

Spatial Distance Between Sites of Sampling Associated With Genetic Variation Among *Neospora Caninum* in Aborted Bovine Foetuses From Northern Italy

Luca Villa

Università degli Studi di Milano: Università degli Studi di Milano

Pavlo Maksimov

Friedrich-Loeffler-Institute Federal Research Institute for Animal Health: Friedrich-Loeffler-Institut Bundesforschungsinstitut für Tiergesundheit

Christine Luttermann

Friedrich-Loeffler-Institute Federal Research Institute for Animal Health: Friedrich-Loeffler-Institut Bundesforschungsinstitut für Tiergesundheit

Mareen Tuschy

Friedrich-Loeffler-Institute Federal Research Institute for Animal Health: Friedrich-Loeffler-Institut Bundesforschungsinstitut für Tiergesundheit

Alessia L. Gazzonis

Università degli Studi di Milano: Università degli Studi di Milano

Sergio A. Zanzani

Università degli Studi di Milano: Università degli Studi di Milano

Michele Mortarino

Università degli Studi di Milano: Università degli Studi di Milano

Franz J. Conraths

Friedrich-Loeffler-Institute Federal Research Institute for Animal Health: Friedrich-Loeffler-Institut Bundesforschungsinstitut für Tiergesundheit

Maria Teresa Manfredi

Università degli Studi di Milano: Università degli Studi di Milano

Gereon Schares (✉ Gereon.Schares@fli.de)

Friedrich-Loeffler-Institut <https://orcid.org/0000-0002-3217-289X>

Research

Keywords: *Neospora caninum*, microsatellite typing, multilocus genotyping, bovine abortion, Holstein Friesian cattle, Italy.

Posted Date: October 14th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-88788/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Parasites & Vectors on January 13th, 2021. See the published version at <https://doi.org/10.1186/s13071-020-04557-6>.

Abstract

Background: *Neospora caninum*, a coccidian protozoan, represents an important cause of bovine abortion. While only a single genotype of *N. caninum* exists world-wide, available *N. caninum* strains show considerable variation *in vitro* and *in vivo*, including different virulence in cattle. To which extent sexual recombination, possible in the intestines of domestic dogs and closely related carnivores as definitive hosts, contribute to this variation is not clear, yet.

Methods: Aborted bovine foetuses were collected between 2015 and early 2019 from Italian Holstein Friesian dairy herds suffering from reproductive problems. A total of 198 samples were collected from 165 intensive farms located in Lombardy, northern Italy. *N. caninum* samples were subjected to multilocus-microsatellite genotyping (MLMG) using ten previously established microsatellite markers. In addition to own data, those from a recent study providing data on five markers from other northern Italian regions were included and analyzed.

Results: Of the 55 samples finally subjected to MLMG, 35 were typed at all or 9 out of 10 loci. Linear regression revealed a statistically significant association between the spatial distance of the sampling sites with the genetic distance of *N. caninum* MLMGs ($P < 0.001$). Including data from a previous North Italian study (eBURST analysis) revealed that part of *N. caninum* MLMGs from northern Italy separate into four groups; most of the samples from Lombardy clustered in one of these groups. Principle component analysis revealed similar clusters and confirmed MLMG groups identified by eBURST. Variations observed between MLMGs were not equally distributed over all loci, but predominantly observed in MS7, MS6A, or MS10.

Conclusions: Our findings confirm the concept of local *N. caninum* subpopulations. The geographic distance of sampling was associated with the genetic distance as determined by MLMGs. Results suggest that multi-parental recombination in *N. caninum* is a rare event, but does not exclude uniparental mating. More comprehensive studies on microsatellites in *N. caninum* and related species like *Toxoplasma gondii* should be undertaken, not only to improve genotyping capabilities, but also to understand possible functions of these regions in the genomes of these parasites.

Background

Neospora caninum is a protozoan coccidian parasite closely related to *Toxoplasma gondii* and *Besnoitia besnoiti*. It has a worldwide distribution and causes foetal losses or stillbirth in livestock, especially cattle [1, 2]. Domestic dogs and other phylogenetically closely related wild carnivores like coyotes, wolves, or dingoes are the only known definitive host of *N. caninum* [1, 3, 4, 5], i.e. hosts in which sexual recombination can take place. The main mode of transmission in cattle, the main intermediate host of the parasite, seems to be endogenous vertical transmission [1, 2], i.e. transmission from a persistently or chronically infected dam to her offspring [6]. Exogenous vertical transmission to a foetus or an unborn calf occurs in cattle that became infected during pregnancy after the ingestion of oocysts shed by a definitive host [6, 7]. Transmission via oocysts by definitive hosts, most likely farm dogs, seems to occur frequently in cattle populations. These infections are the likely cause of epidemic abortions in bovine herds [8]. Such abortion storms were reported to be associated with both (i) low avidity, i.e. a recently established antibody response against *N. caninum* [8, 9] and (ii) an identical microsatellite genotyping pattern in *N. caninum* detected in several foetuses of a single herd [10].

A recent study on the genomes of 50 *N. caninum* isolates collected worldwide from a wide range of hosts using 19 linked and unlinked genetic markers showed that there is only a single genotype of *N. caninum* worldwide [11]. Moreover, whole genome sequencing of seven isolates from two continents revealed less than 10^4 bi-allelic single nucleotide polymorphisms (SNPs), which is very little compared to the situation in the genome of the closely related apicomplexan parasite *T. gondii* with more than 10^6 SNPs between the compared strains [11]. Since more than 50% of the SNP clustered in 6 haploblocks which had been already partially observed in a former study [12], it was concluded that uni-parental reproduction together with a non-sexual expansion formed the actual world-wide *N. caninum* population rather than non-sexual expansion alone [11]. The extremely limited genetic diversity was explained by genetic bottlenecks during the domestication of cattle in the Near East probably about 10 K years ago [11]. Breeding and moving specific cattle breeds including the Holstein Friesian breed examined here during the following centuries may also have contributed [13]. The domestication of dogs may have formed another bottleneck, which shaped the *N. caninum* population that exists today [11].

Nevertheless, *N. caninum* is less uniform than the previously mentioned study [11] suggests. In fact, different *N. caninum* isolates show large differences *in vitro* [14, 15, 16] and *in vivo* [17, 18, 19], also in cattle [20, 21, 22, 23]. The pioneering work of Regidor et al. [24] contributed a number of microsatellite markers that allow to fingerprint *N. caninum* isolates or DNAs and undertake population studies.

We aimed at genotyping *N. caninum* in aborted bovine foetuses from Lombardy, one of the most important dairy cattle production areas in Italy [25]. In this region, where the Italian Holstein Friesian breed prevails, *N. caninum* was recognized as an important cause of abortion; however, data are unpublished, yet (MTM, unpublished data). Moreover, a high *N. caninum* seroprevalence was revealed in cattle from northern Italy [26, 27]. We therefore determined the proportion of *N. caninum* PCR-positive aborted foetuses in this area and characterised the available isolates by multi-locus microsatellite genotyping.

A previous world-wide study [28], a South-American study [29], a local study from Spain [30], and a recent study conducted in northern Italy [31] suggested the existence of *N. caninum* sub-populations specific for particular countries or regions. The North-Italian study conducted predominantly in two areas (Piedmont and Veneto-Trento) close to our study area, suggested at least three sub-populations. We aimed to find out, how *N. caninum* isolates in aborted foetuses from Lombardy fit into the pattern of these sub-populations. To better understand the reasons for the genetic differences, we used detailed geographic information on the sites, where *N. caninum*-positive foetal material had been sampled.

Material And Methods

Study area

All farms, from which aborted fetuses had been sampled, were located in the “Bassa Padana” in the Po valley. This area comprises the territory of the provinces of Lodi, Cremona, Mantova, Pavia, and the south of the provinces of Milano and Brescia (Fig. 1). This is one of the largest dairy production areas in Italy and stands out for the high density of cattle farms, mainly based on Italian Holstein Friesian under the intensive production system. In particular, the Italian National Zootechnical Registry counts 5446 dairy farms hosting 1,071,164 animals in the Lombardy region, corresponding to the 20.7% and 40.8% of all farms and animals in Italy, with an annual production of 5,215,408 tons of milk (http://www.assolatte.it/zpublish/4/uploads/4/news_down/15641362883148578112_RAPPORTO%20ASSOLATTE%202018.pdf; last access 06.08.2020). Most farms in the area host between 100 and 500 animals, but several farms hold more than 500 cattle (National Zootechnical Database, <https://www.vetinfo.sanita.it>; last access 06.08.2020).

The sites have an average altitude of about 69 m above sea level. The climate is continental, typical of the Po valley, with hot muggy summers with a few thunderstorms and cold and foggy winters with some snow. There is a large annual thermal excursion with a mean maximum temperature of 25–28 °C and a mean minimum temperature of -1 to -2 °C. Rainfall is distributed over the year with peaks in spring and autumn. The average total annual rainfall ranges between 700 and 1200 mm.

Sample collection

Aborted bovine fetuses were collected between 2015 and early 2019 from Italian Holstein Friesian dairy herds suffering from reproductive problems. Overall, 198 samples were collected from 165 intensive farms located in the Po valley (Provinces of Cremona, Lodi, Mantova, and Milano). Abortion events had been notified by the farm veterinarian, fetuses were maintained refrigerated and collected directly from the farms or submitted to the laboratory. In the laboratory, a pool of organs was prepared from each aborted foetus, which included brain, lung, and liver. This pool was mechanically homogenized and stored at -20 °C until molecular analysis. The geographical localization (municipality and province as well as the geographic coordinates) of each farm with abortion episodes was recorded along with the data for each aborted foetus.

DNA extraction

Genomic DNA was extracted from tissue homogenates of pooled target organs (brain, lung, liver) of aborted fetuses using a commercial kit (NucleoSpin® Tissue, Macherey-Nagel, Düren, Germany), according to the manufacturer's instructions. To optimize DNA extraction from aborted fetuses, a 200 µl aliquot of the homogenate was pre-lysed with buffer T1 and proteinase K as recommended for the NucleoSpin Tissue Kit, by scaling up eight times the volumes used for the initial digestion. After digestion (56 °C, overnight), 230 µl of the final suspension (1840 µl) were taken and the protocol of the NucleoSpin Tissue kit followed as recommended. A negative extraction control was included in each batch of approximately twenty-four samples, which was subsequently tested by real-time PCR in the same way as the samples from aborted fetuses were analyzed.

Screening *N. caninum* RT-qPCR

Fetal DNA samples were initially subjected to a *N. caninum*-specific RT-qPCR with the primers and probes targeting the Nc5 gene as previously described [32, 33]. Primers were used at a final concentration of 600 nM and the fluorogenic probe (NeoProbe) at 100 nM. Ten µl of 2 x iQ Supermix (Bio-Rad Laboratories GmbH, Feldkirchen, Germany) were added to each reaction. For PCR analysis, 5 µl of DNA from each sample was used in a 20 µl reaction on 96-wells reaction plate. The thermal cycling protocol consisted of an initial polymerase activation and DNA denaturation at 95 °C for 5 min, followed by 46 cycles of amplification including denaturation at 95 °C for 10 s and annealing/extension at 58 °C for 30 s. NC-1 [34] was included as a positive control, while water PCR Reagent (Sigma–Aldrich) was used as a negative control; negative extraction controls were also included.

Microsatellite typing

Samples confirmed as positive for *N. caninum* were subjected to multilocus-microsatellite typing using ten microsatellite markers (MS1B, 2, 3, 5, 6A, 6B, 7, 10, 12, and 21). Nested-PCR (n-PCR) techniques were used for the amplification of microsatellite-containing regions *N. caninum* DNA from aborted fetuses. Microsatellites MS2 and MS10 were amplified and sequenced using primers and protocols previously described [35]. The microsatellites MS1B, 3, 5, 6A, 6B, 7, 12, and 21 were analyzed by nested-PCR and fragment lengths determined by capillary electrophoresis [10]. Forward primers of the internal PCRs were labeled with either 6-FAM, HEX, or NED dyes to allow the simultaneous length determination of three different microsatellites per capillary electrophoresis. In addition, previously sequenced microsatellite-containing fragments from the *in vitro*-grown *N. caninum* NC-1 strain [36] were included in the amplifications, the sequencing, and sizing analyses as a positive control.

For all n-PCRs, primers were used at a final concentration of 0.5 µM, and dNTPs at 250 µM each (Amersham Biosciences, Piscataway, USA). DyNAzyme II DNA polymerase (Finnzymes, Espoo, Finland) was added at 1 U/25 µl with the provided buffer. The reaction mix was supplemented with bovine serum albumin at a concentration of 20 µg/ml. 1.5 µl of genomic DNA or 1 µl of amplification product were used as template for the external and internal PCR amplification steps, respectively. Water PCR Reagent (Sigma–Aldrich) was used as a negative control. DNA from cell culture-derived *N. caninum* tachyzoites (NC-1) was used as a positive control. The PCR reactions were performed in a thermal cycler (Eppendorf Mastercycler, Personal Thermal Cycler). The external PCR was performed with an initial denaturation step of 94 °C for 5 min, followed by 35 cycles of denaturation (1 min at 94 °C), annealing (1 min at 60 °C for all MS, except for MS2: 50 °C, and MS10: 54 °C) and extension (1 min at 72 °C), and a final extension step at 72 °C for 10 min. The internal PCR consisted of an initial denaturation of 94 °C for 5 min, followed by 35 cycles of denaturation: 1 min at 94 °C, annealing: 1 min at 60 °C (MS2, 10) or 30 s at 52 °C (MS1B, 3, 5, 6A, 21), 54 °C (MS6B, 12) or 60 °C (MS7) and extension: 1 min at 72 °C (MS2, 10), 30 s at 70 °C (MS1B, 3, 5, 6A, 6B, 12, 21) or 65 °C (MS7), and a final extension step at 72 °C (MS2,10), 70 °C (MS1B, 3, 5, 6A, 6B, 12, 21) or 65 °C (MS7), for 10 min.

The amplification products were visualized after electrophoresis in 2% agarose gels stained with ethidium bromide using a 100 bp DNA ladder (Invitrogen GmbH, Karlsruhe, Germany) as a size standard.

For sequencing MS2 and MS10, bands of the expected size were excised from the agarose gels and purified with a commercial kit (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel, Düren, Germany), following the manufacturer's instructions. Purified amplification products were then cloned into a commercially available vector (pGEM®-T Easy Vector System I, Promega, Mannheim, Germany) and used to transform chemically competent *Escherichia coli* (OneShot TOP10, Thermo Fisher Scientific, Langenselbold, Germany). The transformed *E. coli* were cultivated, and the plasmid DNA was subsequently collected using a commercial kit (QIAprep Spin Miniprep Kit, Qiagen, Hilden, Germany), according to the manufacturer's instruction. Sequencing was finally performed using the BigDye Terminator v1.1 Cycle SEq. Kit (Thermo Fisher Scientific, Langenselbold, Germany) and passage with NucleoSEQ Columns (Macherey-Nagel, Düren, Germany) for cleaning up nucleic acids, in an ABI 3130 capillary sequencer (Thermo Fisher Scientific, Langenselbold, Germany).

For all other MS, *N. caninum*-positive samples were analyzed on an ABI 3130 capillary sequencer (Thermo Fisher Scientific, Langenselbold, Germany). A ROX dye-labeled standard (GeneScan™ 500 ROX™ dye Size Standard, Thermo Fisher Scientific, Langenselbold, Germany) was included in each analysis as a size reference. The results were analyzed using the Geneious 1.11.5 software (Biomatters, Inc., 2365 Northside Dr., Suite 560, San Diego, CA 92108, USA).

Statistical analysis

Differences in *N. caninum* specific RT-qPCR results (Ct values) in relation to the number of microsatellite markers typed were assessed using the command "wilcox.test" from the package "stats" in R, version 3.5.3 (R Foundation for Statistical Computing, Vienna, Austria; <http://www.R-project.org>). Microsatellite marker-specific differences in the number of repeats between *N. caninum* isolates from different regions were assessed using the command "pairwise.wilcox.test" from the package "stats" in R, version 3.5.3 "BH adjusted" option due to multiple testing [37].

Pairwise Bruvo's genetic distance was calculated using MLMG data of all DNAs from aborted fetuses in Lombardy, for which a complete MLMG pattern was obtained (n = 20). Since MS10 and MS2 were combined microsatellite markers, which consist of three sub-loci (or motifs) each, these sub-loci were analyzed separately in a dendrogram. Calculations were done by R, version 3.5.3., employing "Poppr", version 2, a R package designed for the analysis of populations with mixed modes of reproduction [38], by using the command "bruvo.dist". To obtain an overview of the relatedness of MLMGs a dendrogram was established using "Poppr" package representing the results of K-means clustering using Bruvo's Distance with non-parametric bootstrapping.

For testing of a standardized index of association (I_A^S) by multilocus linkage disequilibrium (LD) among different genotypes including nine loci the LIAN v 3.7 web interface program (<http://guanine.evolbio.mpg.de/cgi-bin/lian/lian.cgi.pl/query>) was applied as described [28, 39].

Testing population differentiation was done by likelihood ratio G-statistic [40], using the function "test.between" from the R-package "HierFstat" [41].

Geographic distance based on the geographic coordinates of the farms were calculated (straight-line distance, accounting for curvature of the earth) using the command "spDists" of the package sp, version 1.4-2, in R. To determine the relatedness of Bruvo's distance, geographic distance and the time between samplings linear regression was performed using the "lm" command in R, version 3.5.3 from the package "stats".

The eBURST software [42] was used to generate networks based on MLMGs using the double-locus option (DVL – at least 7 shared loci of 9). Samples with missing data were excluded from the analysis. As the eBURST analysis was performed on samples from this study and those from a previous study, the analysis was restricted to those microsatellite loci that were used in both studies. Since MS10 is a combined microsatellite marker, which consists of three sub-loci (or motifs), these sub-loci were analysed separately. The same data (MLMGs data collected in the present study on foetuses from Lombardy and data from a previous study) were also analysed by a Principal Coordinates Analysis (PCoA) using the command "dudi.pco" employing the R package "Poppr", version 2.

Figures were assembled using R, version 3.5.3 or 4.0.0 (packages "ggplot2", "reshape" and "scales").

Results

Overall *N. caninum* DNA findings in aborted foetuses

One hundred and ninety-eight aborted bovine fetuses from 165 farms located in the Po valley of Lombardy, i.e. the provinces of Cremona, Lodi, Mantova, and Milano, were included in the study. Out of 198 aborted fetuses 55 were positive for *N. caninum* by RT-qPCR, yielding a prevalence of 27.8% (presence of the parasite in pooled tissue homogenates of brain, lung, and liver); 43 farms recorded at least one positive fetus (26.1%). Overall, 55 samples from 43 herds of fetuses collected from 2015 to early 2019 were subjected to microsatellite typing.

Relationship between real time PCR results and success in genotyping

The majority of the 55 DNA samples subjected to multilocus microsatellite typing were typed at all 10 (n = 20), or at 9 (n = 15) loci. Nine DNA samples could be typed at 8 (n = 2), 7 (n = 2), or 6 (n = 6) loci. The remaining 10 DNAs could only be typed at 5 (n = 3), 4 (n = 3), 3 (n = 2), 2 (n = 1), or 1 (n = 1) of the loci.

The numbers of genotyped loci were related to the amount of parasitic DNA in the sample as reflected by the Ct values determined by real time PCR (Table 1).

Microsatellite typing on individual farms at different time points revealed differences in the MLMG results

In seven farms, more than one fetus had been sampled and detected *N. caninum* DNA typed in the study period from 2015–2018. Most of the differences in typing comprised of only one or two repeat units in MLMG (Table 2). However, in five farms, more prominent differences per locus (i.e. differences of 3 or even more repeat units per locus) were observed. In herd 6, i.e. the farm with the largest number of fetuses analysed, 48.1% of the comparisons of individual loci revealed differences (Table 2, Table 3). These more prominent differences affected the loci MS7 and MS21 (Chromosome 7a), MS10 (MS10.2, Chromosome 8), and MS6A (Chromosome 10) (Table 3).

Relationship between the spatial distance of sampling sites and the genetic distance of *N. caninum* isolates

For $n = 20$ of the *N. caninum* positive samples of bovine fetuses from Lombardy, which had been typed at the complete set of microsatellites loci (MS1B, 2, 3, 5, 6A, 6B, 7, 10, 12 and 21, this study), pairwise Bruvo's genetic distances were calculated. Bruvo's genetic distance and the geographical distance between sampling sites of the individual *N. caninum*-positive fetuses (i.e. based on the geographical coordinates of 17 farms) were tested for a correlation. Linear regression revealed that Bruvo's genetic distance correlated statistically significantly with the geographical distance between the sampling sites (Fig. 1a; Table 4). The model (**Model 1**, Table 4) had an adjusted R^2 of 19.5%. Including the number of days between the sampling of fetuses into the model did not improve the model significantly (**Model 2**, Table 4). Bruvo's genetic distance with non-parametric bootstrapping allowed the identification of at least four groups, separated by bootstrap values > 50 (Fig. 1b). While the farms, from which the largest group of samples originated from (Fig. 1b, **red**), were located with two exceptions central-north in the Cremona district (Fig. 1c, **red**), the farms, where the remaining samples had been derived from, were located in the far north (Fig. 1c, **green**), in the south of the Cremona district (Fig. 1c, **blue**), or in the neighbouring district of Mantova (Fig. 1c, **yellow**).

An eBURST and PCoA analysis reveal different groups or clusters of *N. caninum* MLMGs in northern Italy

In addition to the microsatellite data established in this study (Lombardy; $n = 25$), further data on $n = 50$ *N. caninum* isolates from northern Italy covering mainly other regions than Lombardy (i.e., Piedmont $n = 17$, Veneto-Trento $n = 6$, including also one goat isolate), Lombardy ($n = 1$), and a further North-Italian bovine *N. caninum* isolate ($n = 1$) were analysed. These additional data were available from a recent study [31]. Since the latter study had employed a set of microsatellite markers that overlapped only partially with the one we used, our analysis was restricted to MS6A, MS6B, MS10, MS12 and MS21, for which data were available from both groups. Linkage disequilibrium (LD) was assessed for the entire population excluding the goat sample; results ($I_A^S = 0.0411$, $V_D = 1.8249$, $L = 15508$, $P = 0.0411$) indicated LD because $V_D > L$ [39].

Network analysis using eBURST with double locus variation (DLV) among 6 loci and 3 MS10 motifs or sub-loci (MS10.1-3) revealed that many of *N. caninum* MLMGs from Lombardy ($n = 11/25$; this study) clustered separately (eBURST G4) from those obtained from Piedmont ($n = 10/17$, eBURST G1, G2, G3) or Veneto-Trento ($n = 3/6$, eBURST G1, G2) (Fig. 2a). The grouping of isolates from Piedmont and Veneto-Trento had already been assessed by others [31]. This previous grouping of particular *N. caninum* isolates into the eBURST groups G1, G2 and G3 matched perfectly ($n = 10/10$) with our grouping (**Table S1**). Genotyping of an Italian isolate, for which the province of origin was not known (violet dot) revealed results close to most isolates from Lombardy (group G4). The full representation of the MLMST network using the MST option in eBURST, including the complete data set ($n = 49$ *N. caninum* samples), shows the G4 samples in the centre with most of the Veneto-Trento samples on the right side (G1) and the Piedmont samples both left (G2, G3) and right (G1) in the network (Fig. 2b).

Investigating *N. caninum* positive samples ($n = 48$, excluding a sample, for which the province of origin was unknown, and the goat sample) by PCoA revealed a clear axis #1 separation of the samples from Piedmont (Fig. 3) and Lombardy (Fig. 3). While $n = 17/17$ samples from Piedmont were located on the left side, all from Lombardy ($n = 17/25$) were on the right side (Fig. 3). The majority ($n = 3/5$) of samples from Veneto-Trento (Fig. 3) were left or right in the PCoA graph.

When the eBURST grouping (DLV option) was compared to the PCoA result, eBURST G1 and eBURST G3 was located separately in the upper left quarter of the PCoA graph, while eBURST G2 was located in the lower left part (Fig. 3). *Neospora caninum* samples that had been separated by eBURST into G4 (exclusively Lombardy samples) clustered in both the upper and lower corners on the right side of the PCoA graph (Fig. 3).

Differences between the *N. caninum* populations in northern Italy were assessed F_{ST} analysis values (F) and Nei's unbiased genetic distance (D) (Table 5). Both, results for F and D suggest that there is a statistically significant genetic difference between *N. caninum* from Piedmont and Lombardy as well as between Lombardy and Veneto-Trento (Table 5).

The graphical representation of the heterogeneity in numbers of repeat units, stratified for various loci, suggested differences between bovine *N. caninum* sampled in the North Italian regions of Lombardy, Piedmont, and Veneto-Trento (Fig. 4). The range between the lowest and highest number of repeats was largest for MS10.2, followed by MS6A, MS5 and MS7. However, the pairwise Wilcoxon test (BH adjusted P value) revealed statistically significant differences between regions only for MS5, MS6B, MS7, MS10.1, and MS10.2 (Fig. 4).

Discussion

This study focussed on microsatellite multilocus genotyping of *N. caninum* in bovine fetuses sampled in Lombardy, northern Italy, between 2015 and early 2019 from dairy herds suffering from reproductive problems. 27.8% of aborted fetuses (55/198) sampled in 43/165 farms tested positive for *N. caninum*, which shows that this parasite is present in many herds of this region.

A similar study conducted had also focussed on MLMGs in *N. caninum* in cattle and a goat from northern Italy [31]. Both beef (Piedmontese breed) and dairy (Italian Holstein Friesian breed) as well as crossbreeds or local cattle breeds were included, whereas our work was focussed on Italian Holstein Friesian cattle under intensive farming conditions. Regarding the geographical area, this previous study had concentrated on bovine fetuses from the regions of Piedmont and Veneto-Trento ($n = 38$) and included only a few samples ($n = 2$) from Lombardy. In addition, the previous study lacked detailed data on the geographic origin of samples. The regions of Piedmont and Veneto-Trento neighbour our study area; Piedmont is located west and Veneto-Trento east from Lombardy. On the basis of microsatellite typing results, it has been hypothesized that local *N. caninum* subpopulations exist in Europe [30, 32]. We therefore aimed at confirming these findings and finding reasons for genotypic differences between *N. caninum*, even if the isolates or specimens had been sampled from neighbouring areas.

Entirely unexpected was our finding that geographic distance in such a small area as Lombardy (less than 30,000 sqkm size) may at least partially explain the MLMGs genetic distance of *N. caninum* isolates. This suggests that MLMGs pattern remains relatively constant over time in particular herds and only slowly disperse in an area, probably due to animal trade. As it can be assumed that trade between herds was mostly local, this may have resulted in the observed correlation of the geographic distance with the MLMGs-based genetic distance between *N. caninum* isolates. Within herds, *N. caninum* seems to be transmitted endogenously mainly, i.e. from latently infected dams to their offspring. However, even an exogenous transmission via oocysts shedding definitive hosts (e.g. farm dogs or wolves) could explain the conserved microsatellite patterns we observed, if the definitive hosts were infected only with a single strain, i.e. the local one, which results in a so-called uniparental mating of *N. caninum* [11]. Both endogenous and exogenous transmission without or only limited contact to neighbouring herds or infectious material might have resulted in local *N. caninum* sub-populations.

All cattle analysed in our study in Lombardy belonged to the Italian Holstein Friesian breed, which has been founded in the late 20th century by the importation of cattle from the Netherlands and North America [43, 44]. This breed originates from the north of the Netherlands and Germany and became the predominant dairy breed world-wide; present Holstein Friesian cattle represent crosses of the Dutch Friesian and North-American Holstein lines [43, 44]. It was hypothesized that individual *N. caninum* strains may have been imported together with the Holstein Friesian cattle breed into Italy [31]. Although MLMGs of *N. caninum* from northern Italy showed a relationship to those from Germany [31], it is almost impossible to clarify, whether the *N. caninum* genotypes now observed in Italian Holstein Friesian originated from Italy (i.e. from local cattle breeds) or were introduced by the importation of cattle from abroad. Attempts to find genetic differences between *N. caninum* isolates from Holstein Friesian and from local breeds (i.e. the Piedmontese breed), which could have supported the hypothesis of importation, failed [31]. Interestingly, it has been observed that Argentinian *N. caninum* isolates are still related to those from Spain, even long after the importation of the first Iberian cattle into South America in 15th century [28].

In a few herds, we were able to sample more than a single foetus, and in four of these herds, several months or even a period of years separated the sampling dates. In contrast to findings in bovine herds with epidemic abortions (i.e. in herds, where a point source exposure to *N. caninum* oocysts shed by a dog was assumed [10]), particular MLMGs loci differed strongly (i.e. by 3 or even more repeats). It may be discussed, if these larger differences in single loci (MS7, MS6A) or sub-loci (MS10.2) might be an indication for sexual recombination in a definitive host (probably by a farm dog) in the past. However, typing details (Table 2) revealed that related loci on the same chromosome (like MS1B, MS6B) or even sub-loci (e.g. MS10.1 or MS10.2) were not affected. Thus, these differences are best explained by the loss or gain of repeat units, which is typical for microsatellites and caused by point or polymerase template-slippage mutations [45, 46]. Nevertheless, sexual recombination events in these herds cannot completely ruled out, if this recombination had occurred with two or more separate strains (i.e., multi-parental) with very similar microsatellite patterns, e.g. the microsatellite pattern of the local *N. caninum* population.

Our findings in individual herds with more than one *N. caninum*-MLMG result, but also the overall analysis of the *N. caninum* MLMG results in northern Italy, suggests that particular microsatellite loci may be more prone to variation than others. In our study, especially the sub-locus MS10.2, followed by the loci MS6A, MS5 and MS7 showed the strongest variation in the number of repeats. It is well known that several factors contribute to these differences between microsatellite loci, including repeat number, the sequence of the repeat motif, the length of the repeat unit, and flanking sequences (reviewed in [47, 48, 49]). Thus, high variability in particular microsatellite loci may indicate that these loci are of limited importance for the parasite. On the other hand, the observation of changes in particular microsatellite loci raises questions regarding the possible effects of such variations on the affected parasite. Microsatellites are highly abundant in the noncoding DNA of all eukaryotic genomes [50] and changes in these loci might be "neutral", but microsatellites may also locate in coding regions and variation can be associated with an altered phenotype (reviewed [48]). A recent study on different *Plasmodium* spp. affecting humans suggested that a significant proportion (one fifth to one third) of microsatellite-related sequences are related to coding sequences. Based on gene ontology, the respective coding sequences can be involved in molecular functions like binding or in biologic processes such as metabolism or reproduction [51]. Although, we are far from understanding the biological relevance of microsatellites for *N. caninum*; it would be intriguing, if future studies could comprehensively address the microsatellites of *N. caninum* to gain more knowledge on their function. It has so far not been possible to link particular microsatellite patterns to particular traits, e.g. virulence for foetuses [52]. Nevertheless, virulence differences have been observed among *N. caninum* strains also in its main intermediate host, the cattle [15, 17, 18, 20, 21, 22, 23], and it would be fascinating to know, to which extent differences in microsatellite loci contribute to virulence. Recently, highly virulent isolates of *N. caninum* were shown to express a subset of particular secreted proteins in more abundance [53]. The reasons for differential expression between strains could at least in part be due to differences in microsatellite loci as shown for other eukaryotic cells and organisms [50].

In *T. gondii*, an apicomplexan parasite closely related to *N. caninum*, the situation is similar. Currently, relatively conserved microsatellite loci are used to differentiate dominant clonal lineages of *T. gondii*, while also a number of other microsatellite loci are known, which were called finger-printing microsatellite loci. They can be used to differentiate strains of a single clonal type on a local level [54].

Conclusions

Our findings confirm the concept of local *N. caninum* sub-populations. For the first time, we could show a correlation between the genetic distance of *N. caninum* isolates based on MLMGs and the geographic distance of the places, where the isolates had been obtained. Our results confirm that sexual recombination in *N. caninum* is a rare event. Possible reasons for this might be that endogenous vertical transmission is dominating and that the chance for a definitive host to feed on intermediate hosts with a mixed infection of viable and non-related *N. caninum* strains is extremely low. More comprehensive studies on microsatellites in *N. caninum* and on related species like *T. gondii* should be undertaken, not only to improve genotyping capabilities, but also to understand the possible function of these regions in the genomes of these important parasites.

Abbreviations

MLMG, multilocus-microsatellite genotyping

MS, microsatellite

RT-qPCR, real time quantitative PCR

PCoA, principle coordinate analysis

DLV, double locus variant

LOM, Lombardy

PMN, Piedmont

VEN, Veneto-Trento

Declarations

Acknowledgements

The authors are grateful to the veterinarians (particularly, C. Basevi, C. Barisani, S. Bonfanti, M. Colombo, E. Fumagalli, G. Gelati, R. Landriscina, P. Bossi, C. Santambrogio, R. Zanchetta) who helped in the collection of bovine aborted foetuses. We have to thank Patrick Wysocki for providing the map of farm location.

Ethics approval and consent to participate

All procedures were approved by the Institutional Animal Care and Use Committee of Università degli Studi di Milano ("Organismo Preposto al Benessere degli Animali," Prot. no. OPBA_34_2017).

Consent for publication

Not applicable.

Availability of data and materials

Data supporting the conclusions of this article are included within the article and its additional files. The raw datasets used and analyzed during the present study are available as supplementary Table S1 or from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

LV was funded by the German Academic Exchange Service (DAAD) Short-Term Research Grant (No. 57440917).

Authors' contributions

LV, MTM and GS designed the study. LV, SAZ and MTM collected the samples. LV, FH, MT and GS collected the data. LV, ALG, MM, CL, MT, PM and GS performed experiments and analysed the samples. GS and PM statistically analysed the data. LV, PM and GS interpreted the data. GS, LV, PM, FJC and MTM made major contributions to the writing of the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Veterinary Medicine, Università degli Studi di Milano, Via dell'Università 6, 26900 Lodi, Italy

² Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Institute of Epidemiology, National Reference Centre for Toxoplasmosis, Greifswald-Insel Riems, Germany.

³ Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Institute for Immunology, Südufer 10, 17493 Greifswald - Insel Riems, Germany

References

1. Dubey JP, Hemphill A, Calero-Bernal R, Schares G. Neosporosis in animals. Boca Rotan: CRC Press; 2017.
2. Dubey JP, Schares G, Ortega-Mora LM. Epidemiology and Control of Neosporosis and *Neospora caninum*. Clin Microbiol Rev. 2007;20 2:323-67.
3. Dubey JP, Jenkins MC, Rajendran C, Miska K, Ferreira LR, Martins J, et al. Gray wolf (*Canis lupus*) is a natural definitive host for *Neospora caninum*. Vet Parasitol. 2011;181 2-4:382-7.
4. Gondim LF, McAllister MM, Pitt WC, Zemlicka DE. Coyotes (*Canis latrans*) are definitive hosts of *Neospora caninum*. Int J Parasitol. 2004;34 2:159-61.

5. King JS, Slapeta J, Jenkins DJ, Al-Qassab SE, Ellis JT, Windsor PA. Australian dingoes are definitive hosts of *Neospora caninum*. *Int J Parasitol.* 2010;40 8:945-50.
6. Trees AJ, Williams DJ. Endogenous and exogenous transplacental infection in *Neospora caninum* and *Toxoplasma gondii*. *Trends Parasitol.* 2005;21 12:558-61.
7. McCann CM, McAllister MM, Gondim LF, Smith RF, Cripps PJ, Kipar A, et al. *Neospora caninum* in cattle: experimental infection with oocysts can result in exogenous transplacental infection, but not endogenous transplacental infection in the subsequent pregnancy. *Int J Parasitol.* 2007;37 14:1631-9.
8. McAllister MM, Wallace RL, Björkman C, Gao L, Firkins LD. A probable source of *Neospora caninum* infection in an abortion outbreak in dairy cows. *The Bovine Practitioner.* 2005;39:69-74.
9. Schares G, Bärwald A, Staubach C, Söndgen P, Rauser M, Schröder R, et al. p38-avidity-ELISA: examination of herds experiencing epidemic or endemic *Neospora caninum*-associated bovine abortion. *Vet Parasitol.* 2002;106 4:293-305.
10. Basso W, Schares S, Minke L, Bärwald A, Maksimov A, Peters M, et al. Microsatellite typing and avidity analysis suggest a common source of infection in herds with epidemic *Neospora caninum*-associated bovine abortion. *Vet Parasitol.* 2010;173 1-2:24-31.
11. Khan A, Fujita AW, Randle N, Regidor-Cerrillo J, Shaik JS, Shen K, et al. Global selective sweep of a highly inbred genome of the cattle parasite *Neospora caninum*. *Proc Natl Acad Sci U S A.* 2019;116 45:22764-73.
12. Calarco L, Barratt J, Ellis J. Genome wide identification of mutational hotspots in the apicomplexan parasite *Neospora caninum* and the implications for virulence. *Genome Biology and Evolution.* 2018;10 9:2417-31.
13. Felius M, Beerling M-L, Buchanan DS, Theunissen B, Koolmees PA, Lenstra JA. On the History of Cattle Genetic Resources. *Diversity* 2014;6 705-50.
14. Regidor-Cerrillo J, Gomez-Bautista M, Sodupe I, Aduriz G, Alvarez-Garcia G, Del PI, et al. In vitro invasion efficiency and intracellular proliferation rate comprise virulence-related phenotypic traits of *Neospora caninum*. *Vet Res.* 2011;42 1:41.
15. Dellarupe A, Regidor-Cerrillo J, Jimenez-Ruiz E, Schares G, Unzaga JM, Venturini MC, et al. Comparison of host cell invasion and proliferation among *Neospora caninum* isolates obtained from oocysts and from clinical cases of naturally infected dogs. *Exp Parasitol.* 2014;145:22-8.
16. Jimenez-Pelayo L, Garcia-Sanchez M, Regidor-Cerrillo J, Horcajo P, Collantes-Fernandez E, Gomez-Bautista M, et al. Differential susceptibility of bovine caruncular and trophoblast cell lines to infection with high and low virulence isolates of *Neospora caninum*. *Parasites & vectors.* 2017;10.
17. Rojo-Montejo S, Collantes-Fernandez E, Regidor-Cerrillo J, Alvarez-Garcia G, Marugan-Hernandez V, Pedraza-Diaz S, et al. Isolation and characterization of a bovine isolate of *Neospora caninum* with low virulence. *Vet Parasitol.* 2009;159 1:7-16.
18. Regidor-Cerrillo J, Gomez-Bautista M, Del PI, Jimenez-Ruiz E, Aduriz G, Ortega-Mora LM. Influence of *Neospora caninum* intra-specific variability in the outcome of infection in a pregnant BALB/c mouse model. *Vet Res.* 2010;41 4:52.
19. Dellarupe A, Regidor-Cerrillo J, Jimenez-Ruiz E, Schares G, Unzaga JM, Venturini MC, et al. Clinical outcome and vertical transmission variability among canine *Neospora caninum* isolates in a pregnant mouse model of infection. *Parasitology.* 2014;141 3:356-66.
20. Rojo-Montejo S, Collantes-Fernandez E, Blanco-Murcia J, Rodriguez-Bertos A, Risco-Castillo V, Ortega-Mora LM. Experimental infection with a low virulence isolate of *Neospora caninum* at 70 days gestation in cattle did not result in foetopathy. *Vet Res.* 2009;40 5.
21. Chryssafidis AL, Canton G, Chianini F, Innes EA, Madureira EH, Gennari SM. Pathogenicity of Nc-Bahia and Nc-1 strains of *Neospora caninum* in experimentally infected cows and buffaloes in early pregnancy. *Parasitol Res.* 2014;113 4:1521-8.
22. Regidor-Cerrillo J, Arranz-Solis D, Benavides J, Gomez-Bautista M, Castro-Hermida JA, Mezo M, et al. *Neospora caninum* infection during early pregnancy in cattle: how the isolate influences infection dynamics, clinical outcome and peripheral and local immune responses. *Vet Res.* 2014;45:10.
23. Jimenez-Pelayo L, Garcia-Sanchez M, Vazquez P, Regidor-Cerrillo J, Horcajo P, Collantes-Fernandez E, et al. Early *Neospora caninum* infection dynamics in cattle after inoculation at mid-gestation with high (Nc-Spain7)- or low (Nc-Spain1H)-virulence isolates. *Vet Res.* 2019;50 1.
24. Regidor-Cerrillo J, Pedraza-Diaz S, Gomez-Bautista M, Ortega-Mora LM. Multilocus microsatellite analysis reveals extensive genetic diversity in *Neospora caninum*. *J Parasitol.* 2006;92 3:517-24.
25. Zucali M, Tamburini A, Sandrucci A, Bava L. Global warming and mitigation potential of milk and meat production in Lombardy (Italy). *Journal of Cleaner Production.* 2017;153:474-82.
26. Otranto D, Llazari A, Testini G, Traversa D, di Regalbono AF, Badan M, et al. Seroprevalence and associated risk factors of neosporosis in beef and dairy cattle in Italy. *Vet Parasitol.* 2003;118 1-2:7-18.
27. Sala G, Gazzonis A, Boccardo A, Coppoletta E, Galasso C, Manfredi MT, et al. Using beef-breed semen in seropositive dams for the control of bovine neosporosis. *Prev Vet Med.* 2018;161:127-33.
28. Regidor-Cerrillo J, Diez-Fuertes F, Garcia-Culebras A, Moore DP, Gonzalez-Warleta M, Cuevas C, et al. Genetic diversity and geographic population structure of bovine *Neospora caninum* determined by microsatellite genotyping analysis. *PLoS ONE.* 2013;8 8:e72678.
29. Cabrera A, Fresia P, Berna L, Silveira C, Macias-Rioseco M, Arevalo AP, et al. Isolation and molecular characterization of four novel *Neospora caninum* strains. *Parasitol Res.* 2019;118 12:3535-42.
30. Pedraza-Diaz S, Marugan-Hernandez V, Collantes-Fernandez E, Regidor-Cerrillo J, Rojo-Montejo S, Gomez-Bautista M, et al. Microsatellite markers for the molecular characterization of *Neospora caninum*: application to clinical samples. *Vet Parasitol.* 2009;166 1-2:38-46.
31. Regidor-Cerrillo J, Horcajo P, Ceglie L, Schiavon E, Ortega-Mora LM, Natale A. Genetic characterization of *Neospora caninum* from Northern Italian cattle reveals high diversity in European *N. caninum* populations. *Parasitol Res.* 2020.
32. Constantin EM, Schares G, Grossmann E, Sauter K, Romig T, Hartmann S. Untersuchungen zur Rolle des Rotfuchses (*Vulpes vulpes*) als möglicher Endwirt von *Neospora caninum*. *Berl Munch Tierarztl Wochenschr.* 2011;124 3-4:148-53.

33. Legnani S, Pantchev N, Forlani A, Zini E, Schares G, Balzer J, et al. Emergence of cutaneous neosporosis in a dog receiving immunosuppressive therapy: molecular identification and management. *Vet Dermatol*. 2016;27 1:49-e14.
34. Dubey JP, Hattel AL, Lindsay DS, Topper MJ. Neonatal *Neospora caninum* infection in dogs: isolation of the causative agent and experimental transmission. *J Am Vet Med Assoc*. 1988;193:1259-63.
35. Basso W, Schares S, Barwald A, Herrmann DC, Conraths FJ, Pantchev N, et al. Molecular comparison of *Neospora caninum* oocyst isolates from naturally infected dogs with cell culture-derived tachyzoites of the same isolates using nested polymerase chain reaction to amplify microsatellite markers. *Vet Parasitol*. 2009;160 1-2:43-50.
36. Dubey JP, Dorough KR, Jenkins MC, Liddell S, Speer CA, Kwok OCH, et al. Canine neosporosis: clinical signs, diagnosis, treatment and isolation of *Neospora caninum* in mice and cell culture. *Int J Parasitol*. 1998;28:1293-304.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57 1:289-300.
38. Kamvar ZN, Tabima JF, Grunwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*. 2014;2:e281.
39. Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics*. 2000;16 9:847-9.
40. Goudet J. HierFstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes*. 2005;5 1:184-6.
41. de Meeüs T, Goudet J. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infect Genet Evol*. 2007;7 6:731-5.
42. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol*. 2004;186 5:1518-30.
43. Mancini G, Gargani M, Chillemi G, Nicolazzi EL, Marsan PA, Valentini A, et al. Signatures of selection in five Italian cattle breeds detected by a 54K SNP panel. *Mol Biol Rep*. 2014;41 2:957-65.
44. Buchanan DS. Breeds of dairy cattle (major *Bos taurus* breeds). Reference Module in Food Science: Elsevier; 2016.
45. Schlötterer C, Tautz D. Slippage synthesis of simple sequence DNA. *Nucleic Acids Res*. 1992;20 2:211-5.
46. Viguera E, Canceill D, Ehrlich SD. Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J*. 2001;20 10:2587-95.
47. Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet*. 2000;16 12:551-8.
48. Schlötterer C. Evolutionary dynamics of microsatellite DNA. *Chromosoma*. 2000;109 6:365-71.
49. Estoup A, Jarne P, Cornuet JM. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol*. 2002;11 9:1591-604.
50. Bagshaw ATM. Functional Mechanisms of Microsatellite DNA in Eukaryotic Genomes. *Genome Biology and Evolution*. 2017;9 9:2428-43.
51. Mathema VB, Nakeesathit S, White NJ, Dondorp AM, Imwong M. Genome-wide microsatellite characteristics of five human *Plasmodium* species, focusing on *Plasmodium malariae* and *P. ovalecurtisi*. *Parasite*. 2020;27:34.
52. Dubey JP, Schares G. Neosporosis in animals-The last five years. *Vet Parasitol*. 2011;180 1-2:90-108.
53. Roman LRS, Horcajo P, Regidor-Cerrillo J, Fernandez-Escobar M, Collantes-Fernandez E, Gutierrez-Blazquez D, et al. Comparative tachyzoite proteome analyses among six *Neospora caninum* isolates with different virulence. *Int J Parasitol*. 2020;50 5:377-88.
54. Ajzenberg D, Collinet F, Aubert D, Villena I, Darde ML, French ToxoBs network g, et al. The rural-urban effect on spatial genetic structure of type II *Toxoplasma gondii* strains involved in human congenital toxoplasmosis, France, 2002-2009. *Infect Genet Evol*. 2015;36:511-6.

Tables

Table 1. *caninum* DNA content in samples of aborted fetuses and success in genotyping.

Typing success, # of loci typed	Ct value, Mean +/- Standard deviation	Minimum – Maximum Ct	P value, Wilcoxon Rank sum test to assess difference in Ct values
10 or 9	30.4 +/- 3.2	26.2 – 42.4	Reference
8 – 1	33.0 +/- 1.8	29.2 – 36.5	P = 0.00013

Table 2. Comparison of the MLMG patterns in dairy cattle farms in northern Italy, where more than one *Neospora caninum* positive foetus had been sampled and genotyped during the study period.

Farm #	Number of fetuses per farm (period)	Number of pairs of loci or sub-loci with differences of...				Differences/total number of comparisons (%)
		1 repeat	2 repeats	3 repeats	>3 repeats	
4	2 (2015)	4	-	-	-	3/14 (21.4)
6 ^a	4 (2015-16)	30	6	3	-	39/81 (48.1)
9	2 (2015)	1	1	-	-	2/3 (66.7)
10 ^a	2 (2016)	2	-	-	1	3/10 (30.0)
15 ^a	2 (2017)	2	-	-	1	3/11 (27.3)
17 ^a	3 (2017-18)	4	-	-	2	6/13 (46.2)
25	2 (2017-18)	2	1	-	-	3/14 (21.4)
35 ^a	2 (2018)	1	-	-	1	2/5 (40.0)

^a Table 3 provides details on microsatellite typing results for farms, where loci with 3 or > 3 repeats were observed.

Table 3. MLMG patterns on dairy cattle farms from northern Italy where more than one *Neospora caninum*-positive foetus had been sampled during the study period, which showed differences (bold numbers) in length for ≥ 3 repeats at microsatellite loci on various chromosomes (Chr2, 7a, 8, 9, 10, 12).

Farm	Date	Chr II	Chr VIIa			Chr VIII	Chr IX	Chr X	Chr X		Chr X
		MS2 , x-y-z, (AT) _x -TTGTATC-(AT) _y -GT-(AT) _z	MS1B , x=y+z, (AT) _y -AC-(AT) _z	MS7 , x, AT-AA-(TA) _x -GG	MS21 , x=y+4, TG-(TACA) ₃ -TACC-(TACA) _y TT	MS10 , x-y-z, AGT-(ACT) _x -(AGA) _y -(TGA) _z -CAA	MS5 , x, CG-(TA) _x -TGTAGG	MS6A , x, GC-(TA) _x -AC	MS6B , x, CC-(AT) _x -GT	MS12 , GC-(G)GC	
6	May 2015	6-8-2*	12	15	6	6-23-10	11	22	13	17	
	February 2016	7-10-2	12	15	6	6-25-10	11	25	13	16	
	April 2016	6-11-2	13	13	NA	5-25-10	11	20	14	16	
	June 2016	6-11-2	12	14	6	6-24-9	11	22	13	16	
10	January 2016	5-10-2	12	15	3	NA	11	21	13	16	
	April 2016	6-11-2	12	15	6	6-27-10	11	21	13	NA	
15	August 2017	6-11-2	12	15	6	6- 19 -10	NA	NA	NA	16	
	August 2017	6-11-2	12	14	6	6- 28 -9	11	18	13	16	
17	May 2017	6-10-2	NA	14	6	NA	11	NA	NA	17	
	November 2017	6-9-2	12	15	6	7-24-12	11	21	13	16	
	February 2018	NA	12	10	6	NA	NA	13	12	NA	
35	July 2018	6-11-2	12	15	6	NA	11	19	13	16	
	October 2018	NA	NA	15	6	NA	12	13	NA	NA	

Table 4. A linear regression models to characterize the association between genetic distance (Bruvos's genetic distance) of *N. caninum* isolates and the spatial distance of the sampling sites or the time between dates of sampling.

Model (adjusted R ²)	Variable	Estimate	Standard error	T value	Pr(> t)
1 (19.5%)	Intercept	0.2378241	0.0100212	23.732	<2e-16 ***
	Geographical distance (km)	0.0025708	0.0002615	9.831	<2e-16 ***
2 (19.8%)	Intercept	0.2502685	0.0143581	17.43	<2e-16 ***
	Geographical distance (km)	0.0026139	0.0002638	9.91	<2e-16 ***
	Time between dates of sampling	0.0013088	0.0010820	-1.21	0.227

Table 5. Genetic differences between northern Italian *N. caninum*. Pairwise population matrix of F_{ST} analysis values (F) and Nei's unbiased genetic distance (D). Values written in italic style indicate statistical significance based on of goodness-of-fit test (G-test) assessing the significance of the effect of region on genetic differentiation.

		F_{ST} analysis values (F_{ST})			
		Region	Piedmont	Lombardy	Veneto-Trento
Nei's genetic distance (D)	Piedmont	NA	<i>0.1559**</i>	0.0279	
	Lombardy	<i>0.1561**</i>	NA	<i>0.0973*</i>	
	Veneto-Trento	0.0331	<i>0.1082*</i>	NA	

* $P = 0.012$; ** $P = 0.001$

Figures

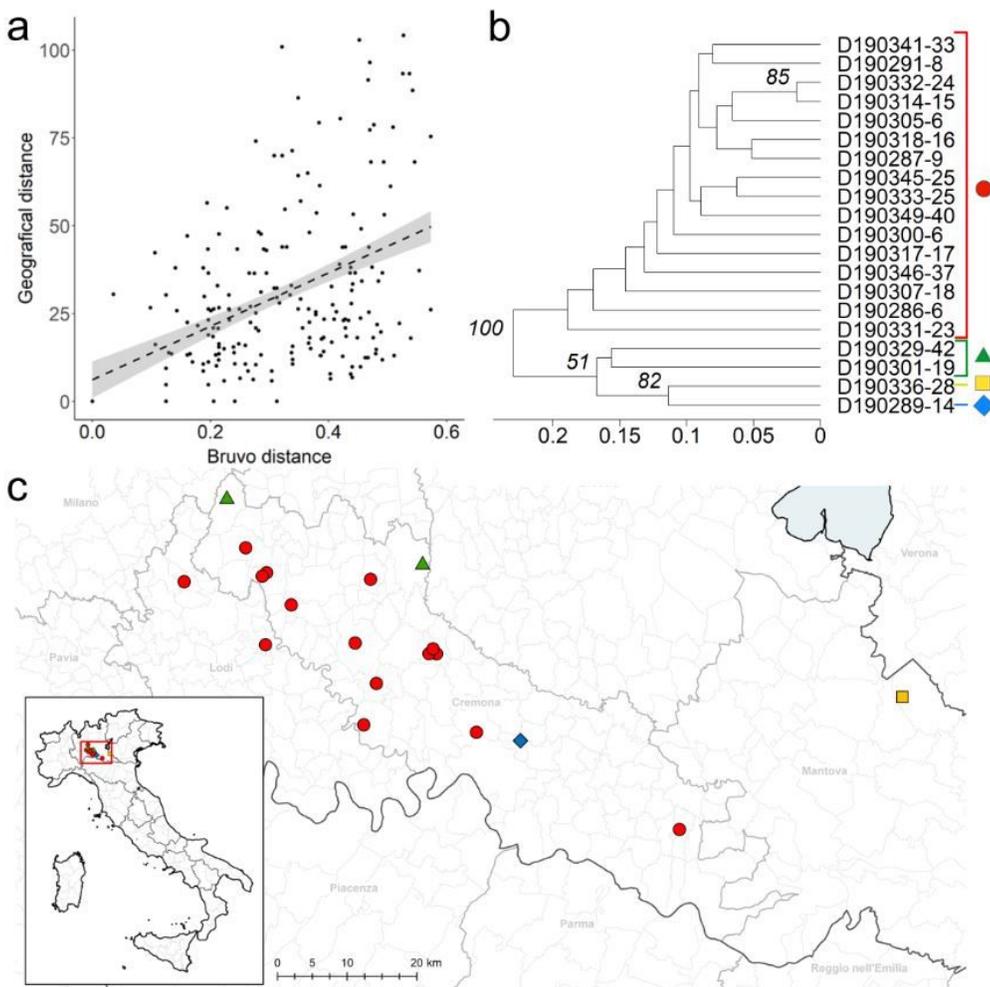


Figure 1

Genetic distance between *N. caninum* isolates from bovine foetuses collected in Lombardy, northern Italy, is associated with the geographical distance of the sampling sites. (a) Graphical representation of a linear regression model to associate Bruvo's genetic distance of *N. caninum* in bovine foetuses and spatial distance of sampling sites (km). (b) Dendrogram representing the results of K-means clustering using Bruvo's distance with non-parametric bootstrapping; the analysis identified at least four groups (different colours of sample designation, consisting of a string of sample ID, village name, and farm number), separated by bootstrap values > 50. (c) Map showing farm locations, using the same colours as for the groups in the dendrogram.

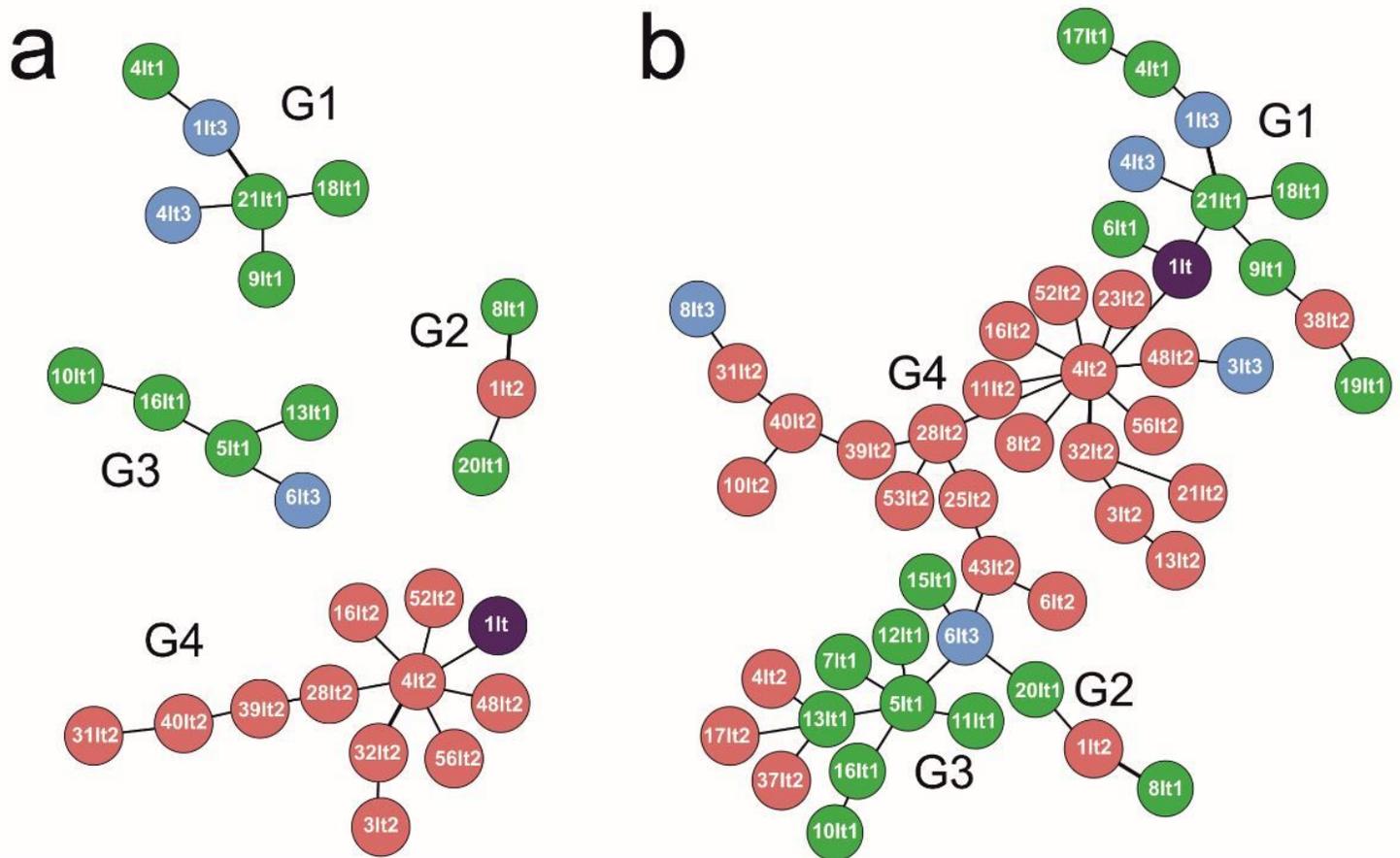


Figure 2

eBURST analysis of 49 northern Italian *Neospora caninum* samples. Specimens from Lombardy (red dots) tend to cluster separately from those from Piedmont (green dots) or Veneto-Trento (blue dots). (a) Analysis using the Double Locus Variant (eBURST DLV) option shows that *N. caninum* multilocus genotypes can be separated into four groups (G1-4). Genotyping of an Italian isolate, for which the province of origin was not known (violet dot) revealed results close to most isolates from Lombardy (group G4). (b) The full MST option in eBURST, including the complete data set (n=50 *N. caninum* samples) shows the location of the individual groups within a network of all samples. Note: The analysis was restricted to microsatellite markers available from this and a previous study [31]. Only samples that could be typed for all markers (i.e. MS5, 6A, 6B, 7, 10, 12, 21) were included. Moreover, MS10, which combines variation in three separate motifs (sub-loci), was analysed per each motif individually. Groups G1-G3 resemble the grouping reported in [31], while G4 represents a new group including 11 of 25 bovine *N. caninum* samples from Lombardy that were added to the analysis by this study.

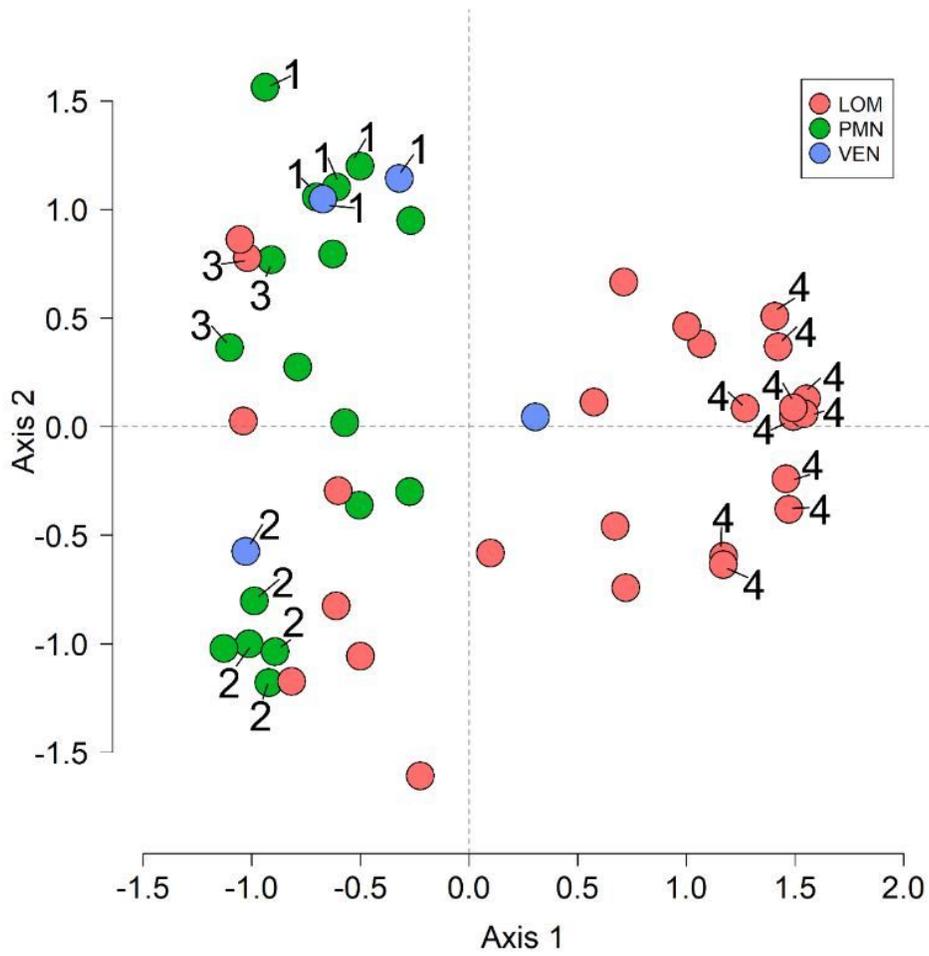


Figure 3
 Principle coordinate analysis (PCoA) separates north Italian *N. caninum* based on their MLMGs into clusters. Data from the present and a previous study conducted in cattle from northern Italy [31] were used for this analysis. The two dominant Eigenvalues of a PCoA based on the allele-sharing coefficient separated all *N. caninum* positive samples isolated from Piedmont and most from Veneto-Trento so that they clustered on the left side of the PCoA graph, while most of the samples from Lombardy were located on the right side of the PCoA graph. Groups G1-4 (represented by circles with numbers 1-4), as determined by eBURST, were located clearly in separate positions. While G1-3 was found on the left side, G4 (exclusively Lombardy samples) was localised on the right side of the PCoA graph.

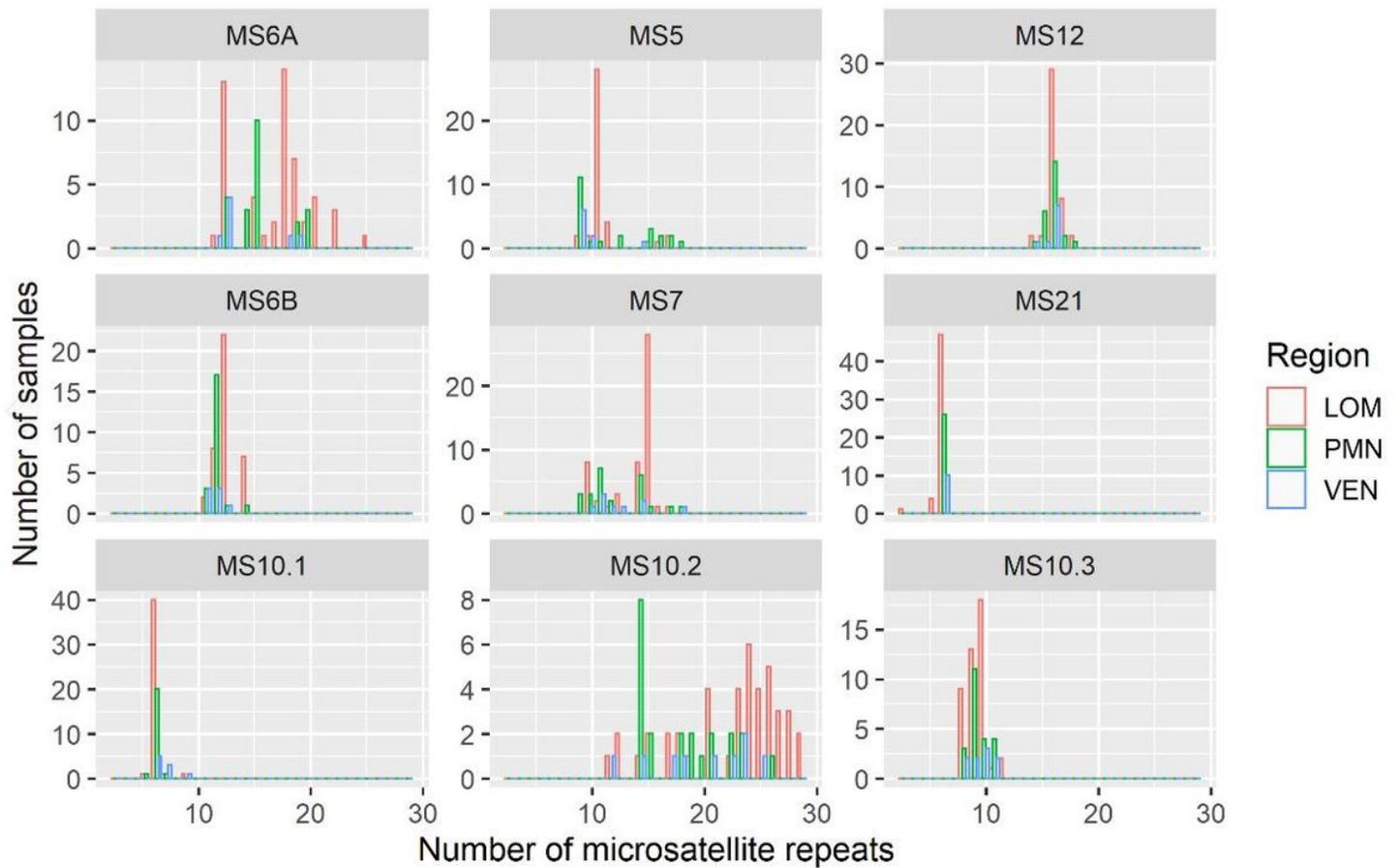


Figure 4

Differences in the heterogeneity of the numbers of repeat units characteristic for *Neospora caninum* microsatellite loci examined in the present and a previous study conducted in cattle from northern Italy [31], i.e. in the regions Lombardy (LOM), Piedmont (PMN), and Veneto-Trento (VEN). The range between the lowest and highest number of repeats was largest for MS10.2 (17 repeats), followed by MS6A (13), MS5 (9) and MS7 (9). For the remaining repeats, the range was 4 (MS12, MS21, MS10.1, MS10.3) or 3 repeats (MS6B). Pairwise Wilcoxon test (BH adjusted P value) revealed statistically significant differences in MS5 (LOM vs VEN), MS6B (LOM vs PMN, LOM vs VEN), MS7 (LOM vs PMN), MS10.1 (LOM vs VEN, PMN vs VEN) and MS10.2 (LOM vs PMN, LOM vs VEN).

Supplementary Files

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

- [SupplementarydataTableS1.xlsx](#)