Open access • Journal Article • DOI:10.1159/000506221

Identification of Some Errors in the Genome Assembly of Bovidae by FISH.

Lisa De Lorenzi, Pietro Parma

Published on: 02 Apr 2020 - Cytogenetic and Genome Research (Cytogenet Genome Res)

Topics: Genome, Sequence assembly and Human genome

Original Article

Cytogenetic and Genome Research

Cytogenet Genome Res
DOI: 10.1159/000506221

Accepted: January 3, 2020
Published online: ■■■

Identification of Some Errors in the Genome Assembly of Bovidae by FISH

Lisa De Lorenzi Pietro Parma

Department of Agricultural and Environmental Sciences (DiSAA), University of Milan, Milan, Italy

Keywords
Bovidae · Comparative cytogenetics · Evolution · FISH · Genome assembly

Abstract
From an economic point of view, Bovidae represent the most important family of the Ruminantia suborder. Thus, the mitochondrial and nuclear genomes of *Bos taurus* were among the first genomes to be sequenced after the sequencing of the human genomes. Over the millennia, the evolution of the genomes of the 3 main species belonging to the Bovidae family – *B. taurus* (BTA), *Ovis aries* (OAR), and *Capra hircus* (CHI) – has led to few chromosome rearrangements. Certainly, the availability and free access to the animal genomes significantly contributed to the improvement of animal genetics; however, some errors may exist due to the high automation in the genomic assembly construction process. In this work, some differences between the genomes of cattle, goat, and sheep highlighted by bioinformatics analysis have been verified by FISH, confirming that some errors persist even in the most recent genome assemblies. This type of approach has allowed us to detect a misassembly of a region belonging to BTA16 and to the homologues OAR12 and CHI16, a misassembly of a short tract in BTA22, OAR19, and CHI22, an incorrect mapping of a region of BTA21 and of CHI27 and OAR26, a discrepancy in the BTA26, OAR22, and CHI26 assemblies, a missed inversion in CHI1 compared to BTA1 and OAR1, and the exact assembly of a region of about 7 Mb in OAR10 and CHI12. Incorrect positioning of genomic tracts can cause unintended consequences in genetic analyses, especially when the data represent a starting point for the construction of genetic tools. In the new genomic assemblies published after the conclusion of our experiments, however, the accuracy in the construction of animal assemblies has been much improved, even if the new assemblies present more extended unmapped portions than the previous versions. The gap could be filled by comparative analyses between similar species or FISH.

Bovidae, from an economic point of view, represent the most important family among the 6 families belonging to the Ruminantia suborder and comprise more than 100 species [Chen et al., 2019]. For this reason, Bovidae, especially the species *Bos taurus*, have been the subject of many genetic and genomics studies. The mitochondrial [Anderson et al., 1982] and nuclear genomes [Bovine Genome Sequencing and Analysis Consortium, 2009] of *B. taurus*, in fact, were among the first genomes to be sequenced, only a few years after the sequencing of the 2 human genomes. After this first sequencing, both updates/improvements and independent sequencing fol-

KARGER

karger@karger.com
www.karger.com/cgr

Pietro Parma
Department of Agricultural and Environmental Sciences (DiSAA), University of Milan
Via G. Celoria, 2
IT–20133 Milan (Italy)
E-Mail pietro.parma@unimi.it

Table 1. Characteristics of the assemblies currently available regarding *Bos taurus*, *Ovis aries*, and *Capra hircus*

Species	Assembly	Release date	Submitter	Assembly content	Mapped base pairs	Unplaced base pairs	Coverage	Sequencing strategy	Reference
Cattle	UMD_3.1.1	11/25/2014	Center for Bioinformatics and Computational Biology, University of Maryland	Autosomes; X chromosome; unplaced scaffolds	2,660,906,405	9,216,905	9×	Sanger	Zimin et al., 2009
Cattle	Btau_5.0.1	11/19/2015	Cattle Genome Sequencing International Consortium	Autosomes; X chromosome; unplaced scaffolds	2,715,765,904	9,214,836	19×	Sanger; PacBio RS II	Bovine Genome Sequencing and Analysis Consortium, 2009
Cattle	ARS-UCD1.2	04/11/2018	USDA ARS	Autosomes; X chromosome; unplaced scaffolds	2,628,394,923	87,442,531	80×	PacBio; Illumina NextSeq 500; Illumina HiSeq; Illumina GAII	Unpublished
Sheep	Oar_v4.0	11/20/2015	International Sheep Genome Consortium	Autosomes; X chromosome; unplaced scaffolds	2,584,815,894	30,683,789	166×	Illumina GAII; 454; PacBio RSII	International Sheep Genomics Consortium, 2010
Sheep	Oar_rambouillet_v1.0	11/02/2017	Baylor College of Medicine Human Genome Sequencing Center	Autosomes; X chromosome; unplaced scaffolds	2,809,021,901	60,875,879	126×	HiSeq X Ten; PacBio RS II	Unpublished
Goat	CHIR_2.0	09/16/2015	Beijing Genomics Institute	Autosomes; X chromosome; unplaced scaffolds	2,685,484,791	123,933,671	175×	Illumina	Dong et al., 2013
Goat	ARS1	08/24/2016	USDA ARS	Autosomes; unplaced scaffolds	2,582,134,882	340,661,721	50×	PacBio	Unpublished
Goat	CVASU_BBG_1.0	03/22/2019	Bangladesh Goat Genome Consortium (BGGC)	Autosomes; unplaced scaffolds	2,791,592,313	250,031,347	14×	Illumina HiSeq	Unpublished

lowed. All these data are now freely available through various online web sites.

Concerning the cattle, goat, and sheep genomes, different research groups have produced assemblies that have been updated gradually. The characteristics of the assemblies currently available are shown in Table 1 (only the most recent update). The information necessary for carrying out this work was taken from the following assemblies: UMD_3.1.1 (cattle), Oar_v4.0 (sheep), and CHIR_2.0 (goat). The availability and assembly of genome sequences have proved to be a great help in animal selection. A selection can now be made on a genomic basis and not only on a morphological basis, as was the case in livestock farms until not long ago. However, due to the highly automated strategies used for the construction of genome assemblages, errors in the genome assembly may occur [De Lorenzi et al., 2010]. Furthermore, several discrepancies between different assemblies have already been highlighted [Partipilo et al., 2011]. Without diminishing the importance of the availability of genomic sequences in animal genetics research, an erroneous assembly can cause various problems such as an incorrect linkage between molecular markers. Considering the 3 main species (from an economic point of view) of the Bovidae family, cattle (*B. taurus,* BTA), sheep (*Ovis aries*, OAR), and goat (*Capra hircus*, CHI), we can assert that their genomic structure is highly conserved, showing a strong degree of similarity. Cattle and goats share the same diploid

Table 2. Genomic characteristics of the BACs used in the experiments

Case	BAC name	Library	GenBank accession numbers for BAC end sequences	BAC length[a], kb	Genome position[b]		
					cattle[c]	sheep[d]	goat[e]
1	182I21	CH240	CL605447/CL605383	136.5	BTA16: 49,510,112	OAR12: 78,760,150	Un:scaffold2905[f]
1	97H07	CH240	BZ956267/BZ956180	136.5	BTA16: 50,120,317	OAR12: 46,878,591	CHI16: 46,405,859
2	364O23	CH240	CC517840/CC517750	184.7	BTA22: 14,409,935	OAR19: 15,479,000	CHI22: 15,495,247
2	203I12	CH240	BZ866440/BZ866405	225.7	BTA22: 15,155,470	OAR19: 14,484,823	CHI22: 14,328,276
3	53D03	INRA	CR841122/CR841121	141.5	BTA21: 59,936,599	OAR26: 628,140	CHI27: 807,842
3	955F02	INRA	CR845072/CR845071	Unknown[g]	BTA21: 60,247,362	OAR26: 441,555	CHI27: 628,278
3	312A06[h]	INRA	DEFB1	Unknown[i]	BTA27: 5,788,753[j]	OAR26: 5,696,505[j]	CHI27: 5,867,577[j]
4	783G01[k]	INRA	RB1	Unknown[i]	BTA12: 18,196,641[j]	OAR10: 18,439,414[j]	CHI12: 16,026,841[j]
4	391O02	CH240	CC591234/CC591147	146.3	BTA26: 23,062,932	OAR22: 49,128,581	CHI26: 48,660,131
4	239G20	CH240	BZ843545/BZ843543	148.5	BTA26: 25,064,333	OAR22: 23,843,622	CHI26: 23,089,034
5	668D03	INRA	CR816168/CR816167	135.7	BTA1: 22,832,127	OAR1: 141,557,009	CHI1: 23,766,882
5	535G01	INRA	CR807858/CR807859	109.1	BTA1: 24,648,163	OAR1: 143,396,027	CHI1: 21,970,551
6	9C04	CH240	BZ896667/BZ896585	211.6	BTA12: 69,602,179	OAR10: 69,522,093	CHI12: 66,688,682
6	519L12	CH240	CZ022887/CZ022695	195.2	BTA12: 72,374,129	OAR10: 71,056,324	CHI12: 67,626,687
6	368K21	CH240	CC520755/CC520663	114.9	BTA12: 75,120,432	OAR10: 71,101,723	Un: scaffold2040
6	379022	CH240	CC583411/CC583319	167.3	BTA12: 76,972,153	OAR10: 72,376,726	CHI12: 68,434,055

[a] The sizes of the BACs refer to the bovine genome. [b] The initial position of the most centromeric gene is reported. [c] UMD_3.1.1 genome assembly. [d] OviAri_4 sheep genome assembly. [e] CHIR_1.0 goat genome assembly. [f] Un, unplaced scaffold. [g] From available bioinformatics data it was not possible to know the BAC size. [h] Reference marker for BTA27 according to Gautier et al. [2001], related to the defensin beta 1 gene (*DEFB1*). [i] The ends of these BACs were not sequenced; thus it was not possible to know the BAC size. [j] CDS starting position of the reference gene. [k] Reference marker for BTA12 according to Gautier et al. [2001], related to the retinoblastoma 1 gene (*RB1*).

number (2n = 60) with structurally similar autosomes, except for a small well-known translocation involving BTA9/CHI14, while sheep have a lower diploid number (2n = 54) than the 2 other species, due to the presence of 3 centric fusions that occurred during species evolution [Iannuzzi et al., 2009]. Several years ago, our laboratory developed a bioinformatics approach combining physical mapping of markers by fluorescence in situ hybridization (FISH) with informatics data which allowed detecting both small evolutionary divergences and errors in genomic assemblies [De Lorenzi et al., 2015].

In the present work, some differences between the cattle, goat, and sheep genomes highlighted by bioinformatics analysis have been verified by FISH, confirming that even the most recent assemblies contain errors.

Materials and Methods

Bioinformatics Analysis

Briefly, end sequences from the INRA Bt BAC library [Eggen et al., 2001] were used as e-probes to compare the cattle, sheep, and goat genomes. The genomic assemblies that were considered were UMD3.1.1 (cattle), CHIR_1.0 (goat), and OviAri4 (sheep). These probes underwent stringent quality control in order to eliminate sequences that might produce incorrect results. The localization of e-probes on genomes was performed using the BLAST-like Alignment Tool (BLAT) software [Kent, 2002]. A detailed description of the production of the e-probes and their localization on genomes is given in De Lorenzi et al. [2015].

FISH Analysis

BACs from 2 libraries were used for the FISH experiments: INRA Bt [Eggen et al., 2001] and the CH-240 Library (constructed by C.L. Shu and K. Osoegawa from CHORI). The characteristics of the BACs and their location in the various genomes under consideration are reported in Table 2 and online supplementary Table 1 (for all online suppl. material, see www.karger.com/doi/10.1159/000506221). The BACs were grown overnight at 37°C in Luria Broth supplemented with chloramphenicol, and DNA was extracted using the PhasePrep™ BAC DNA Kit (Sigma Aldrich), following the manufacturer's instructions. Next, 250 ng of DNA was labeled with biotin 16-dUTP or Cy3-dUTP using a Nick Translation Kit (Merck). Overnight hybridization was performed as reported in De Lorenzi et al. [2014]. The images were taken using an epifluorescence microscope equipped with a charge-coupled device camera. Avidin-FITC (to detect biotin-labeled DNA) signals, Cy3 (direct labeling) signals, and DAPI fluorescence of the chromosomes were detected using specific filters. The images were

recorded and merged using the QFISH software (Leica Microsystems, Wetzlar, Germany).

Cytogenetic Localization

In order to provide a cytogenetic localization of the BACs used in the FISH experiments, we compared the hybridization signal images to a reference ideogram of the chromosomes [Cribiu et al., 2001], as previously proposed by Iannuzzi et al. [2015].

Results

Case 1. Discrepancies in the BTA16, OAR12, and CHI16 Assemblies

From an evolutionary point of view, BTA16 and OAR12 are homologous, so we expected that the order of the markers from the centromere to the telomere of the 2 chromosomes would be respected. From the bioinformatics analysis, it emerged instead that a region of about 210 kb, characterized by the presence of 9 BACs, is positioned in the sheep at about 78.8 Mb (almost terminal on OAR12) instead of around 49.5 Mb as expected from the bovine database. Considering the goat database, it was not possible to determine the position of the BACs used as probes in FISH experiments, as they were not mapped. The data obtained from the databases were verified by physical mapping of 2 bovine BACs: 182I21, which in sheep should be located in an almost terminal region of OAR12, and 97H07, which maps downstream of the misassembled region (Fig. 1a). The results obtained by FISH, performed in all 3 species, confirmed there was actually a misassembly of the fragment in the sheep genome, but also in the cattle and the goat genome, thus highlighting an assembly error in the bovine assembly. BAC 182I21 is always positioned in a terminal region in BTA16, OAR12, and CHI16 (Fig. 2a–c).

Case 2. Discrepancies in the BTA22, OAR19, and CHI22 Assemblies

In this case, the discrepancy found between the bovine assembly and the sheep and goat assemblies comprised a rather small misassembled region involving BTA22, corresponding to CHI22 and OAR19. It amounted to about 0.7 Mb. The sheep genome showed the same assembly as the goat genome (not shown). Two BACs (364O23 and 203I12; Fig. 1b) were used in FISH experiments for the physical mapping of the genomic region involved. The results revealed the more centromeric position of BAC 203I12 compared to 364O23 in all 3 species. The results confirm the error in the assembly of the bovine genome compared to the sheep and goat genomes (Fig. 2d–f).

Case 3. Discrepancies in the BTA21, OAR26, and CHI27 Assemblies

A genomic region of about 400 kb has a discordant location in the 3 species. In bovines, it is assembled near the telomere of BTA21, while in the sheep and goat genomes it is located on homologous chromosomes, OAR26 and CHI27, near the centromere (Fig. 1c). The exact position was verified using the BACs 53D03 and 955F02, both included in this region. Surprisingly, the physical mapping showed their location on 2 different chromosomes: BAC 53D03 maps close to the centromere of BTA27 (homologous to OAR26 and CHI27), whereas BAC 955F02 maps close to the centromere of BTA12. The correct chromosome was identified using reference BACs of BTA27 and BTA12, according to Gautier et al. [2001] (Fig. 2g–j). In comparing the databases of the 3 species, it emerged that a short region of about 400 kb (including BACs 53D03 and 955F02) is not correctly assembled. Particularly, in the bovine genome, it is assembled in a telomeric region of BTA21, while in goat and sheep it is assembled in a centromeric region of CHI27 and OAR26, respectively. The regions containing the 2 BACs (53D03 and 955F02) which we used as markers in FISH experiments to physically map the region of interest are represented in purple in Figure 1c. The physical mapping, instead, pointed out that the 2 markers do not belong to the same region and that, in the bovine assembly, neither of them maps on BTA21. FISH demonstrated that BAC 53D03 hybridizes in the centromeric region of BTA27 (corresponding to CHI27 and OAR26), while BAC 955F02 maps to a subcentromeric region of BTA12, CHI12, and OAR10. Regarding BAC 53D03, we can conclude that the bovine database is incorrect, while the goat and sheep databases are correct. For 955F02, all 3 databases are incorrect. For the exact identification of chromosomes, we used double FISH, combining the markers with the reference BACs of BTA27 (BAC 312A06; Fig. 2h–j) and BTA12 (BAC 783G01; Fig. 2g), according to Gautier et al. [2001]. Images of double FISH performed with BACs 955F02 and 783G01 on goat and sheep chromosomes are not shown.

Case 4. Discrepancies in the BTA26, OAR22, and CHI26 Assemblies

This situation is similar to those reported in cases 1 and 2. A small region (of about 160 kb) maps on BTA26 at about 23 Mb, whereas the same region seems to map on CHI26 at 48 Mb (Fig. 1d) and on OAR22 (homologous to BTA26 and CHI26) at 49 Mb (data not shown). Double FISH analysis using BACs 391O02, belonging to the hy-

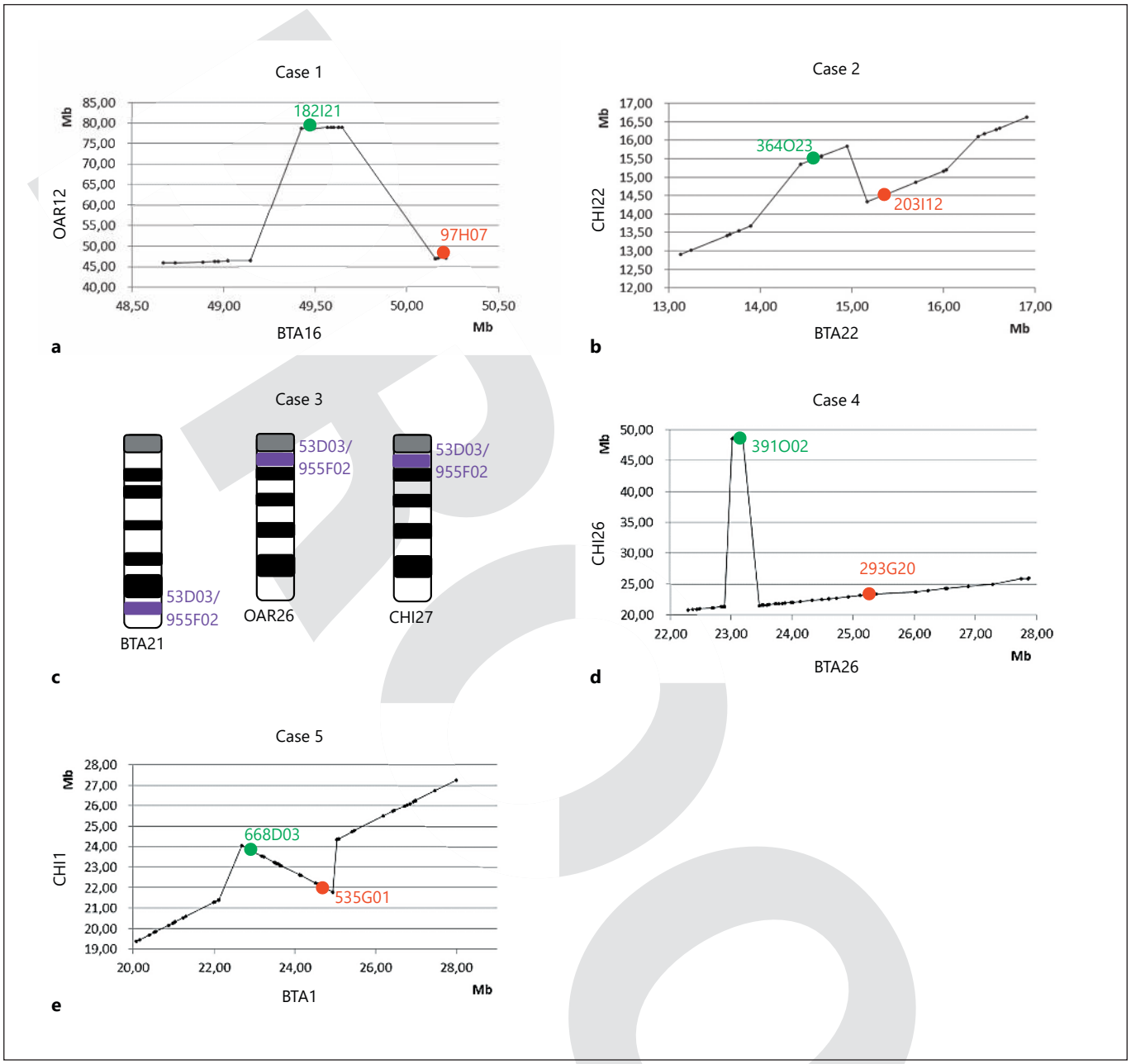

Fig. 1. a–e Genomic position of BACs as highlighted by bioinformatics. **a** BACs 182I21 and 97H07 in OAR12 compared to the bovine homologue BTA16. **b** BACs 364O23 and 203I12 in CHI22 compared to BTA22. **c** Genomic location of the region including the BACs 53D03 and 955F02. The region is misassembled in a telomeric region of BTA21 and in a centromeric region of CHI27 and its homologue OAR26. **d** BACs 391O02 and 293G20 in CHI26 compared to BTA26. **e** Representation of an inverted fragment, including BACs 668D03 and 535G01, on CHI1 compared to BTA1. In **a**, **b**, **d**, and **e**, the BACs are labeled in green and red in accordance to their signal colors in FISH (green, probe labeled with biotin and detected with avidin-FITC; red, probe directly labeled with Cy3). The dots along the lines represent the probes available from which we have chosen the markers used in the FISH experiments.

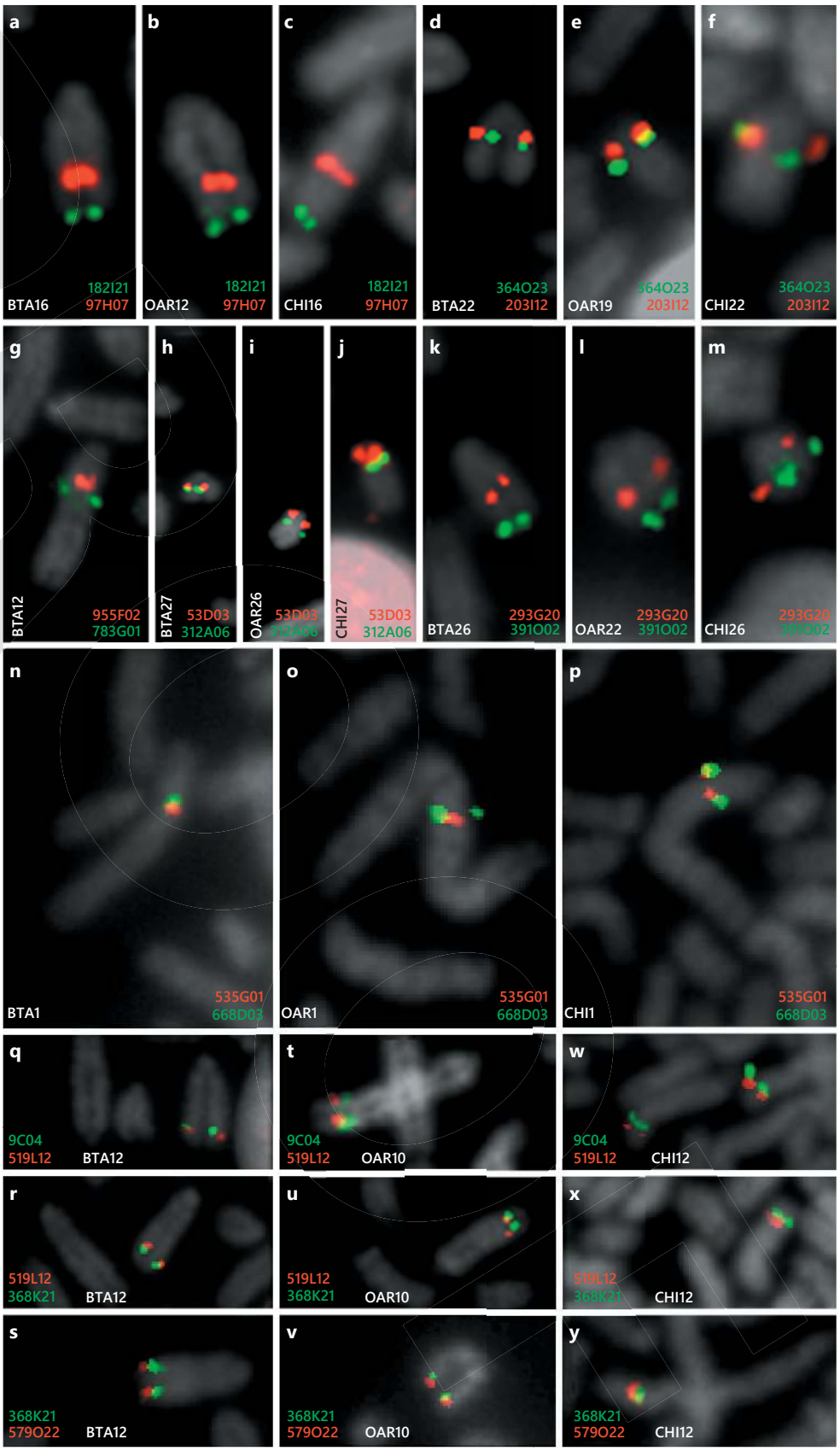

Fig. 2. Physical mapping of different BACs on bovine (BTA), sheep (OAR), and goat (CHI) chromosomes. **a–c** BACs 182I21 (green signals in an almost terminal region of the homologous chromosomes) and BAC 97H07 (red signals). **d–f** BACs 364O23 (green signals) and 203I12 (red signals). In all 3 species, the more centromeric position of BAC 203I12 compared to 364O23 is evident. **g** BAC 955F02 (red signals) in combination with the reference BAC of BTA12 (783G01, green signals). BAC 955F02 maps close to centromere. **h–j** BAC 53D03 (red signals) coupled with the reference BAC of BTA27 (312A06, green signals). BAC 53D03 maps close to centromere of the homologous chromosomes. **k–m** BACs 391O02 (green signals) and 293G20 (red signals). In all 3 species the telomeric position of BAC 391O02 is shown. **n–p** BACs 668D03 (green signals) and 535G01 (red signals). In all 3 species the more centromeric position of BAC 668D03 compared to 535G01 is visible. **q–y** BACs 9C04 and 368K21 (green signals) and BACs 519L12 and 579O22 (red signals). For FISH experiments the BACs were coupled 2 by 2 to verify the exact order of the markers, which is conserved for all 3 species.

Table 3. Cytogenetic localization of the BACs used in the FISH experiments

BAC name	Cattle/goat	Sheep
182I21	16q26	12q25
97H7	16q22	12q17-21
203I12	22q13	19q13
364O23	22q13-21	19q13
53D03	27q13	26q14
312A06[a]	27q13-14	26q14-15
955F02	12q13	10q13
783G01[a]	12q13	10q13
391O02	26q23	22q23-24
293G20	26q21	22q21
668D03	1q21	1q12-13
535G01	1q21	1q13
9C04	12q23	10q23
519L12	12q23-24	10q24
368K21	12q24	10q24
379022	12q25	10q24-25

[a] Localization in cattle already reported by Gautier et al. [2001].

pothetical misassembled region, and 293G20, located downstream of the misassembly, confirmed the existence of a misassembly in both sheep and goat, but also in cattle, highlighting an error in the bovine genome assembly (Fig. 2k–m).

Case 5. Discrepancies in the BTA1, OAR1, and CHI1 Assemblies

Here, a large genomic portion (of about 2.3 Mb), comprising 24 consecutive BACs and positioned at about 24 Mb on BTA1, appears to be inverted in the goat genome compared to the cattle and sheep genomes (Fig. 1e). FISH experiments performed using the BACs 668D03 and 535G01, included in the inverted region at the opposite poles, revealed that the hypothetical inversion was an artifact of genomic assembly. The physical mapping showed the same orientation of the genetic portion in all 3 species; this orientation agrees with the cattle and sheep databases. In this case, the error lies in the goat genome assembly (Fig. 2n–p).

Case 6. Discrepancies in the BTA12, OAR10, and CHI12 Assemblies

In the last case, we considered a genomic region of about 8 Mb, located on BTA12 at about 75 Mb, and incorrectly mapped in the sheep and goat genomes (online suppl. Table 2). In order to map this large genomic region, 4 different BACs were used: 9C04, 519L12, 368K21, and 579O22. The first and last BAC show the correct mapping (i.e., mapping on homologous chromosomes in the 3 species), while the other 2 show a non-coherent mapping. The results of physical mapping have shown that even the region comprising the 2 external BACs presents consistent mapping; in fact, all BACs map on BTA12, OAR10, and CHI12, respectively, preserving order and marker distances (Fig. 2q–y).

Cytogenetic Localization

To complete the work, we also propose a cytogenetic localization of the BACs used in FISH experiments. The assignment of the cytogenetic band in the 3 species considered is shown in Table 3.

Discussion

Incorrect positioning of genomic tracts in databases can cause unintended consequences in genetic analysis, especially when the assembly data are used as a starting point for the construction of a genetic tool used by researchers around the world. If we take Case 1 as an example, a region of about 250 kb (BTA16:49,423,434–49,648,057) has been arranged in a wrong position in the bovine genome (assembly UMD3.1.1), since the physical mapping by FISH analysis clearly showed that it is placed in a telomeric region. As this assembly was used for the construction of the mostly employed single-nucleotide polymorphism (SNP) arrays, the logical consequence is that a certain series of SNPs do not map where they really are. In one of the most used HD SNP arrays, the aforementioned region includes 6 SNPs (BovineHD1600013730, BovineHD1600013736, ARS-BFGL-BAC-32069, ARS-BFGL-NGS-76897, BovineHD1600013759, and BTA-119822-no-rs). Therefore, during the mapping of elements concerning phenotypic aspects, any association involving these SNPs should consider not only the genes included in this genomic portion, but also the genes positioned in a much more telomeric position. The negative consequence is even clearer in Case 3, in which the genomic portion is placed on another chromosome. The bioinformatics analysis that highlighted these anomalies took into consideration the genomes available at the time of the work: UMD3.1.1 (bosTau8) for cattle, made available in January 2015, CHIR_1.0 for goat, released in September 2015, and finally OviAri4 for sheep, made available in

December 2015. Subsequently, new independently produced assemblies were published for all 3 species. In April 2018, the USDA ARS released the cattle ARS-UCD1.2 genome assembly, in November 2019, the Baylor College of Medicine released the sheep Oar_rambouillet_v1.0 genome assembly, and in August 2016, the USDA ARS also released the goat ARS1 genome assembly. In 4 of the 6 analyzed regions, we found an assembly error in the cattle UMD3.1.1 genome, and we confirmed the correct assembly for the sheep and goat genomes. In reconsidering the positions of the BACs used for the present work in the newly published assemblies of the 3 species (online suppl. Table 1), we can affirm that many of the discrepancies observed between the old databases and physical mapping of the markers have been corrected and are currently in agreement, demonstrating the greater accuracy of the data provided by the latest assemblies. Better accuracy is probably the result of using better sequencing technologies such as the PacBio sequencing strategy, which allows the production of longer reading strings. The data recorded in this work, in fact, certify the improvement work done in the last genomic assemblies. There is a further consideration, however, regarding the assemblies of the goat genome: CHIR_1.0 (from the International Goat Genome Consortium) and Goat ARS1 (from USDA ARS). BAC 53D03 is correctly localized in the new bovine genome assembly (ARS-UCD1.2) and in the 2 sheep genomes, old and new (oviAri4 and Oar_rambouillet_v1.0, respectively), while it is incorrectly localized in the new goat genome ARS1 assembly. It is mapped, in fact, in a telomeric region. Regarding BAC 955F02, we can state that it is incorrectly mapped not only in the old version of the genome assembly but also in the most recent ones. Traditionally, in a genomic assembly, the base pair number 1 is located in a centromeric region; but in the case of ARS1, it is reported in an inverted manner, so it is located in the telomeric region. Subsequent analyses have verified this, and we can state that while 16 pairs of chromosomes (numbers 1, 5, 6, 8, 9, 11, 13, 16, 18, 19, 20, 21, 22, 24, 25, and 29) map in a comparable direction in the 2 assemblies, the remaining 13 pairs are inverted. However, despite the proven higher accuracy of the most recent assemblies, they present more extended unmapped portions than the previous versions (87 Mb vs. 9 Mb for the cattle genome, 60 vs. 30 Mb for the sheep genome, and 340 vs. 111 Mb for the goat genome). The remaining unplaced fragments could be mapped in 2 ways: by comparative analysis between similar species or by FISH [De Lorenzi et al., 2013].

Conclusion

Bioinformatics analysis has previously shown an evolutionary rearrangement between the cattle and the goat assemblies [De Lorenzi et al., 2015]. The physical mapping (FISH) carried out during our tests was a fundamental tool to unveil some errors in the UMD3.1.1 bovine assembly which has often been used as a basis for producing tools dedicated to genetic analysis. In particular, this approach allowed us to detect assembly discrepancies in the following chromosomes: BTA16, BTA22, BTA21, CHI27, OAR26, BTA26, CHI1, OAR10, and CHI12. Our work once again underlines how FISH technology can be of great assistance in genomic analysis at a time when sequencing seems to be the ultimate type of analysis.

Acknowledgement

We are grateful to the ANABIC breeders' associations that required cytogenetic analysis and provided material that could also be used in research.

Statement of Ethics

During the researches no animals were used. The cytological preparations used in FISH experiments were already available in our laboratory for routine analyzes previously authorized and carried out.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Sources

This work was funded with private funds from the authors. L.D.L. is supported with research grant from Italian Ministry of Education, University and Research.

Author Contributions

L.D.L performed FISH analyses. P.P. analyzed the bioinformatics data. Both authors chose and selected the BACs used as probes and wrote the paper.

References

Anderson S, de Bruijn MH, Coulson AR, Eperon IC, Sanger F, Young IG: Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J Mol Biol 156:683–717 (1982).

Bovine Genome Sequencing and Analysis Consortium: The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528 (2009).

Chen L, Qiu Q, Jiang Y, Wang Y, Lin Z, et al: Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science 364:eaav6202 (2019).

Cribiu EP, Di Berardino D, Di Meo GP, Eggen A, Gallagher DS, et al: International System for Chromosome Nomenclature of Domestic Bovids (ISCNDB 2000). Cytogenet Cell Genet 92:283–299 (2001).

De Lorenzi L, Molteni L, Parma P: FISH mapping in cattle (*Bos taurus* L.) is not yet out of fashion. J Appl Genet 51:497–499 (2010).

De Lorenzi L, Genualdo V, Perucatti A, Iannuzzi A, Iannuzzi L, Parma P: Physical mapping of 20 unmapped fragments of the btau_4.0 genome assembly in cattle, sheep and river buffalo. Cytogenet Genome Res 140:29–35 (2013).

De Lorenzi L, Rossi E, Gimelli S, Parma P: De novo reciprocal translocation t(5;6)(q13;q34) in cattle: cytogenetic and molecular characterization. Cytogenet Genome Res 142:95–100 (2014).

De Lorenzi L, Planas J, Rossi E, Malagutti L, Parma P: New cryptic karyotypic differences between cattle (*Bos taurus*) and goat (*Capra hircus*). Chromosome Res 23:225–235 (2015).

Dong Y, Xie M, Jiang Y, Xiao N, Du X, et al: Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). Nat Biotechnol 31:135–141 (2013).

Eggen A, Gautier M, Billaut A, Petit E, Hayes H, et al: Construction and characterization of a bovine BAC library with four genome-equivalent coverage. Genet Sel Evol 33:543–548 (2001).

Gautier M, Laurent P, Hayes H, Eggen A: Development and assignment of bovine-specific PCR systems for the Texas nomenclature marker genes and isolation of homologous BAC probes. Genet Sel Evol 33:191–200 (2001).

Iannuzzi A, Perucatti A, Pauciullo A, Genualdo V, De Lorenzi L, et al: Fluorescent in situ hybridization mapping of three fecundity genes on cattle, river buffalo, sheep and goat. Caryologia 68:9–12 (2015).

Iannuzzi L, King WA, Di Berardino D: Chromosome evolution in domestic bovids as revealed by chromosome banding and FISH-mapping techniques. Cytogenet Genome Res 126:49–62 (2009).

International Sheep Genomics Consortium: The sheep genome reference sequence: a work in progress. Anim Genet 41:449–453 (2010).

Kent WJ: BLAT – the BLAST-like alignment tool. Genome Res 12:656–664 (2002).

Partipilo G, D'Addabbo P, Lacalandra GM, Liu GE, Rocchi M: Refinement of *Bos taurus* sequence assembly based on BAC-FISH experiments. BMC Genomics 12:639 (2011).

Zhang Z, Schwartz S, Wagner L, Miller W: A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214 (2000).

Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, et al: A whole-genome assembly of the domestic cow, *Bos taurus*. Genome Biol 10:R42 (2009).