in years. All statistics were calculated with the R software.

## 2.5.7 | Model comparison by approximate Bayesian computation

In order to investigate the recent demographic history of the considered groups, we compared three different models for each Common/Control pair. All models follow the same demographic history (Figure 2) whose events are: (a) a divergence between 7000 and 15,000 years ago representing a putative separation between Italian and non-Italian (Pop 2) groups; (c) a second divergence between 2000 and 5000 years ago representing the segregation of local populations (i.e., the ancestors of Common/Control pairs) from the general Italian population (Pop 1); (c) a third divergence event 1000 years ago representing the separation between Common and Control. In all models the local populations (from event 1 to event 2) and the Controls (after



event 2) keep exchanging migrants with Pop 1 (general Italian population). The three tested models instead differ for the degree of admixture from non-Italian Pop 2 (external source of genetic variation) to the Common. According to the “Only-Drift model” (Figure 2a), there are no exchanges between these two demes, and the Common remains completely isolated until present times. Under the “Admix-30 model,” 30% of the Common Y-chromosomal composition and 1% of the corresponding Control composition originated from an external source of variation (Pop 2), whereas under the “Admix-50 model” these percentages are 50% and 5%, respectively (Figure 2b). For a detailed description of all models' parameters see Table S1.

We generated 5000 simulated data sets (21 Y-chromosome microsatellites) under each model with *fastsimcoal* (Excoffier et al., 2013) within the ABCToolbox suite (Wegmann et al., 2009). We sampled 40 individuals from the Common deme and 40 individuals from the corresponding Control deme at each iteration, while Pop 1 and Pop 2 were set as “ghost” populations. The considered summary statistics are: the number of alleles (K), Heterozygosity (H), Garza-Williamson statistic (GW), allelic range (R), Fixation index (Fst), Pairwise differences (Pi), and mean delta-mu squared (DMUSQR). We compared models using Random Forest Approximate Bayesian Computation (ABC-RF: Pudlo et al., 2016). Under ABC-RF, the Bayesian model selection is rephrased as a classification problem, in which the classifier is constructed from simulations via a machine learning RF algorithm. Once the classifier is constructed and applied to the observed data, the posterior probability of the resulting model can be approximated through another RF that regresses the selection error over the statistics used to summarize the data. To perform the model selection procedure, we used the function *abcrf* from the R package *abcrf* and employed a forest of 500 trees, as suggested by Pudlo et al. (2016). Before calculating the posterior probabilities of the most supported model, we computed the confusion matrix and evaluated the out-of-bag classification error.

In order to better define the patterns emerged from models' comparison, we performed the estimation of the admixture proportion in the considered Commons (introgression from Pop 2 to the Common). To do this, we designed a demographic model in which the admixture parameter is free to vary at each iteration from 0% to 50%. We performed 500,000 simulations and then estimated the posterior distribution through the *ABCestimator* tool of the ABCToolbox package (Wegmann et al., 2009).

### 3 | RESULTS

#### 3.1 | Reconstructed pedigrees, generation time, social endogamy

In total, we reconstructed 61 paternal pedigrees (24 for S. Agata B. and 37 for Grignano P.; Figure S1, Table S2), encompassing 960 generations and 124 individuals. Of them, 28 refer to a single individual, 18 comprise two individuals and 14 more than two individuals (up to 6). Of these 124 individuals, 115 were actually genotyped

while nine were excluded given their recent relatedness with other individuals, that is, a number of generations separating them from other individuals of the same paternal pedigree lower than 7. This was done for ethical reasons as in Claerhout et al. (2019) and Boattini et al. (2019).

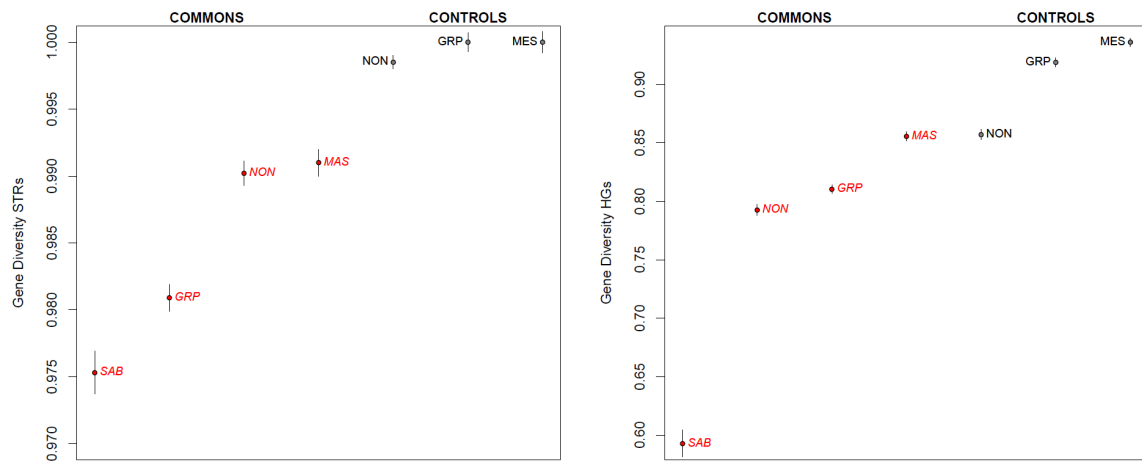
Based on all 61 pedigrees, we estimated the average generation time and obtained 33.72 years per generation (95% CI: 33.22, 34.28). We then calculated values for each population, obtaining 34.18 (95% CI: 33.62, 34.74) for Grignano P. and 32.61 (95% CI: 31.64083, 33.58142) for S. Agata B. Hence, Grignano P. seems to exhibit slightly higher average generation times. Since Grignano P. pedigrees are averagely deeper than S. Agata B. ones, we also sub-sampled pedigrees from the former in order to match the temporal depth of the latter. Results show that sub-sampled pedigrees from Grignano P. yielded 33.91 years per generation (95% CI: 33.16, 34.77), which largely overlaps with the above estimate. This fact suggests that some, even if slight, population-dependent variability indeed exists.

Social endogamy was evaluated for one of the central communities (S. Agata B.) and for one of the eastern communities (Grignano P.). Nonantola was excluded due to the current unavailability of historical-demographical data and Massenzatica for not having a set of Common-specific (“founder”) surnames. Grignano P. genealogies comprise 630 marriages, 569 of them including the surname of the bride. Social endogamy in these 569 marriages was 59.22% (95% CI: 55.01, 63.09). As for S. Agata B., our data set (pedigrees) comprises 335 marriages, 269 of them including the surname of the bride. Social endogamy here is much lower, being only 28.25% (95% CI: 22.68, 33.46). Therefore, these results show that, in these communities, patrilineal isolation was balanced by exogamous marriages.

#### 3.2 | Diversity indexes

We calculated classic indexes of genetic diversity for all the considered populations (Commons and Controls) using both Y-STR haplotypes and haplogroup frequencies (Table S3, Figure 3). The obtained results (particularly those based on Y-STRs) clearly show that all the considered Commons exhibit lower diversity values than the Controls. Interestingly this feature is observed also in Massenzatica, despite its higher social “openness” (membership is here conditioned only by residency) compared with the remaining Commons. As for these groups, the lowest diversity values were observed for S. Agata B., which in fact exhibits the lowest population size. Indeed, STR-based Gene Diversity values for the three Commons based on founder families faithfully correspond to the number of stakeholders (Nonantola > Grignano > S. Agata B.).

When explicitly testing for genetic structuring among Commons, we obtain high and significant Fst values with both STR haplotypes and haplogroup frequencies (Fst = 0.042 with *p* value <0.001, Fst = 0.116 with *p* value <0.001, respectively). The same analysis with Control groups instead yields much lower Fst values, in one case (STRs) not even reaching the nominal significance threshold (STRs: Fst = 0.002 with *p* value = 0.330; haplogroups: Fst = 0.020 with



**FIGURE 3** Y-chromosome diversity indexes in Commons and Controls

$p$  value = 0.006). These results confirm the strong differentiation among Commons, while Controls exhibit higher homogeneity between each other.

### 3.3 | Haplogroup distribution

Haplogroup-wise (Table S4), the considered Commons—despite generally harboring haplogroups typical of Western-European populations—show some significant differences from their Controls, both from an overall point of view and for specific lineages. Let us consider first the two Central Commons, that is, Nonantola and S. Agata Bolognese. When performing overall  $F_{st}$  tests (AMOVA) based on haplogroup frequencies, both Commons appear as significantly different from the Control (Nonantola:  $F_{st}$  = 0.04,  $p$  value = 0.004; S. Agata B:  $F_{st}$  = 0.08;  $p$  value <0.001). In particular, these Commons, similarly to their Control (Nonantola), show a typical prevalence of R haplogroups, which make up to ~70% of their repertoire. However, Nonantola exhibits a significantly higher presence of R1b-L2, when compared with the Control (33.33% vs. 7.27%,  $p$  value = 0.014), while S. Agata is strongly characterized by lineage R1b-U152\* (63.27% vs. 30.91%,  $p$  value = 0.0026).

Moving to Eastern Commons, again overall  $F_{st}$  tests revealed significant differences between them and their controls (Grignano P:  $F_{st}$  = 0.04,  $p$  value = 0.004; Massenzatica/Mesola:  $F_{st}$  = 0.03,  $p$  value = 0.005). As for specific lineages, Massenzatica, compared with its Control (Mesola), shows an overall higher frequency of R lineages (64% vs. 47.62%). Such difference is mainly due to a higher frequency of paragroup R1b-L51\*, however not reaching the statistical significance after Bonferroni correction (18% vs. 2.38%,  $p$  value = 0.386).

The case of the Common of Grignano P. is particularly interesting, since its main feature is the very high frequency (22.73%) of a non-R haplogroup, T-M70, which is absent in its control ( $p$  value = 0.002) and in general quite sparse in Italian populations (1.58%; Boattini et al., 2013).

## 3.4 | Multivariate analyses

### 3.4.1 | MDS

In order to check the position of the considered groups within the Italian and the European Y-chromosomal genetic landscape, we performed STR-based MDS using several reference populations (Figure 4).

As for MDS results at an Italian level, we observe that Control groups are located towards the center of the plot, with Nonantola falling between North-Western Italian populations and the two Eastern-most Controls (Mesola and Grignano P.) mapped quite near to each other and in an intermediate position between North-Western and South-Eastern Italian population groups. Interestingly, all Commons occupy peripheral positions in the plot, which can be explained according to two hypotheses: (a) as a result of prolonged genetic drift (in agreement with diversity indexes results), and (b) as a result of admixture/introgression involving populations with a different genetic background.

When considering a European-wide genetic space, the observed patterns are highly consistent with the previous case. In addition, the Common of Nonantola is placed along with Iberian populations, while the Common of S. Agata B. still occupies an outlier position.

### 3.4.2 | DAPC and time estimates

Within-haplogroup analyses were performed for the most frequent haplogroups in the considered Commons, that is, R-U152\*, R-L51, R-L2, T-M70 and G2a-U8 (Table S5, Figure S2). DAPC results (Table 2, Table S5) show that these haplogroups revealed a number of well-defined clusters variable from 3 to 7. In addition, all of them exhibit at least one Common-specific cluster, that is, a cluster in which the majority (>50%) of the haplotypes is associated to a given Common. More precisely, S. Agata B. presents two associated clusters, both of them within R-U152\* (namely 4 and 5); Nonantola also presents two





**TABLE 2** Time estimates and standard errors (in years before present) based on DAPC clusters exceeding nine individuals

Haplogroup	Cluster	Ntot	Common	Ncommon	Date_All		Date_Common	
					SD estimator	Batwing	SD estimator	Batwing
R-U152*	1	12	GP	12	-	-	1206.62 +/- 257.25	1280.524 +/- 141.75
R-U152*	2	10	-	-	2541.43 +/- 541.83	3999.52 +/- 415.80	-	-
R-U152*	3	19	NON	11	4714.93 +/- 1005.23	6996.91 +/- 575.771	1016.09 +/- 216.63	1535.75 +/- 206.22
R-U152*	4	11	SAB	9	2674.84 +/- 570.28	2627.54 +/- 324.46	1040.1 +/- 221.75	1138.26 +/- 163.48
R-U152*	5	15	SAB	15	-	-	520.59 +/- 110.99	594.17 +/- 76.30
R-U152*	6	15	-	-	3139.15 +/- 669.27	4384.90 +/- 278.73	-	-
R-U152*	7	11	-	-	3508.49 +/- 748.01	5832.41 +/- 491.53	-	-
R-L51	1	7	-	-	-	-	-	-
R-L51	2	16	MA/ME	10	2863.21 +/- 610.44	4457.97 +/- 377.79	1743.77 +/- 371.77	2276.28 +/- 311.76
R-L51	3	17	-	-	2126.93 +/- 453.46	4814.91 +/- 321.75	-	-
R-L2	1	7	GP	6	-	-	-	-
R-L2	2	9	-	-	7405.99 +/- 1578.96	5023.37 +/- 589.37	-	-
R-L2	3	23	-	-	3638.76 +/- 775.79	5938.69 +/- 290.09	-	-
R-L2	4	10	NON	10	-	-	1135.13 +/- 242.01	996.21 +/- 143.45
R-L2	5	11	GP	10	678.33 +/- 144.62	687.86 +/- 103.95	700.31 +/- 149.31	721.29 +/- 122.16
T-M70	1	5	-	-	-	-	-	-
T-M70	2	16	GP	15	628.68 +/- 134.03	1946.89 +/- 187.71	562.65 +/- 119.96	792.43 +/- 89.86
T-M70	3	3	-	-	-	-	-	-
G2a-U8	1	9	GP	9	-	-	774.65 +/- 165.15	871.71 +/- 136.85
G2a-U8	2	20	-	-	3140.2 +/- 669.49	6764.46 +/- 348.90	-	-
G2a-U8	3	5	GP	5	-	-	-	-

Note: Calculations are performed for whole clusters (Date\_All) and for common-specific individuals (Date\_common).

Abbreviations: Common, name of the Common, if any, to which the cluster is specific (GP, Grignano P; MAS, Massenzatica; NON, Nonantola; SAB, S. Agata B.); Date\_all, time estimate for all cluster individuals; Date\_Common, time estimate for Common-specific individuals only; Ncommon, number of Common-specific individuals; Ntot, total number of individuals.

the considered models proved to be able to reproduce the observed variation, as shown by linear discriminant analysis (LDA; Figure S3).

### 3.6 | Parameters estimation

When performing model selection, we tested demographics accounting for different proportions of admixture (0%, 30%, and 50%) from an external source of genetic variation. This was done in order to

maximize the identifiability of the considered models, which correlates with the degree of differentiation among the tested demographic histories. However, the admixture parameter could actually vary in a wider range than that explored by the above tested models. Accordingly, we estimated admixture rate for all the considered Commons using ABC (Table 4, Figure S4). Consistently with the model selection analysis, posterior probabilities in Nonantola and S. Agata B. appear narrower and shifted toward lower values than those observed for Grignano P. and Massenzatica. Median admixture rates in the Eastern

**TABLE 3** Model comparison and selection for each Common/Control pair

Common	Control	Model	Classification error	Votes	Posterior probability
<i>All models in a single run</i>					
Grignano P.	Grignano P.	Drift only	0.28	154	
		Drift +30% admix	0.62	165	
		<b>Drift + 50% admix</b>	<b>0.39</b>	<b>181</b>	<b>0.49</b>
Massenzatica	Mesola	Drift only	0.28	158	
		<b>Drift + 30% admix</b>	<b>0.62</b>	<b>279</b>	<b>0.60</b>
		Drift +50% admix	0.39	63	
S. Agata B.	Nonantola	<b>Drift only</b>	<b>0.28</b>	<b>308</b>	<b>0.69</b>
		Drift +30% admix	0.62	144	
		Drift +50% admix	0.38	48	
Nonantola	Nonantola	<b>Drift only</b>	<b>0.29</b>	<b>365</b>	<b>0.70</b>
		Drift +30% admix	0.62	124	
		Drift +50% admix	0.39	11	
<i>Drift only vs. Drift + 30% admix</i>					
Grignano P.	Grignano P.	Drift only	0.26	225	
		<b>Drift + 30% admix</b>	<b>0.29</b>	<b>275</b>	<b>0.60</b>
Massenzatica	Mesola	Drift only	0.26	200	
		<b>Drift + 30% admix</b>	<b>0.29</b>	<b>300</b>	<b>0.53</b>
S. Agata B.	Nonantola	<b>Drift only</b>	<b>0.26</b>	<b>350</b>	<b>0.69</b>
		Drift +30% admix	0.29	150	
Nonantola	Nonantola	<b>Drift only</b>	<b>0.26</b>	<b>396</b>	<b>0.78</b>
		Drift +30% admix	0.29	104	
<i>Drift only vs. Drift + 50% admix</i>					
Grignano P.	Grignano P.	Drift only	0.20	233	
		<b>Drift + 50% admix</b>	<b>0.20</b>	<b>267</b>	<b>0.59</b>
Massenzatica	Mesola	<b>Drift only</b>	<b>0.20</b>	<b>389</b>	<b>0.77</b>
		Drift +50% admix	0.20	111	
S. Agata B.	Nonantola	<b>Drift only</b>	<b>0.20</b>	<b>393</b>	<b>0.85</b>
		Drift +50% admix	0.20	107	
Nonantola	Nonantola	<b>Drift only</b>	<b>0.20</b>	<b>485</b>	<b>0.94</b>
		Drift +50% admix	0.20	15	

Note: Bold characters indicate the selected models.

Common/control	Mean	Mode	Median	95% lower HPD	95% upper HPD
Grignano P./Grignano P.	0.56	0.38	0.29	0.05	0.49
Massenzatica/Mesola	0.41	0.12	0.19	$4 \times 10^{-5}$	0.42
S. Agata B./Nonantola	0.27	0.05	0.11	$2 \times 10^{-5}$	0.34
Nonantola/Nonantola	0.21	0.04	0.08	$3 \times 10^{-5}$	0.26

**TABLE 4** ABC estimator results describing admixture estimates between a non-Italian ghost populations (Pop 2; see Figure 2) and each Common, showing mean, mode, median, and 95% highest posterior density (HPD) interval

Commons (Grignano P. and Massenzatica) were around 30% (with mean value of 56%) in Grignano P. and 20% (with mean value of 41%) in Massenzatica. In Central ones (S. Agata B. and Nonantola) they were around 10% with slightly narrower confidence intervals. In other words, a small amount of admixture/introgression in Nonantola and S. Agata B., which showed support for the model with “only drift,” is still compatible with the observed data.

## 4 | DISCUSSION

The main aim of this study is to understand how admixture and drift may contribute to the genetic make-up of peculiar populations such as those associated to the presence of Commons. At the same time, these groups, investigated at a micro-geographic scale after a careful sampling, may reveal aspects of the recent genetic history of the

region in which they are located—aspects that are normally hidden or more difficult to detect in the “general” population. Similar considerations were also expressed in a study about four Anatolian villages by Gokcumen et al. (2011), according to whom “broad, ethnicity-based sampling is inadequate to capture the genetic signatures of recent social and historical dynamics, which have had a profound influence on contemporary genetic and cultural regional diversity”.

Indeed, a previous study about the *Partecipanza* of S. Giovanni in Persiceto (Boattini et al., 2015) suggested that social-cultural features such as the presence and the persistence through the centuries of Commons seem particularly apt to produce, as a biological effect, peculiar micro-geographic genetic structures in which both admixture and drift could have played an important role. Accordingly, we sampled four Commons from the Padana Plain, Northern Italy, each of them characterized by its peculiar set of rules to keep the property of the shared goods in the hands of a restricted and well-defined group of stakeholders (Figure 1, Table 1). Then, we compared their Y-chromosomal structure to that of Control populations, that is, groups that share the same environmental and cultural features of the Common, excepted for the affiliation to the Common itself.

Our results showed that in all cases Commons are characterized by a significant reduction of genetic diversity, compared with their Controls (Figure 3, Table S3). This was somehow anticipated, since the social-cultural rules determining the affiliation to a given Common are likely to have caused a certain degree of isolation between the Common and the neighboring populations, at least from the paternal side. Such rules are particularly stringent for Nonantola, S. Agata B. and Grignano P., where both local residence and legitimate patrilineal descent are required, while in Massenzatica only the former is applied (Table 1). Accordingly, only a moderate diversity reduction was observed in Massenzatica, while such phenomenon is more evident in the three remaining Commons, in addition showing a clear association with population size. Furthermore, as revealed by AMOVA, Control groups show an overall higher similarity between them than Commons do. A significant reduction of Y-chromosomal diversity was not detected in the *Partecipanza* of S. Giovanni in Persiceto, but this fact could be due to the lower number of Y-STRs and Y-SNPs typed in that study (Boattini et al., 2015).

Differences between Commons and their Controls are also apparent when inspecting their haplogroup composition (Table S4). In fact, in all cases we detected an overall significant difference in terms of *F<sub>st</sub>* between the considered pairs of populations. In particular, few haplogroups clearly characterize each of the considered populations. On that respect, a subtle difference between Central (Nonantola, S. Agata B.) and Eastern (Grignano P., Massenzatica) Commons emerged: Central Commons harbor haplogroups/paragroups that are among the most frequent in Northern Italy, i.e. R-U152\* (S. Agata B.) and R-L2 (Nonantola); instead Eastern Commons show high frequencies of more uncommon or rare haplogroups/paragroups in Italy such as R-L51\* and T-M70 (Boattini et al., 2013). All these cases can be explained based on the effects of genetic drift, according to which the frequency of some haplogroups may have increased (or decreased) by random fluctuations. However, it is possible that some of them—

particularly in Eastern Commons—could be the result of an introgression/admixture event around the time in which the Common was formed. The same observations hold also for MDS results (Figure 3), in which the four Commons, compared with reference Italian and European populations, tend to occupy peripheral or even outlier positions in the genetic space: indeed, both drift and admixture could explain such configuration. In addition, the case of the *Partecipanza* of S. Giovanni in Persiceto—where similar patterns were observed, most notably the higher-than-expected frequency of the otherwise rare haplogroup I1-L22—suggested that external admixture could have played a significant role in the genetic history of Commons (Boattini et al., 2015).

However, it is important to underline that isolation in the considered populations is here only examined from the Y-chromosomal (patrilineal) point of view. Indeed, estimates of social endogamy in Grignano P. is ~60% and in S. Agata B. only ~30%, suggesting that marriages did not occur based on affiliation to the Common. In other words, Commons appear as “open” to the rest of the population from the maternal side. This fact implies that, differently from “classic” isolates, such as ethno-linguistic minorities as the Arbereshe of Southern Italy or German-speaking groups from Northern Italy (Anagnostou et al., 2017; Anagnostou et al., 2019; Boattini et al., 2011; Coia et al., 2012; Coia et al., 2013; Colonna et al., 2013; Esko et al., 2013; Sarno et al., 2016), no inbreeding increase is expected in these populations, nor significant genetic differentiation, if autosomal variants or mitochondrial DNA (mtDNA) would have been considered. Indeed, no significant difference was observed in S. Giovanni in Persiceto between the Common and its Control from the mtDNA point of view (Boattini et al., 2015).

We then explicitly tested the admixture-drift model in our populations via ABC. A similar approach was recently used at a larger geographic scale in Kutanan et al. (2019), which however is here aimed at disentangling specific micro-geographic patterns and reconstructing recent demographic histories. In particular, we considered three different models namely: (a) Drift-only, (b) Admix-30 (i.e., 30% admixture followed by drift), and (c) Admix-50 (i.e., 50% admixture followed by drift) (Figure 2). Our simulations suggested that the Drift-only model is favored for Central Commons (Nonantola, S. Agata B.), while for Eastern Commons (Grignano P., Massenzatica) some degree of introgression from an external source of genetic variation is more probable (Table 3, Figure S3). However, admixture cannot be excluded also for Central Commons, albeit with a much lower admixture rate, as suggested by ABC parameter estimation (Figure S4). In fact, after estimating the admixture parameter, we obtained values around 17%–21% with large confidence intervals for Eastern Commons, and 11%–15% with tight confidence intervals for Central ones (Table 4). Interestingly, these results agree with the haplogroup dissection of the considered populations, in which Eastern Commons are characterized by a remarkable presence of infrequent haplogroups in Northern Italy.

As for the observed heterogeneity among the considered Commons, it should be mentioned here that similar results were found in different contexts, suggesting that dynamics of drift/admixture alike

those described here may be at work also in these cases. For instance, Chaix et al. (2007) observed that in Central Asia lifestyle differences were associated to a substantial loss of Y-chromosome diversity in pastoral populations compared with farmer ones. Similarly, Gokcumen et al. (2011) in central Turkey observed micro-geographic patterns of paternal genetic structuring mostly associated to cultural isolation (ancestry, religion). Interestingly, both these examples revealed that such differences/structures were Y-chromosome-specific, being absent when considering mitochondrial DNA or autosomal DNA.

Our second aim was to use this information in order to reconstruct the genetic history of these groups in the wider context of Italian genetic history. Accordingly, we proceeded to a more detailed analysis of the detected haplotypes/haplogroups/paragroups using DAPC (Table 2, Table S5, Figure S2). Our results revealed 11 Common-specific clusters of haplotypes within the most frequent haplogroups. The great majority (9/11) of them were exclusive or almost exclusive of one of the considered Commons, albeit with two exceptions, namely Cluster 3 in paragroup R-U152\*, which is associated to the Nonantola Common but is sporadically observed in other Commons and Controls, and Cluster 2 in hg R-L51\*, which is associated to the Common of Massenzatica, but is sporadically observed in all Controls and also in three central-northern European populations (Bavarian, Danish, English). Interestingly, Massenzatica and Nonantola are also the Commons that showed the lowest amount of diversity reduction, suggesting a lower degree of paternal isolation than the other ones (Figure 3, Table S3).

In general, it seems plausible that Common-specific clusters would mark expansion events within the corresponding populations, therefore estimates of their time depth should work as a lower bound for the time of origin of the Common itself.

A key parameter, when evaluating time estimates based on molecular markers, is the average generation time. This study suggests that, based on documented pedigrees, a generation time of 33.72 years could be adequate at least for Northern Italian populations. Indeed, our estimate overlaps with similar pedigree-based estimations for the Common of S. Giovanni in Persiceto (33.38 years; Boattini et al., 2015) and, even more interestingly, for Emilia-Romagna populations not including Commons (33.57 years; Boattini et al., 2019). These values substantially agree with previous estimates based on demographic cross-cultural comparisons (Fenner, 2005), according to which male generation interval in “developed nations” is 30.8, and in “less developed nations” is 31.8. Such figures, which are slightly lower than ours, are referred to 20th centuries populations, while pedigree-based estimates encompassing larger time intervals vary between 31.9 (Iceland, 1742–2002; Helgason et al., 2003) and 34.5 (French Canada, 1850–1990s; Tremblay & Vézina, 2000). Interestingly, our results hint at a moderate but significant local variability, with Grignano P. showing slightly higher values than S. Agata B.

Despite considering the fact that STR-based estimates should be taken with the greatest caution, our SD and Batwing time estimates, besides being in agreement with each other, seem to be coherent with historical information about the origin of these communities (Ariotti

et al., 1990; Cori, 2011; Costato, 1968; Fregni, 1992; Venturoli, 2004). Indeed, dates (Table 2) range between central Middle Ages and early Modern Age, which is in agreement with an origin of the Commons at least 1000 years ago and following re-expansion events after the segregation of the founder families (in Nonantola, S. Agata B., Grignano P.) around 500 years ago. Again, Massenzatica is an exception pointing towards an earlier origin—likely in Late Antiquity—which seems coherent with local archaeological remains (Cori, 2011). Whole clusters age estimates instead point especially to the interval between 2000 and 5000 years before present, which is in agreement with the more general make-up of the Italian population (Antonio, 2019; Boattini et al., 2013; Cocca, 2020; Fernandes et al., 2020; Marcus et al., 2020; Sarno et al., 2017; Sazzini et al., 2016; Sazzini et al., 2020).

Recent studies based on modern and ancient genomes suggested that the “core” of Italian genetic variation was already in place in the early antiquity. Late antiquity and medieval migrations, as far as we are concerned, seem to have left only minor traces in the Italian genomic background, which are not apparent in “general” populations (Antonio, 2019). For instance, traces of late medieval migrations from Southern Balkans to Southern Italy are detectable in some ethnic-linguistic minorities that still conserve their original language (Arbereshe) but not in the “average” Southern Italian population (Sarno et al., 2016; Sarno et al., 2017). Similarly, the *Partecipanza* of S. Giovanni in Persiceto was hypothesized to spot traces of an early medieval migration from Northern Europe which were lost in other groups (Boattini et al., 2015).

As above discussed, our analyses suggested that Eastern Commons, that is, Grignano P. and Massenzatica, are the most likely cases for admixture-drift, while the Y-chromosomal variability of Central Commons (Nonantola and S. Agata B.) is more easily explained by drift only. In light of these results together with DAPC and dating experiments, the most likely scenario could be the following. The three Commons with the patrilineal descent rule (Nonantola, S. Agata B., and Grignano P.) were founded in the central Middle Ages, while Massenzatica could have been a few centuries older. All communities likely stem from local populations that lived in the area from a long time, however Grignano P. and Massenzatica probably incorporated a ~ 20% contribute from an external population around the same time in which the Common was founded. Unfortunately, it is not possible to be more precise about the origin/identity of these external populations, limiting ourselves to the observation that Grignano P. exhibits an otherwise rare haplogroup, T-M70, which could refer to a Mediterranean background, while R-L51\*, which is typical of Massenzatica, is mostly observed in central Europe (Busby et al., 2012; Harney et al., 2018; Mendez et al., 2011; Myres et al., 2011).

Later, isolation and genetic drift were induced in Nonantola, S. Agata B., and Grignano P. in virtue of their patrilineal separation from the neighboring populations. In Massenzatica, instead, we believe that the peculiar environment in which the community was founded—a lagoon-like environment subjected to frequent floods from the Po river—played a fundamental role in its segregation.

## 5 | CONCLUSION

This study shows how the co-presence of admixture and drift forms a suitable model for explaining the genetic variability of at least two of the four considered Commons, namely Grignano P. and Massenzatica. At the same time, we observed that the peculiar social-cultural features of Commons—based on patrilineal descent and local residence— influence their Y-chromosomal variability in a way reminiscent of ethnic-linguistic minorities, where phenomena such as isolation and/or admixture are frequently observed. The collected results allowed to reconstruct some aspects of the genetic history of the considered communities. For instance, our estimates suggest that the Commons of Nonantola, S. Agata B., and Grignano P. probably originated in the central Middle Ages from a set of mainly but not exclusively local populations, while the case of Massenzatica seems to suggest a more ancient origin. The same admixture-drift model could be proposed as a reference model for the interpretation of the genetic structure of isolated populations in which social-cultural features play a significant role.

### ACKNOWLEDGMENTS

Authors would like to thank all the participants to the study. We are particularly grateful to the personnel of Partecipanza of Nonantola, Partecipanza of S. Agata Bolognese, Antichi Beni Originari of Grignano Polesine and Consorzio Uomini Massenzatica for their invaluable help in designing the sampling campaign and organizing results dissemination. Open Access Funding provided by Università degli Studi di Bologna within the CRUI-CARE Agreement. [Correction added on 18 May 2022, after first online publication: CRUI funding statement has been added.]

### AUTHOR CONTRIBUTIONS

**Stefania Sarno:** Data curation; formal analysis; methodology; supervision; visualization; writing-original draft. **Rajiv Boscolo Agostini:** Formal analysis; methodology; writing-original draft. **Sara De Fanti:** Formal analysis; methodology. **Gianmarco Ferri:** Formal analysis; methodology. **Silvia Ghirotto:** Formal analysis; methodology; writing-original draft. **Giorgia Modenini:** Data curation; investigation. **Davide Pettener:** Conceptualization; resources; supervision. **Alessio Boattini:** Conceptualization; data curation; formal analysis; investigation; methodology; supervision; writing-original draft.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in YHRD at [yhrd.org](http://yhrd.org), reference numbers YA004724 and YA004725.

### ORCID

Alessio Boattini  <https://orcid.org/0000-0002-0758-4454>

### REFERENCES

Alfani, G. (2015). Closing a network: A tale of not-so-common lands (Nonantola, sixteenth to eighteenth centuries). In G. Fertig (Ed.), *Social*

- networks, political institutions, and rural societies* (pp. 153–182). Brepols.
- Alfani, G., & Rao, R. (2011). *La gestione delle risorse collettive. Italia settentrionale, secoli XII-XVIII*. Franco Angeli.
- Anagnostou, P., Dominici, V., Battaglia, C., Lisi, A., Sarno, S., Boattini, A., Calò, C., Francalacci, P., Vona, G., Tofanelli, S., Vilar, M. G., Colonna, V., Pagani, L., & Destro Bisol, G. (2019). Inter-individual genomic heterogeneity within European population isolates. *PLoS One*, *14*(10), e0214564. <https://doi.org/10.1371/journal.pone.0214564>
- Anagnostou, P., Dominici, V., Battaglia, C., Pagani, L., Vilar, M., Wells, R. S., Pettener, D., Sarno, S., Boattini, A., Francalacci, P., Colonna, V., Vona, G., Calò, C., Destro Bisol, G., & Tofanelli, S. (2017). Overcoming the dichotomy between open and isolated populations using genomic data from a large European dataset. *Scientific Reports*, *7*, 41614. <https://doi.org/10.1038/srep41614>
- Antonio, M. L. (2019). Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science*, *366*(6466), 708–714. <https://doi.org/10.1126/science.aay6826>
- Ariotti, E., Fregni, E., & Torresani, S. (1990). *Le Partecipanze Agrarie Emiliane. La storia, le fonti, il rapporto col territorio*. Nonantola, Italy: Grafiche 4-Esse.
- Balaresque, P., Bowden, G. R., Adams, S. M., Leung, H. Y., King, T. E., Rosser, Z. H., Goodwin, J., Moisan, J. P., Richard, C., Millward, A., Demaine, A. G., Barbujani, G., Previderè, C., Wilson, I. J., Tyler-Smith, C., & Jobling, M. A. (2010). A predominantly neolithic origin for European paternal lineages. *PLoS Biology*, *8*(1), e1000285. <https://doi.org/10.1371/journal.pbio.1000285>
- Ballantyne, K. N., Goedbloed, M., Fang, R., Schaap, O., Lao, O., Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vézina, H., Knoblach, H., Lessig, R., Roewer, L., ... Kayser, M. (2010). Mutability of Y-chromosomal microsatellites: Rates, characteristics, molecular bases, and forensic implications. *The American Journal of Human Genetics*, *87*, 341–353.
- Boattini, A., Lisa, A., Fiorani, O., Zei, G., Pettener, D., & Manni, F. (2012). General method to unravel ancient population structures through surnames, final validation on Italian data. *Human Biology*, *84*(3), 235–270. <https://doi.org/10.3378/027.084.0302>
- Boattini, A., Luiselli, D., Sazzini, M., Useli, A., Tagarelli, G., & Pettener, D. (2011). Linking Italy and the Balkans. A Y-chromosome perspective from the Arbereshe of Calabria. *Annals of Human Biology*, *38*(1), 59–68. <https://doi.org/10.3109/03014460.2010.491837>
- Boattini, A., Martinez-Cruz, B., Sarno, S., Harmant, C., Useli, A., Sanz, P., Yang-Yao, D., Manry, J., Ciani, G., Luiselli, D., Quintana-Murci, L., Comas, D., Pettener, D., & the Genographic Consortium. (2013). Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata. *PLoS One*, *8*(5), e65441. <https://doi.org/10.1371/journal.pone.0065441>
- Boattini, A., Sarno, S., Bini, C., Pesci, V., Barbieri, C., de Fanti, S., Quagliarello, A., Pagani, L., Ayub, Q., Ferri, G., Pettener, D., Luiselli, D., & Pelotti, S. (2016). Mutation rates and discriminating power for 13 rapidly-mutating Y-STRs between related and unrelated individuals. *PLoS One*, *11*(11), e0165678. <https://doi.org/10.1371/journal.pone.0165678>
- Boattini, A., Sarno, S., Mazzarisi, A. M., Viroli, C., de Fanti, S., Bini, C., Larmuseau, M. H. D., Pelotti, S., & Luiselli, D. (2019). Estimating Y-Str mutation rates and Tmrc through deep-rooting Italian pedigrees. *Scientific Reports*, *9*(1), 9032. <https://doi.org/10.1038/s41598-019-45398-3>
- Boattini, A., Sarno, S., Pedrini, P., Medoro, C., Carta, M., Tucci, S., Ferri, G., Alù, M., Luiselli, D., & Pettener, D. (2015). Traces of medieval migrations in a socially stratified population from northern Italy. Evidence from uniparental markers and deep-rooted pedigrees. *Heredity*, *114*(2), 155–162. <https://doi.org/10.1038/hdy.2014.77>
- Busby, G. B. J., Brisighelli, F., Sánchez-Diz, P., Ramos-Luis, E., Martínez-Cadenas, C., Thomas, M. G., Bradley, D. G., Gusmão, L., Winney, B.,



- Bodmer, W., Vennemann, M., Coia, V., Scarnicci, F., Tofanelli, S., Vona, G., Ploski, R., Vecchiotti, C., Zemunik, T., Rudan, I., ... Capelli, C. (2012). The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proceedings of Biological Sciences*, 279(1730), 884–892. <https://doi.org/10.1098/rspb.2011.1044>
- Chaix, R., Quintana-Murci, L., Hegay, T., Hammer, M. F., Mobasher, Z., Austerlitz, F., & Heyer, E. (2007). From social to genetic structures in Central Asia. *Current Biology*, 17(1), 43–48. <https://doi.org/10.1016/j.cub.2006.10.058>
- Clærhout, S., Vandenbosch, M., Nivelles, K., Gruyters, L., Peeters, A., Larmuseau, M. H. D., & Decorte, R. (2018). Determining Y-STR mutation rates in deep-rooting genealogies: Identification of haplogroup differences. *Forensic Science International. Genetics*, 34, 1–10. <https://doi.org/10.1016/j.fsigen.2018.01.005>
- Cocca, M. (2020). A bird's-eye view of Italian genomic variation through whole-genome sequencing. *European Journal of Human Genetics*, 28, 435–444.
- Coia, V., Boschi, I., Trombetta, F., Cavulli, F., Montinaro, F., Destro-Bisol, G., Grimaldi, S., & Pedrotti, A. (2012). Evidence of high genetic variation among linguistically diverse populations on a micro-geographic scale: A case study of the Italian Alps. *Journal of Human Genetics*, 57(4), 254–260. <https://doi.org/10.1038/jhg.2012.14>
- Coia, V., Capocasa, M., Anagnostou, P., Pascali, V., Scarnicci, F., Boschi, I., Battaglia, C., Crivellaro, F., Ferri, G., Alù, M., Brisighelli, F., Busby, G. B. J., Capelli, C., Maixner, F., Cipollini, G., Viazzo, P. P., Zink, A., & Destro Bisol, G. (2013). Demographic histories, isolation and social factors as determinants of the genetic structure of alpine linguistic groups. *PLoS One*, 8(12), e81704. <https://doi.org/10.1371/journal.pone.0081704>
- Colonna, V., Pistis, G., Bomba, L., Mona, S., Matullo, G., Boano, R., Sala, C., Viganò, F., Torroni, A., Achilli, A., Hooshiar Kashani, B., Malerba, G., Gambaro, G., Soranzo, N., & Toniolo, D. (2013). Small effective population size and genetic homogeneity in the Val Borbera isolate. *European Journal of Human Genetics*, 21, 89–94.
- Cori, G. (2011). *Massenzatica. Dal comune al C.U.M. Tipo-Lito Grafiche Riunite*.
- Costato, L. (1968). *I domini collettivi nel medio Polesine*. Ed. Giuffrè.
- Destro Bisol, G., Anagnostou, P., Batini, C., Battaglia, C., Bertoncini, S., Boattini, A., Caciagli, L., Calò, M. C., Capelli, C., Capocasa, M., Castrí, L., Ciani, G., Coia, V., Corrias, L., Crivellaro, F., Ghiani, M. E., Luiselli, D., Mela, C., Melis, A., ... Pettener, D. (2008). Italian isolates today: Geographic and linguistic factors shaping human biodiversity. *Journal of Anthropological Sciences*, 86, 179–188.
- Esko, T., Mezzavilla, M., Nelis, M., Borel, C., Debniak, T., Jakkula, E., Julia, A., Karachanak, S., Khrunin, A., Kisfalvi, P., Krulisova, V., Aušrelė Kučinskienė, Z., Rehnström, K., Traglia, M., Nikitina-Zake, L., Zimprich, F., Antonarakis, S. E., Estivill, X., Glavač, D., ... D'Adamo, P. (2013). Genetic characterization of northeastern Italian population isolates in the context of broader European genetic diversity. *European Journal of Human Genetics*, 21, 659–665.
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, 9, e1003905.
- Excoffier, L., Laval, G., & Schneider, S. (2007). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47–50.
- Fenner, J. N. (2005). Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal of Physical Anthropology*, 128(2), 415–423. <https://doi.org/10.1002/ajpa.20188>
- Fernandes, D. M., Mitnik, A., Olalde, I., Lazaridis, I., Cheronet, O., Rohland, N., Mallick, S., Bernardos, R., Broomandkhoshbacht, N., Carlsson, J., Culleton, B. J., Ferry, M., Gamarra, B., Lari, M., Mah, M., Michel, M., Modi, A., Novak, M., Oppenheimer, J., ... Reich, D. (2020). The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology and Evolution*, 4(3), 334–345. <https://doi.org/10.1038/s41559-020-1102-0>
- Fregni, E. (1992). *Terre e Comunità nell'Italia Padana. Il caso delle Partecipanze Agrarie Emiliane: da beni comuni a beni collettivi*. Edizioni Centro Federico Odierici.
- Gokcumen, O., Gultekin, T., Alakoc, Y. D., Tug, A., Gulec, E., & Schurr, T. G. (2011). Biological ancestries, kinship connections, and projected identities in four central Anatolian settlements: Insights from culturally contextualized genetic anthropology. *American Anthropologist*, 113(1), 116–131. <https://doi.org/10.1111/j.1548-1433.2010.01310.x>
- Hallast, P., Batini, C., Zadić, D., Maisano Delser, P., Wetton, J. H., Arroyo-Pardo, E., Cavalleri, G. L., de Knijff, P., Destro Bisol, G., Dupuy, B. M., Eriksen, H. A., Jorde, L. B., King, T. E., Larmuseau, M. H., Lopez de Munain, A., Lopez-Parra, A. M., Loutradis, A., Milasin, J., Novelletto, A., ... Jobling, M. A. (2015). The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. *Molecular Biology and Evolution*, 32(3), 661–673. <https://doi.org/10.1093/molbev/msu327>
- Harney, E., May, H., Shalem, D., Rohland, N., Mallick, S., Lazaridis, I., Sarig, R., Stewardson, K., Nordenfelt, S., Patterson, N., Hershkovitz, I., & Reich, D. (2018). Ancient DNA from chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nature Communications*, 9(1), 3336. <https://doi.org/10.1038/s41467-018-05649-9>
- Helgason, A., Hrafnkelsson, B., Gulcher, J. R., Ward, R., & Stefánsson, K. (2003). A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: Evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *American Journal of Human Genetics*, 72(6), 1370–1388. <https://doi.org/10.1086/375453>
- Huang, Y. Z., Pamjav, H., Flegontov, P., Stenzl, V., Wen, S. Q., Tong, X. Z., Wang, C. C., Wang, L. X., Wei, L. H., Gao, J. Y., Jin, L., & Li, H. (2018). Dispersals of the Siberian Y-chromosome haplogroup Q in Eurasia. *Molecular Genetics and Genomics*, 293(1), 107–117. <https://doi.org/10.1007/s00438-017-1363-8>
- Jombart, T. (2008). A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 94.
- Kutanan, W., Kampuansai, J., Srikumool, M., Brunelli, A., Ghirrotto, S., Arias, L., Macholdt, E., Hübner, A., Schröder, R., & Stoneking, M. (2019). Contrasting paternal and maternal genetic histories of Thai and Lao populations. *Molecular Biology and Evolution*, 36(7), 1490–1506. <https://doi.org/10.1093/molbev/msz083>
- Mantovani, L. (2017). *Antropologia dei commons. Ambiente, storia e memoria nelle partecipanze agrarie emiliane*. University of Bologna, Bologna.
- Marcus, J. H., Posth, C., Ringbauer, H., Lai, L., Skeates, R., Sidore, C., Beckett, J., Furtwängler, A., Olivieri, A., Chiang, C. W. K., al-Asadi, H., Dey, K., Joseph, T. A., Liu, C. C., der Sarkissian, C., Radzevičiūtė, R., Michel, M., Gradoli, M. G., Marongiu, P., ... Novembre, J. (2020). Genetic history from the middle Neolithic to present on the Mediterranean Island of Sardinia. *Nature Communications*, 11(1), 939. <https://doi.org/10.1038/s41467-020-14523-6>
- Mendez, F. L., Karafet, T. M., Krahn, T., Ostrer, H., Soodyall, H., & Hammer, M. F. (2011). Increased resolution of Y chromosome haplogroup T defines relationships among populations of the near east, Europe, and Africa. *Human Biology*, 83(1), 39–53. <https://doi.org/10.3378/027.083.0103>
- Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16, 1215.
- Myres, N. M., Rootsi, S., Lin, A. A., Järve, M., King, R. J., Kutuev, I., Cabrera, V. M., Khusnutdinova, E. K., Pshenichnov, A., Yunusbayev, B., Balanovsky, O., Balanovska, E., Rudan, P., Baldovic, M., Herrera, R. J., Chiaroni, J., di Cristofaro, J., Vilems, R., Kivisild, T., & Underhill, P. A. (2011). A major Y-chromosome haplogroup R1b Holocene era founder

- effect in central and Western Europe. *European Journal of Human Genetics*, 19(1), 95–101. <https://doi.org/10.1038/ejhg.2010.146>
- Platt, D. E., Haber, M., Dagher-Kharrat, M. B., Douaihy, B., Khazen, G., Ashrafian Bonab, M., Salloum, A., Mouzaya, F., Luiselli, D., Tyler-Smith, C., Renfrew, C., Matisoo-Smith, E., & Zalloua, P. A. (2017). Mapping post-glacial expansions: The peopling of Southwest Asia. *Scientific Reports*, 7, 40338. <https://doi.org/10.1038/srep40338>
- Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866. <https://doi.org/10.1093/bioinformatics/btv684>
- Purps, J., Siegert, S., Willuweit, S., Nagy, M., Alves, C., Salazar, R., Angustia, S. M. T., Santos, L. H., Anslinger, K., Bayer, B., Ayub, Q., Wei, W., Xue, Y., Tyler-Smith, C., Bafalluy, M. B., Martínez-Jarreta, B., Egyed, B., Balitzki, B., Tschumi, S., ... Roewer, L. (2014). A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Science International. Genetics*, 12(100), 12–23. <https://doi.org/10.1016/j.fsigen.2014.04.008>
- R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sarno, S., Boattini, A., Carta, M., Ferri, G., Alù, M., Yao, D. Y., Ciani, G., Pettener, D., & Luiselli, D. (2014). An ancient Mediterranean melting pot: Investigating the uniparental genetic structure and population history of Sicily and southern Italy. *PLoS One*, 9(4), e96074. <https://doi.org/10.1371/journal.pone.0096074>
- Sarno, S., Boattini, A., Pagani, L., Sazzini, M., de Fanti, S., Quagliariello, A., Gnechi Ruscone, G. A., Guichard, E., Ciani, G., Bortolini, E., Barbieri, C., Cilli, E., Petrilli, R., Mikerezi, I., Sineo, L., Vilar, M., Wells, S., Luiselli, D., & Pettener, D. (2017). Ancient and recent admixture layers in Sicily and southern Italy trace multiple migration routes along the Mediterranean. *Scientific Reports*, 7(1), 1984. <https://doi.org/10.1038/s41598-017-01802-4>
- Sarno, S., Petrilli, R., Abondio, P., de Giovanni, A., Boattini, A., Sazzini, M., de Fanti, S., Cilli, E., Ciani, G., Gentilini, D., Pettener, D., Romeo, G., Giuliani, C., & Luiselli, D. (2021). Genetic history of Calabrian Greeks reveals ancient events and long term isolation in the Aspromonte area of southern Italy. *Scientific Reports*, 11, 3045. <https://doi.org/10.1038/s41598-021-82591-9>
- Sarno, S., Tofanelli, S., de Fanti, S., Quagliariello, A., Bortolini, E., Ferri, G., Anagnostou, P., Brisighelli, F., Capelli, C., Tagarelli, G., Sineo, L., Luiselli, D., Boattini, A., & Pettener, D. (2016). Shared language, diverging genetic histories: High-resolution analysis of Y-chromosome variability in Calabrian and Sicilian Arbereshe. *European Journal of Human Genetics*, 24(4), 600–606. <https://doi.org/10.1038/ejhg.2015.138>
- Sazzini, M., Abondio, P., Sarno, S., Gnechi-Ruscone, G. A., Ragno, M., Giuliani, C., de Fanti, S., Ojeda-Granados, C., Boattini, A., Marquis, J., Valsesia, A., Carayol, J., Raymond, F., Pirazzini, C., Marasco, E., Ferrarini, A., Xumerle, L., Collino, S., Mari, D., ... Garagnani, P. (2020). Genomic history of the Italian population recapitulates key evolutionary dynamics of both continental and southern Europeans. *BMC Biology*, 18(1), 51. <https://doi.org/10.1186/s12915-020-00778-4>
- Sazzini, M., Gnechi Ruscone, G. A., Giuliani, C., Sarno, S., Quagliariello, A., de Fanti, S., Boattini, A., Gentilini, D., Fiorito, G., Catanoso, M., Boiardi, L., Croci, S., Macchioni, P., Mantovani, V., di Blasio, A. M., Matullo, G., Salvarani, C., Franceschi, C., Pettener, D., ... Luiselli, D. (2016). Complex interplay between neutral and adaptive evolution shaped differential genomic background and disease susceptibility along the Italian peninsula. *Scientific Reports*, 6, 32513. <https://doi.org/10.1038/srep32513>
- Sengupta, S., Zhivotovsky, L. A., King, R., Mehdi, S. Q., Edmonds, C. A., Chow, C. E. T., Lin, A. A., Mitra, M., Sil, S. K., Ramesh, A., Usha Rani, M. V., Thakur, C. M., Cavalli-Sforza, L. L., Majumder, P. P., & Underhill, P. A. (2006). Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of central Asian pastoralists. *The American Journal of Human Genetics*, 78, 202–221.
- Sineo, L., Crivellaro, F., Boattini, A., Brisighelli, F., Anagnostou, P., Franceschi, Z. A., Corrias, L., Sanna, D., Capocasa, M., Bachis, V., De Fanti, S., Dominici, V., Carta, M., Sarno, S., Vona, G., Bertocchini, S., Taglioli, L., Coia, V., Francalacci, P., ... Boschi, I. (2014). Linguistic, geographic and genetic isolation: A collaborative study of Italian populations. *Journal of Anthropological Sciences*, 92, 201–231. <https://doi.org/10.4436/JASS.92001>
- Tremblay, M., & Vézina, H. (2000). New estimates of intergenerational time intervals for the calculation of age and origins of mutations. *American Journal of Human Genetics*, 66(2), 651–658. <https://doi.org/10.1086/302770>
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.
- Venturoli, R. (2004). *La Partecipanza Agraria di Nonantola*. Centro studi storici nonantolani.
- Wegmann, D., Leuenberger, C., & Excoffier, L. (2009). Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, 182(4), 1207–1218.
- Wilson, I., Weale, M., & Balding, D. (2003). Inferences from DNA data: Population histories, evolutionary processes and forensic match probabilities. *Journal of the Royal Statistical Society: Series A*, 166(2), 155–188.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Sarno, S., Boscolo Agostini, R., De Fanti, S., Ferri, G., Ghirotto, S., Modenini, G., Pettener, D., & Boattini, A. (2021). Y-chromosome variability and genetic history of Commons from Northern Italy. *American Journal of Physical Anthropology*, 175(3), 665–679. <https://doi.org/10.1002/ajpa.24302>