SUPPLEMENTARY MATERIAL

Cold agglutinin-associated B-cell lymphoproliferative disease shows highly recurrent gains of chromosome 3 and 12 or 18

Agnieszka Małecka, Jan Delabie, Ingunn Østlie, Anne Tierens, Ulla Randen, Sigbjørn Berentsen, Geir E. Tjønnfjord and Gunhild Trøen

Supplementary Methods

Material and DNA extraction

Primary CAD patient samples were collected during the CAD5 study and detailed clinical characteristics are reported previously.[1] Most of the patients in the current study were not treated (9 patients) prior to procurement of bone marrow samples, while 3 patients were treated once and 1 patient was treated twice. In addition we present data from two samples published previously.[2] Mononuclear cells were purified from fresh bone marrow and/or blood as previously described.[3] Clonal B cells and T cells (normal

control) were acquired using fluorescence-activated cell sorting[2,3] DNA was purified using AllPrep DNA/RNA Micro Kit (Qiagen) according to the manufacturer recommendations.

Exome sequencing and bioinformatic analysis

Samples CAD-5, CAD-7 and CAD-1.06, CAD-1.07, CAD-1.22, CAD-1.25, CAD-1.31 (previously marked as CAD-1, CAD-2, CAD-19, CAD-20, CAD-22 respectively) were exome sequenced as previously described.[2] The additional 8 samples were exome sequenced at Helse Sør-Øst Genomics Core Facility at Oslo University Hospital by using the Twist Core Exome enrichment system and the Illumina NextSeq instrument with High Output flow cell, generating an average coverage of 60X.

The resulting reads were aligned to the hg38 reference genome with BWA[4]. Post-processing involved Picard (https://broadinstitute.github.io/picard/) and GATK[5-7] tools and consisted of quality score recalibration, realignment around indels and marking of duplicates. Finally, copy number variants were analyzed using GATK4 (DenoiseReadCounts) (Supplemental Figure 1B) and Control-FREEC[8](Supplemental Figure 2). All CAD samples from previous study[2] were reanalyzed as described above in order to compare all the samples. Analysis of small copy number changes was not possible for two samples: for one sample (CAD-1.06) only targeted sequencing was performed and for the other sample (CAD-1.07) exome sequencing was performed using whole genome amplified material. The main advantage of the Twist Core Exome library kit for exome sequencing is the possibility to use limited amounts of starting material (DNA input <50 ng). However, the disadvantage of using a low DNA input is uneven coverage resulting in less reliable results for small copy number changes. Large changes are clearly detected, however smaller changes are only reported if also detected by cytogenetic microarrays.

Cytogenetic microarrays - OncoScan CNV Plus Assay

Analysis was performed using cytogenetic microarrays, OncoScan CNV Plus Assay (Thermo Fisher Scientific), following the manufacturer recommendations. Arrays were scanned on the GeneChip Skanner 300 7G and array fluorescence intensity (CEL) files were automatically generated by the Affymetrix GeneChip Command Console (AGCC) Software. The CEL-files were further processes using the Chromosome Analysis Suite (ChAS 4.0) software using human genome version hg38 to produce OSCHP files.

The main challenge, as expected with rare disease, was very limited amount of material available for analysis. For 6/13 samples the DNA input was below recommended amount. Therefore, based on our experience with limited DNA and our previous projects, we applied a 3 Mb cutoff for reporting gains and losses from cytogenetic microarrays.

In accordance with Norwegian legislation and the etic approval of the study, all sensitive data are stored in protective systems at Oslo University Hospital. On request the data will be made available for other institutions. For the original data please contact: gto@ous-hf.no

Comparison of cytogenetic microarrays and exome sequencing CNV analysis.

Table 1 show both major chromosomal changes and smaller changes (> 3 Mb) detected by both methods. Changes determined by cytogenetic microarrays smaller than the 3 Mb cutoffs are included only if also detected by exome sequencing CNV analysis.

Statistical analysis

Correlation analysis was applied to examine significance of response to therapy vs presence or absence of trisomy 12 or 18. The categorical samples were coded as indicated: no trisomy 12 or 18 as 0, trisomy 12 as 1 and trisomy 18 as 2. For sample CAD-1.25 we assigned value 1 since it has some minor CNVs in both chromosome 12 and 18 (Supplementary Table 1). In this model we performed Pearson correlation and calculated P value for all 13 samples.

Correlation with response to therapy was calculated for all recurrent CNVs. In addition the correlation of recurrent CNVs with *KMT2D* and *CARD11* mutations (unpublished data) was calculated. Due to the small sample size we report only the results that are significant at $p < 0.01$.

Comparison of recurrent chromosomal changes with published data

Recurrent large CNVs detected in CAD that are also frequently found in hematological malignancies are discussed in main text. However, small regions of recurrent gains or losses detected (+1p36.31-p36.13, -8p21.3-p21.1, +9q34.2-q34.3, +11q13.1-q13.3, +17q25.1-q25.3, +21q22.2-q22.3, +22q13.31-q13.33) are rarely found in other hematological malignancies.[9-14] Nevertheless, some of these changes are found to some extend in other diseases. Deletion of 8p was found by Rinaldi et al.[9] in splenic marginal zone lymphoma (SMZL) and it is consistent with -8p21.3-p21.1 in CAD, moreover this deletion was also detected in exome sequencing data (Table 1). Amplification of 17q was also found in some cases of SMZL and follicular lymphoma.[9,14] Minimal common deleted and gained regions identified in >10% of cases of marginal zone and lymphoplasmacytic lymphomas [12] do not overlap with small

regions of recurrent gains or losses detected in CAD, with the exception of gain of 9q34.2-q34.3 that was also detected in 15% of cases of lymphoplasmacytic lymphoma.

Supplementary Table and Figures

Supplemental Table 1. Copy number variations (larger than 3Mb) in CAD patient samples detected by cytogenetic microarray assay. Changes in chromosomes 3, 12 and 18 presented in Table 1 in the main text are excluded.

	Minor recurrent changes (>2 samples)													
Samples	**chr 1 gains**	**chr 4 gains**	**chr 8 gains**	**chr 8 losses**	**chr9 gains**	**chr 11 gains**	**chr 14 losses**	**chr 15 losses**	**chr 16 gains**	**chr 17 gains**	**chr 20 gains**	**chr 20 losses**	**chr 21 gains**	**chr 22 gains**
CAD-1.06	1p36.33-p36.12				9q33.3-q34.3		14q32.2-q32.33**			17q24.3-q25.3				
CAD-1.07	1p36.31-p36.13	4p16.3-p16.1			9q34.11-q34.3				16q24.1-q24.3				21q22.2-q22.3	22q13.2-q13.31
CAD-1.22					9q34.13-q34.3									
CAD-1.23					9q34.2-q34.3	11q13.1-q13.3							21q22.2-q22.3	22q12.3-q13.33
CAD-1.24			8q24.3		9q34.13-q34.3	11q13.1-q13.3							21p11.2-q22.3	22q13.31-q13.33
CAD-1.25	1p36.32p36.13	4p16.3		8p23.1-p21.1*	9q34.12-q34.3	11q13.1-q13.3		15q23-q24.1	16p13.3; 16p13.13-p12.3	17q25.1-q25.3	20q13.33			22q13.2-q13.33
CAD-1.26					9q34.13-q34.3	11q13.1-q13.3								22q13.31-q13.33
CAD-1.30	1p36.33-p36.12			8p21.3-p21.1*	9q33.3-q34.3	11q12.2-q13.1	14q32.2-q32.33**		16q24.1-q24.3	17q25.3	20q13.33			22q12.2-q13.33
CAD-1.31	1p36.33-p36.13				9q34.13-q34.3	11q12.2-q13.3			16p13.3; 16p13.13-p12.3	17q24.3-q25.3	20q13.33	20p11.21q-11.21	21q22.2-q22.3	22q12.3-q13.33
CAD-1.32					9q34.13-q34.3	q13.2, q13.3				17q24.3-q25.3				22q13.2-q13.31
CAD-1.34	1p36.33-p36.13		8q24.3		9q33.3-q34.3	11q13.1-q13.3		15q23-q24.3	16p13.3	17q25.1-q25.3		20p11.21-q11.21	21q22.3	22q13.1-q13.33
CAD-1.37	1p36.33-p36.13	4p16.3-p16.1	8q24.3	8p23.3-p11.23*	9q34.11-q34.3	11q13.1-q13.2	14q32.32-q32.33**	15q23-q24.1	16q24.1-q24.3			20p11.21-q11.1	21q22.13-q22.3	22q12.2-q13.33
CAD-2.02				8p23.3-p12										
consensus region	1p36.31-p36.13	4p16.3	8q24.3	8p21.3-p21.1	9q34.2-q34.3	11q13.1-q13.3	14q32.32-q32.33**	15q23-q24.1	16q24.1-q24.3; 16p13.3	17q25.1-q25.3	20q13.33	20p11.21-q11.1	21q22.2-q22.3	22q13.31-q13.33
number of patients	7/13	3/13	3/13	4/13	12/13	6/13	3/13	3/13	3/13; 3/13	5/13	3/13	3/13	5/13	8/13

Table continues on next page.

	Not recurrent changes (<=2 samples)	
Samples	**gains**	**losses**
CAD-1.06	14q32.13-q32.2; 19p13.3-p13.2	5q35.2-q35.3; 6q27; 11q24.3-q25; 19p13.3
CAD-1.07	11p11.12-q11; 19p13.11-q13.2	19p13.3
CAD-1.22	16p11.2-p11.1	19q13.2
CAD-1.23		7q31.2-q32.3; 13q31.3-q33.2
CAD-1.24		
CAD-1.25	1q23.1-q31.3; 1q42.13-q44; 2q35-q37.1; 2q37.3; 5q21.1-q21.2; +12q14.1; +12q24.31-q24.33; +18q21.2-q21.31; +18q22.1-q22.3; 19p13.3-p13.2; 20q11.22-q11.23;21p11.2-q21.1; Xp21.1-p11.4	3p21.31-p21.2; 6p11.2-q11.1; 6q23.3-q24.1; 10q11.21-q11.22; 10q23.31*
CAD-1.26		
CAD-1.30	17q22-q23.2; 22q11.21-q11.22; Yp11.2-q12	5q35.3; 6p22.2-p21.31; 6q16.1*; 6q27; 13q34; 17q11.1-q11.2
CAD-1.31	10q24.32*;	2q37.3*; 19q13.2
CAD-1.32		
CAD-1.34	11p15.5-p15.4; +14q12*; 15q11.1-q12; 19p13.11-p12; 19q13.32-q13.33; 20q11.21-q11.23; Xp11.21-q12; Xq21.1-q22.1; Xq23q-26.1; Yp11.2-q12	1q22-q23.2; 1q32.1-q41; 10q11.22-q21.1; 17q11.2-q12;
CAD-1.37	2p22.1-p21; 5p12-q11.1; 22q11.21-q11.23	1p36.11-p35.2; 17p13.1; 17q21.31q-21.32; 20q11.23q12
CAD-2.02		8q11.1q12.1; chrX

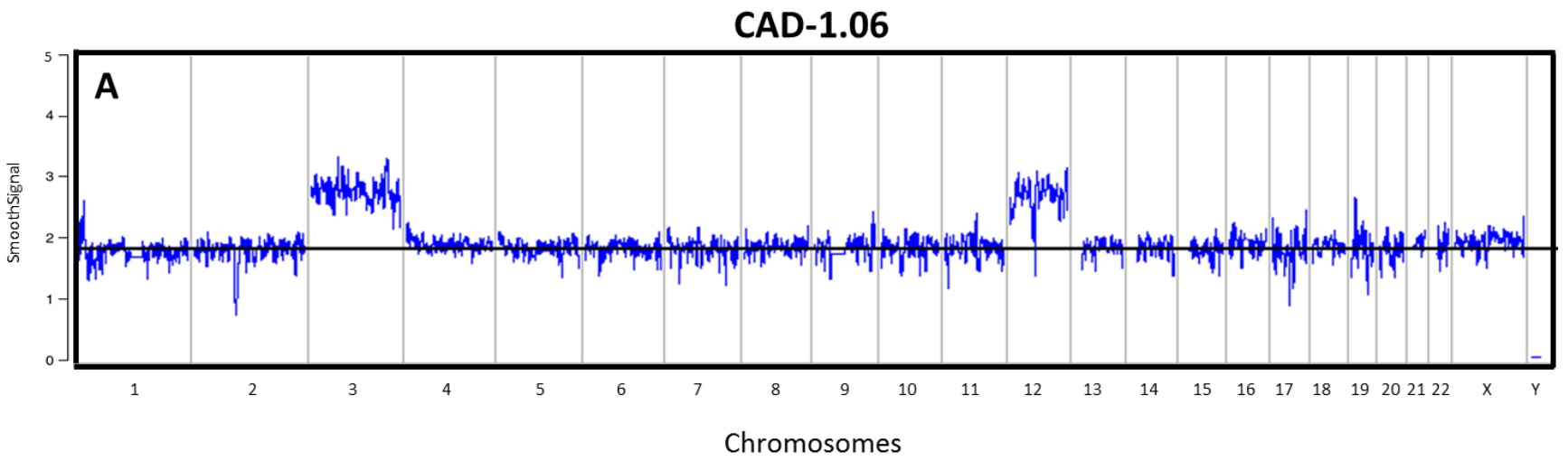
* manually corrected data -small change removed due to 3 Mb filter, however retained since it also appear in exome sequencing data.

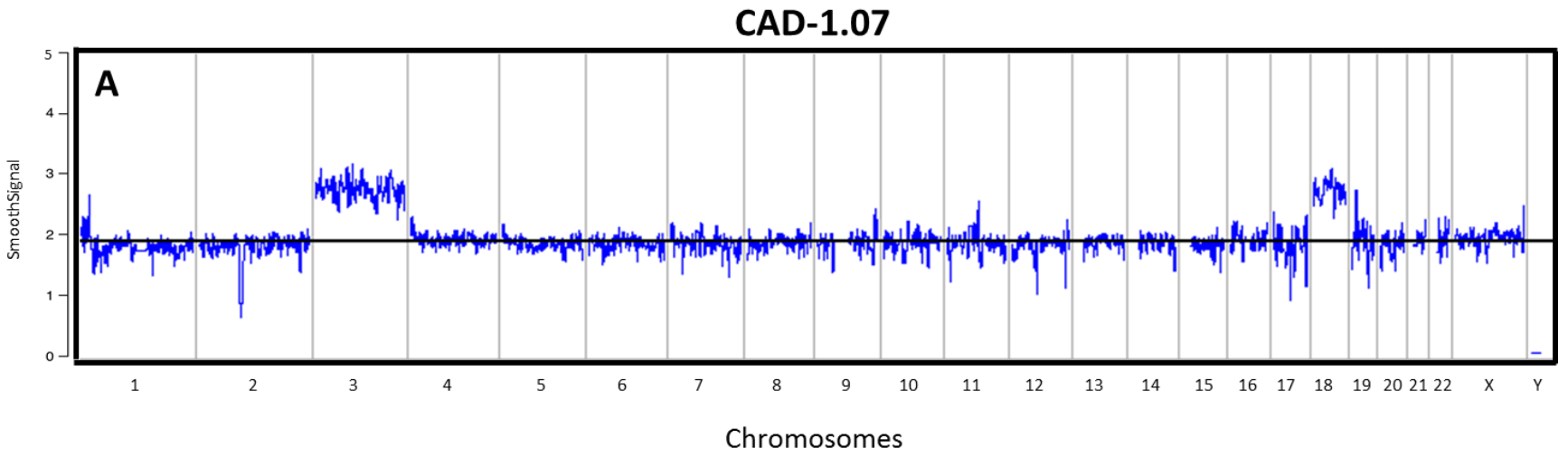
** IGH gene complex locus (#14q32.3)

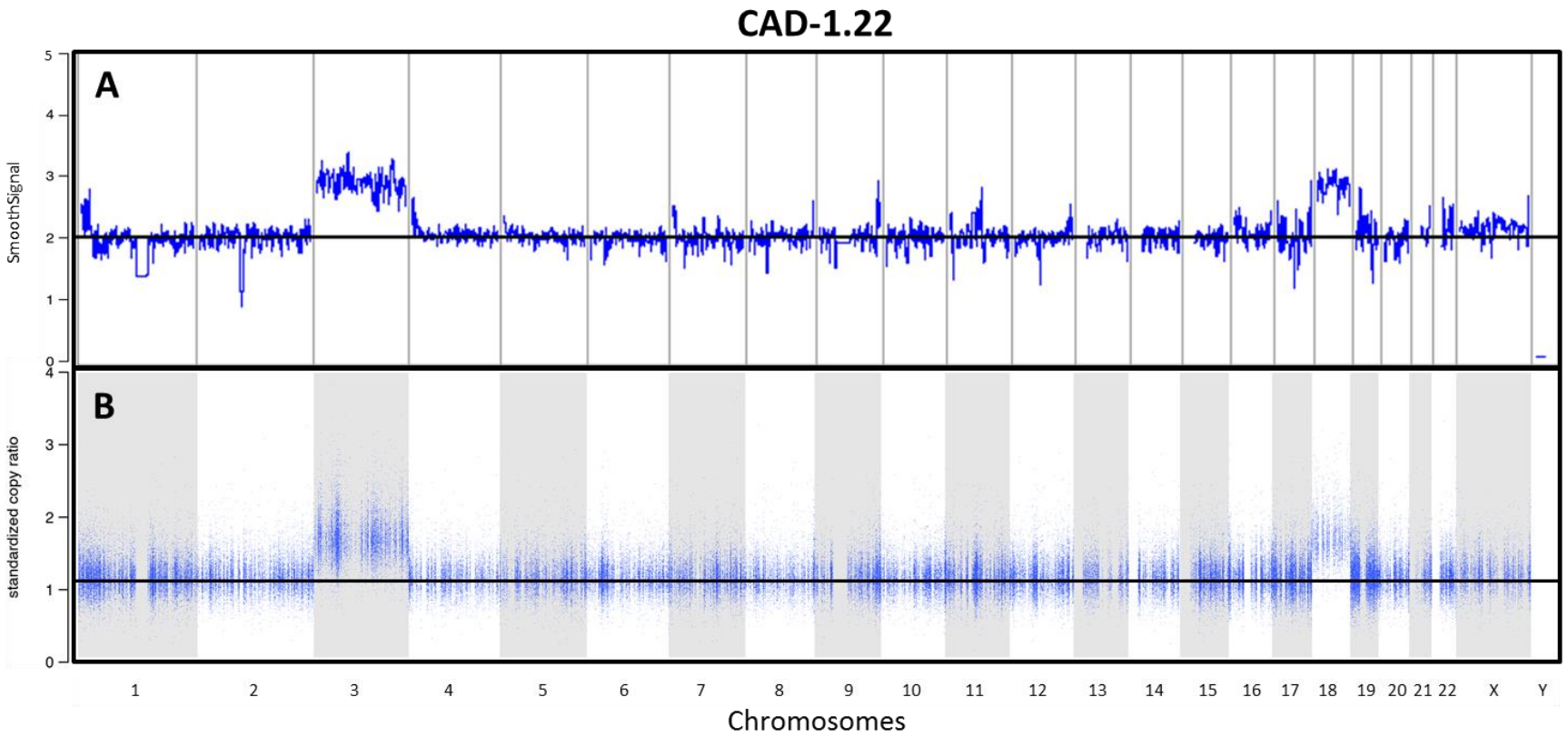
Supplemental Figure 1. Chromosomal profile of gains and losses in each CAD sample.

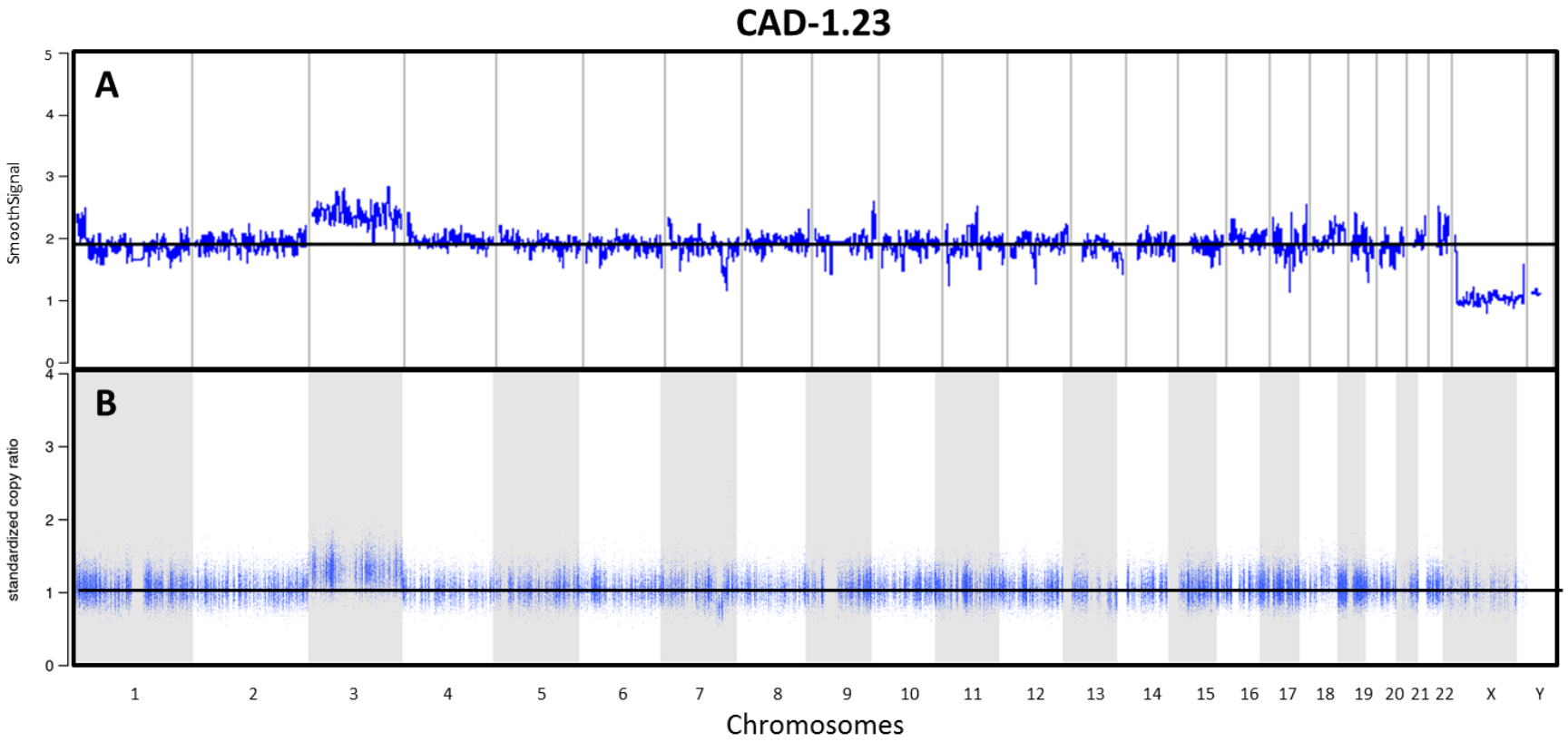
Upper panels for each case (A) show results from cytogenetic microarray OncoScan CNV Plus Assay (Smooth Signal Copy Number) and horizontal black line marks normal (diploid) DNA copy number.

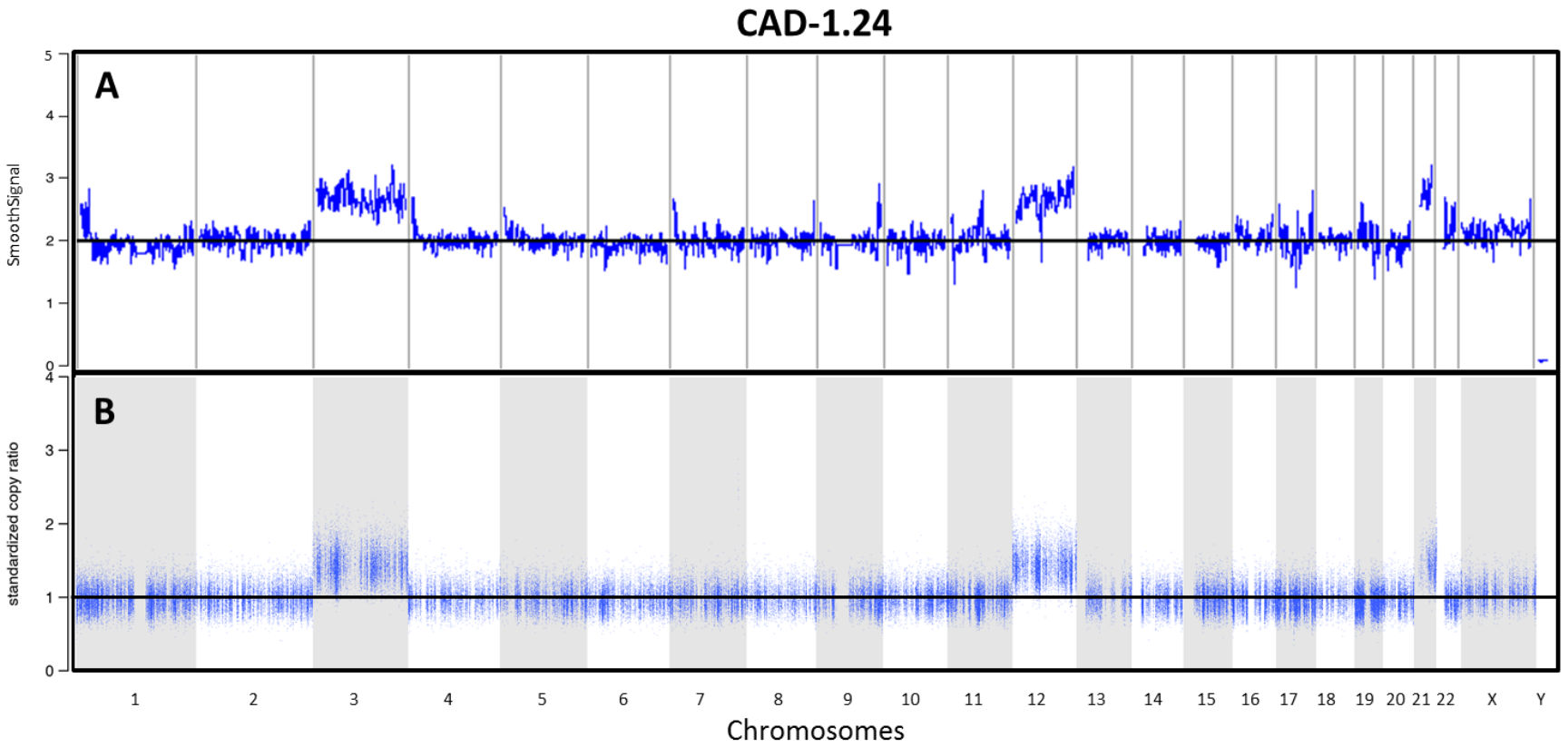
Lower panels for each case (B) show results from exome sequencing CNV analysis using GATK4 and horizontal black line marks normal standardized copy ratio as compared to normal sample.

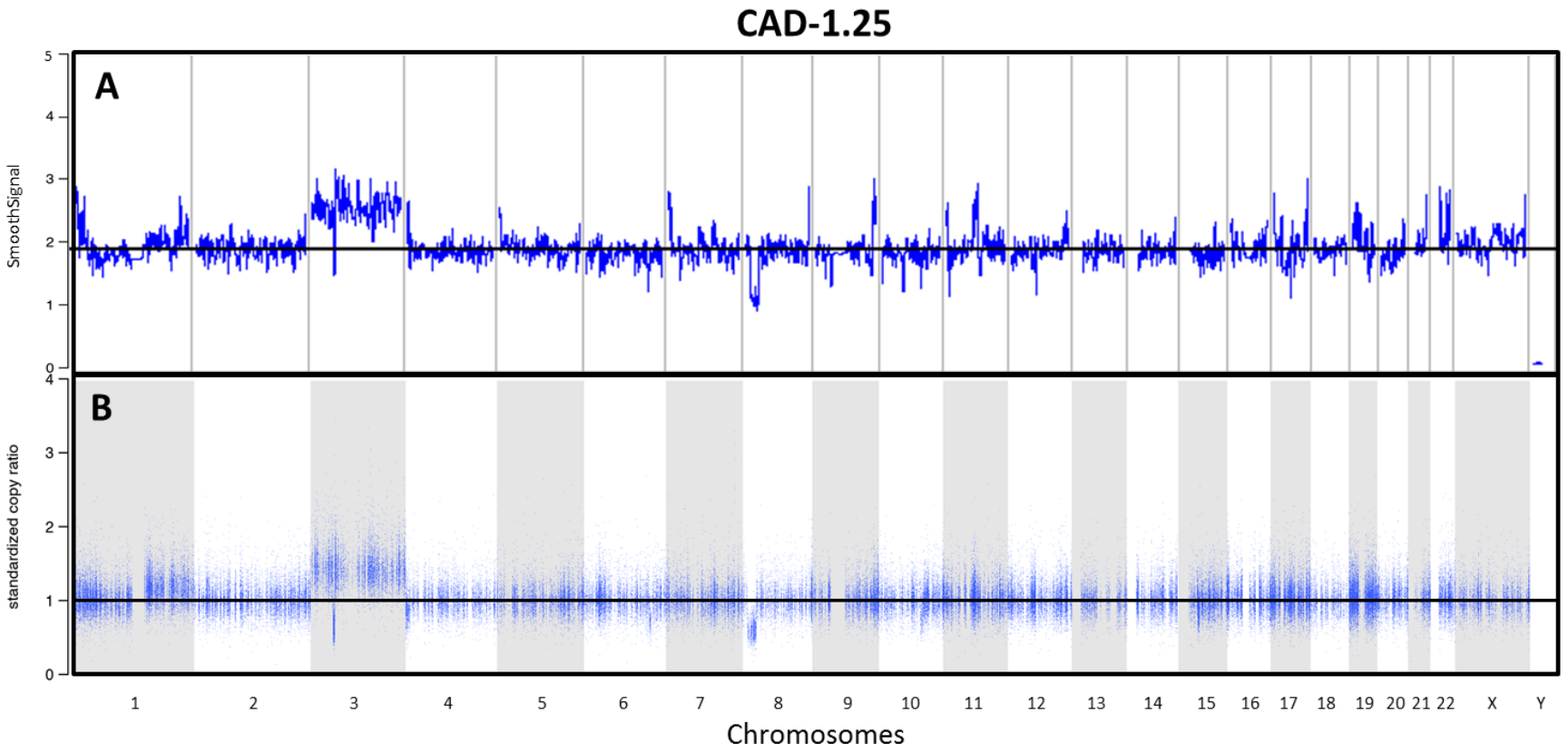

CAD-1.06
A
SmoothSignal
0
1
2
3
4
5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Chromosomes

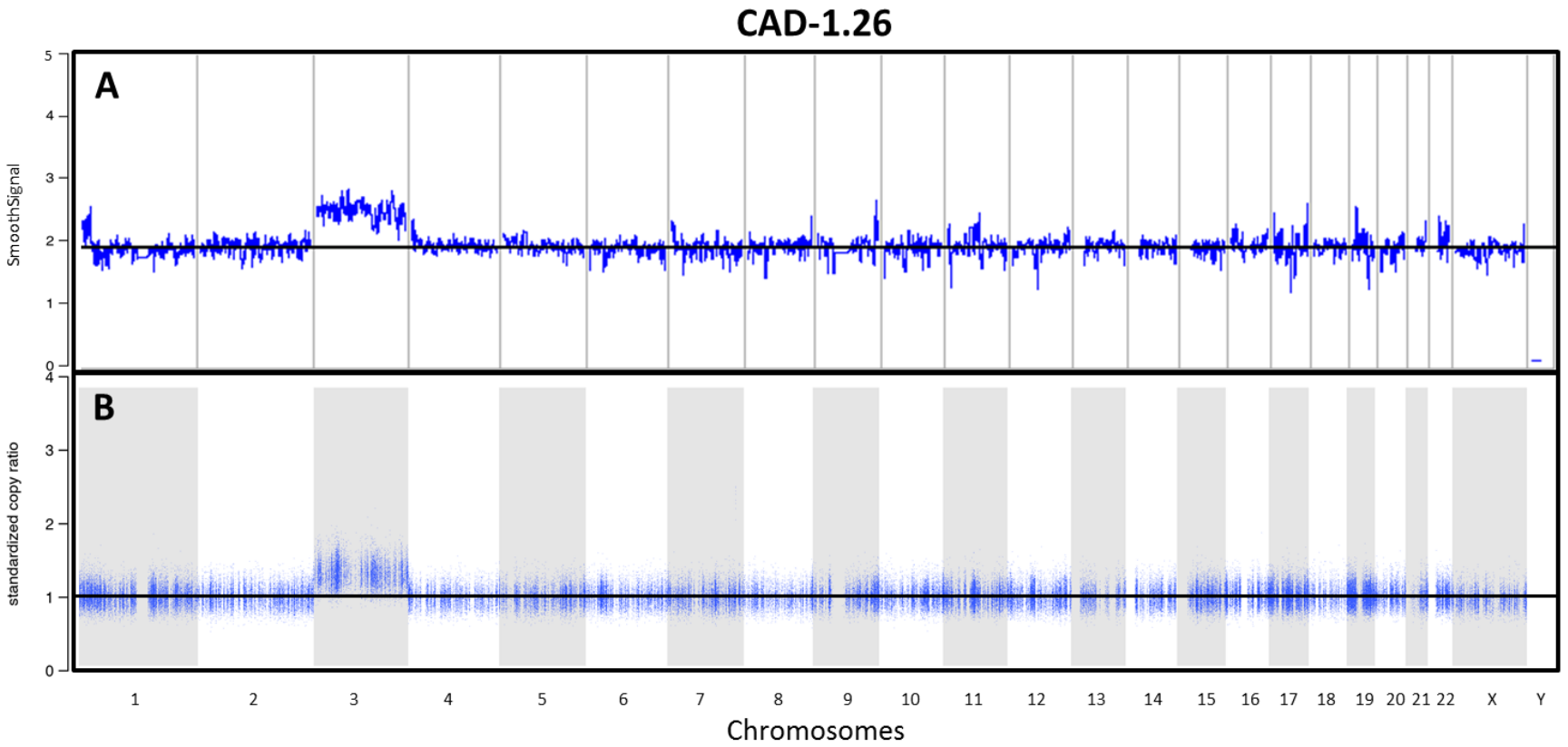

CAD-1.07
A
SmoothSignal
0
1
2
3
4
5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Chromosomes

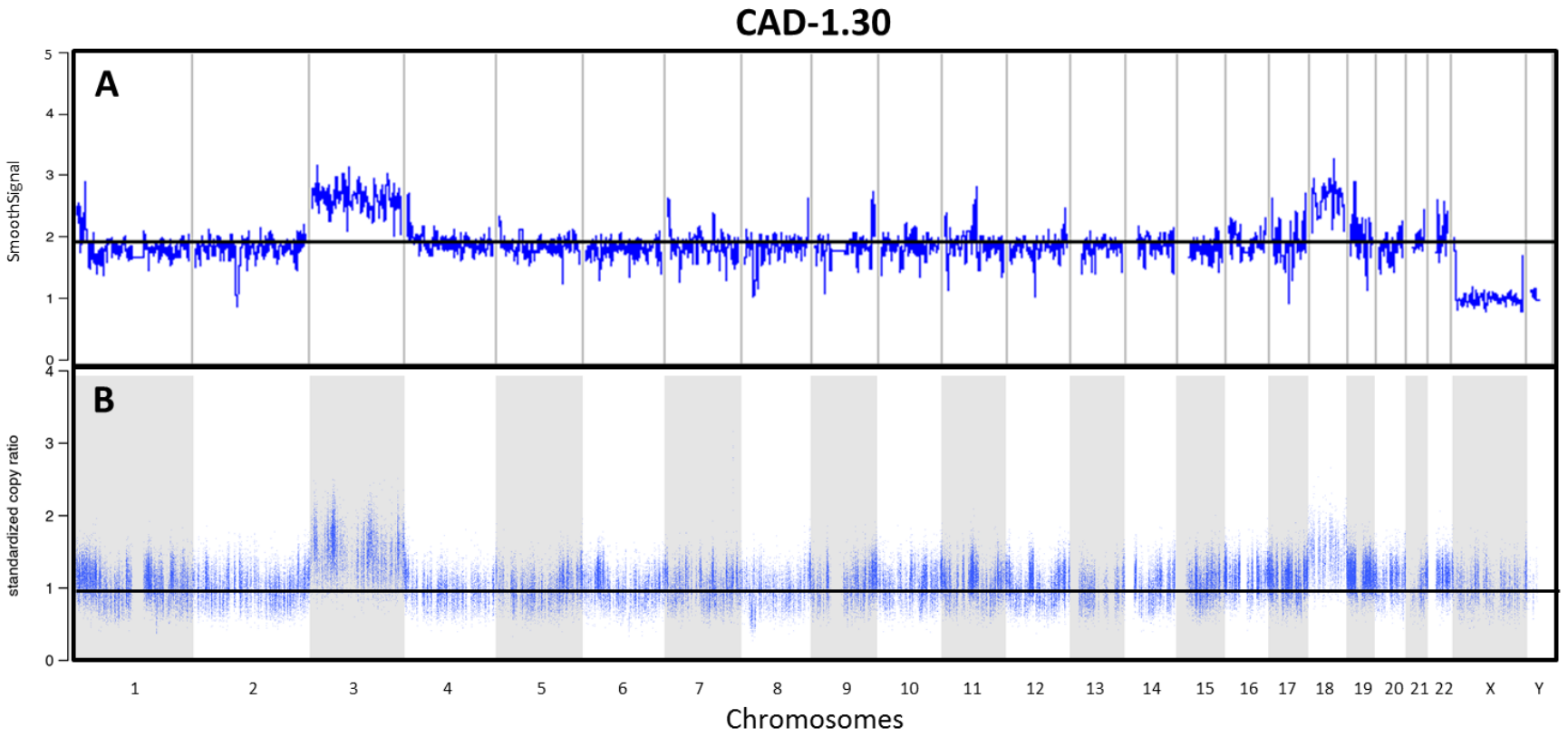

CAD-1.22
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

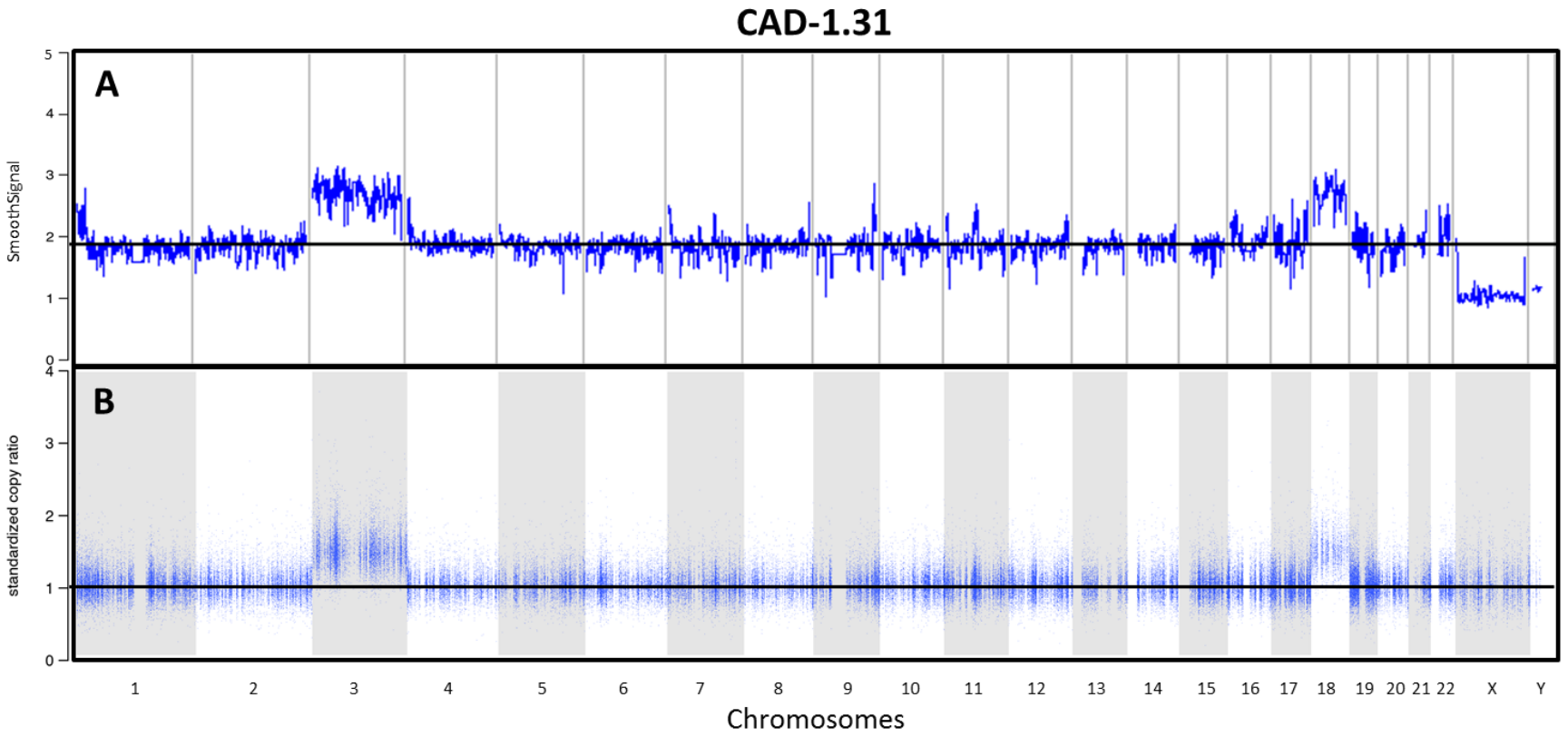

CAD-1.23
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

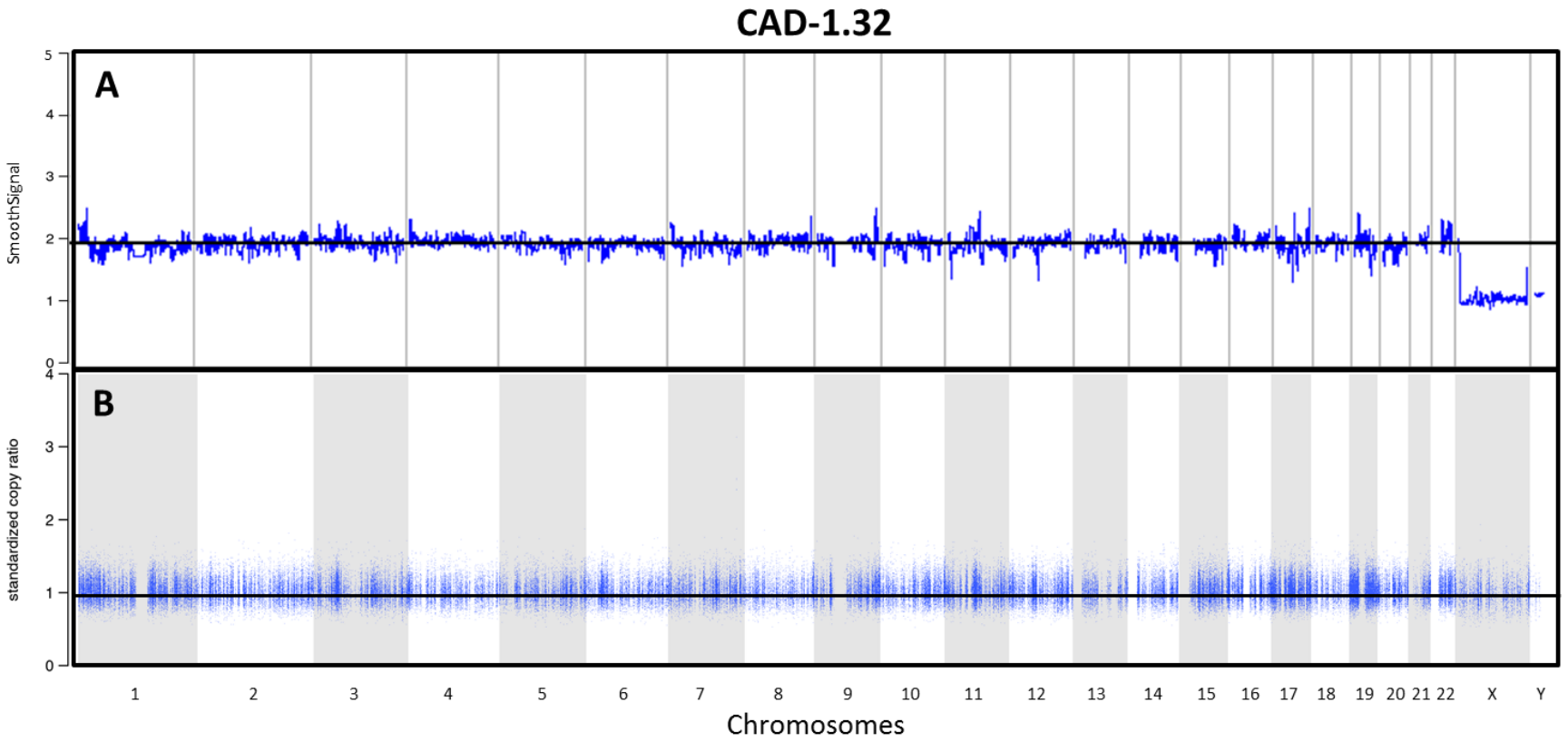

CAD-1.24
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

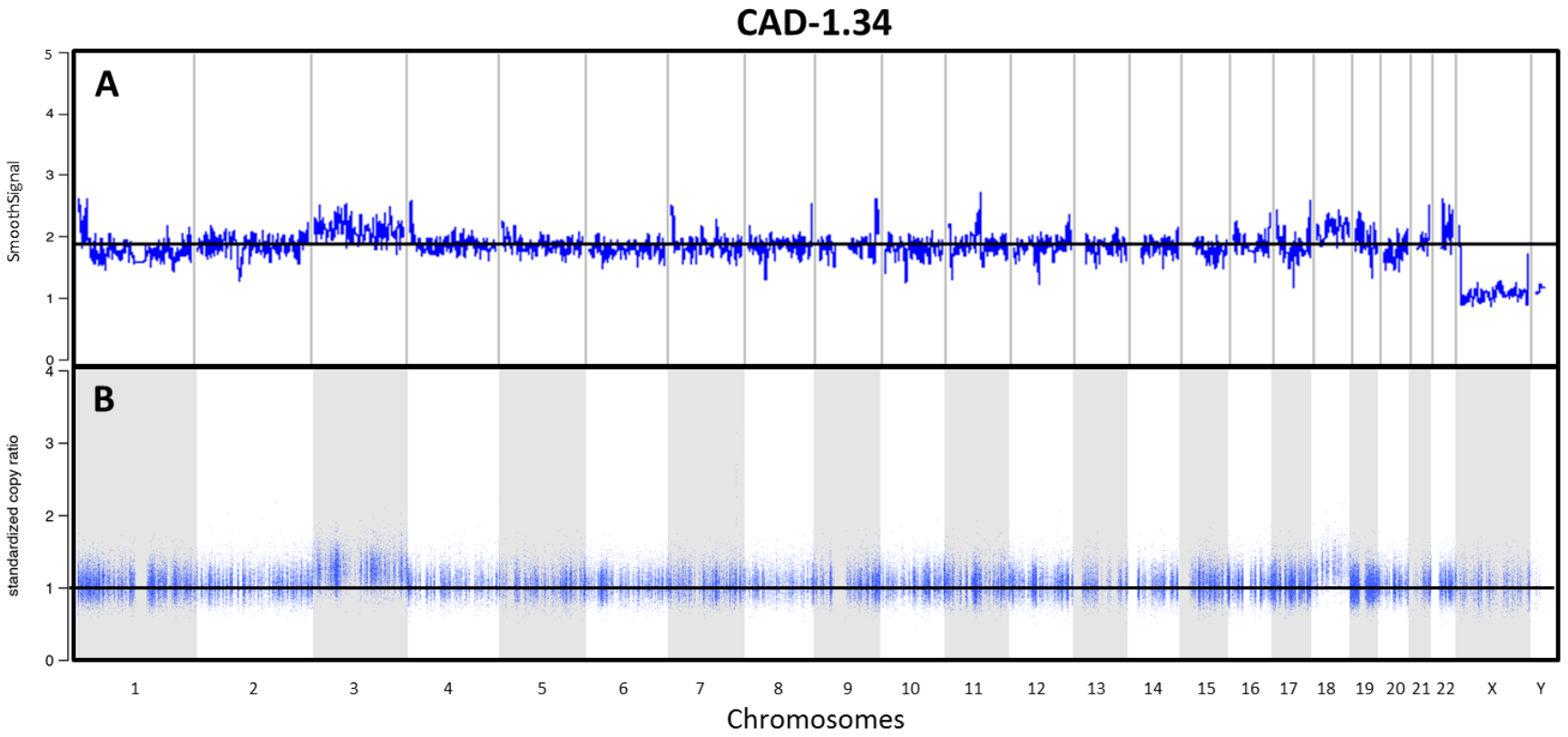

CAD-1.25
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

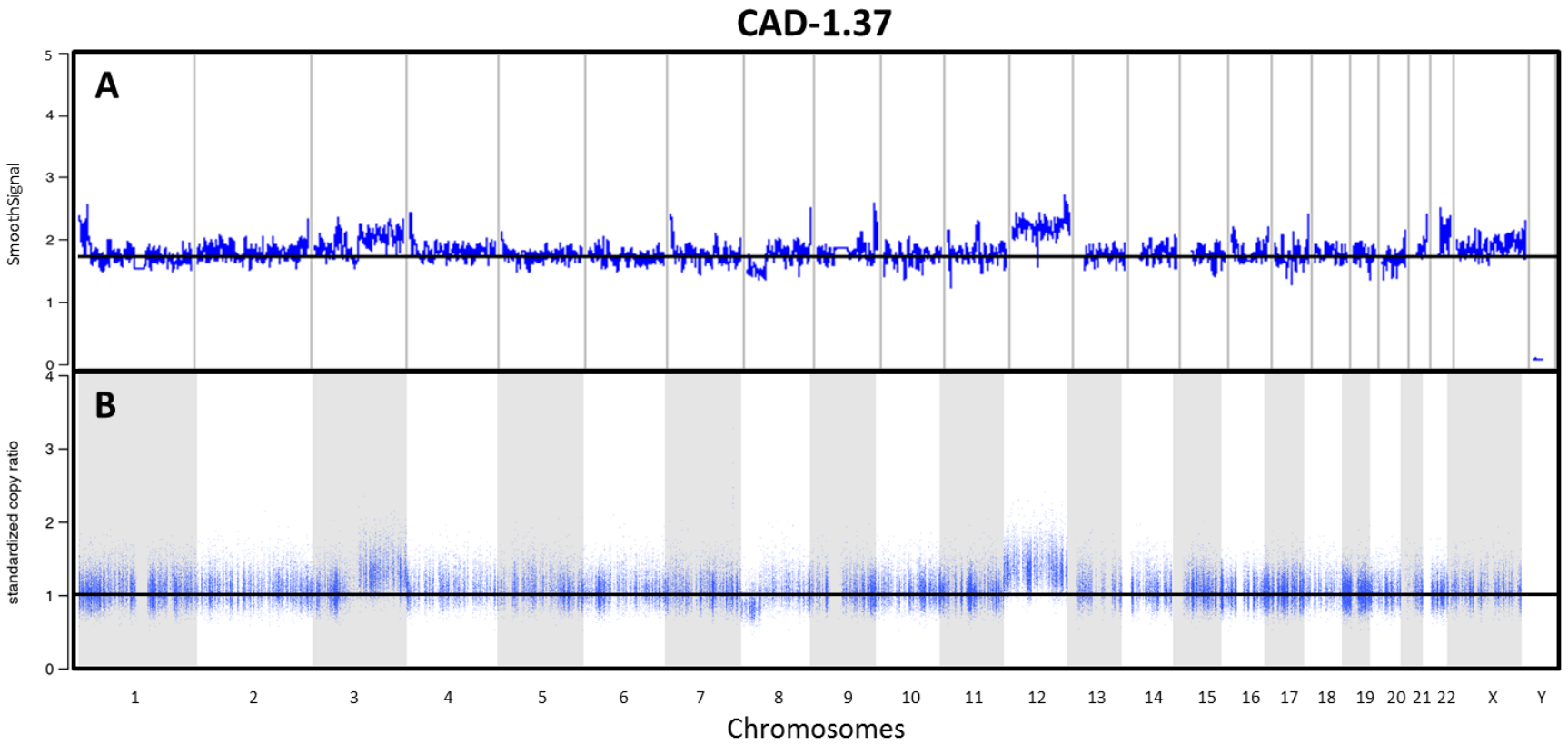

CAD-1.26
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

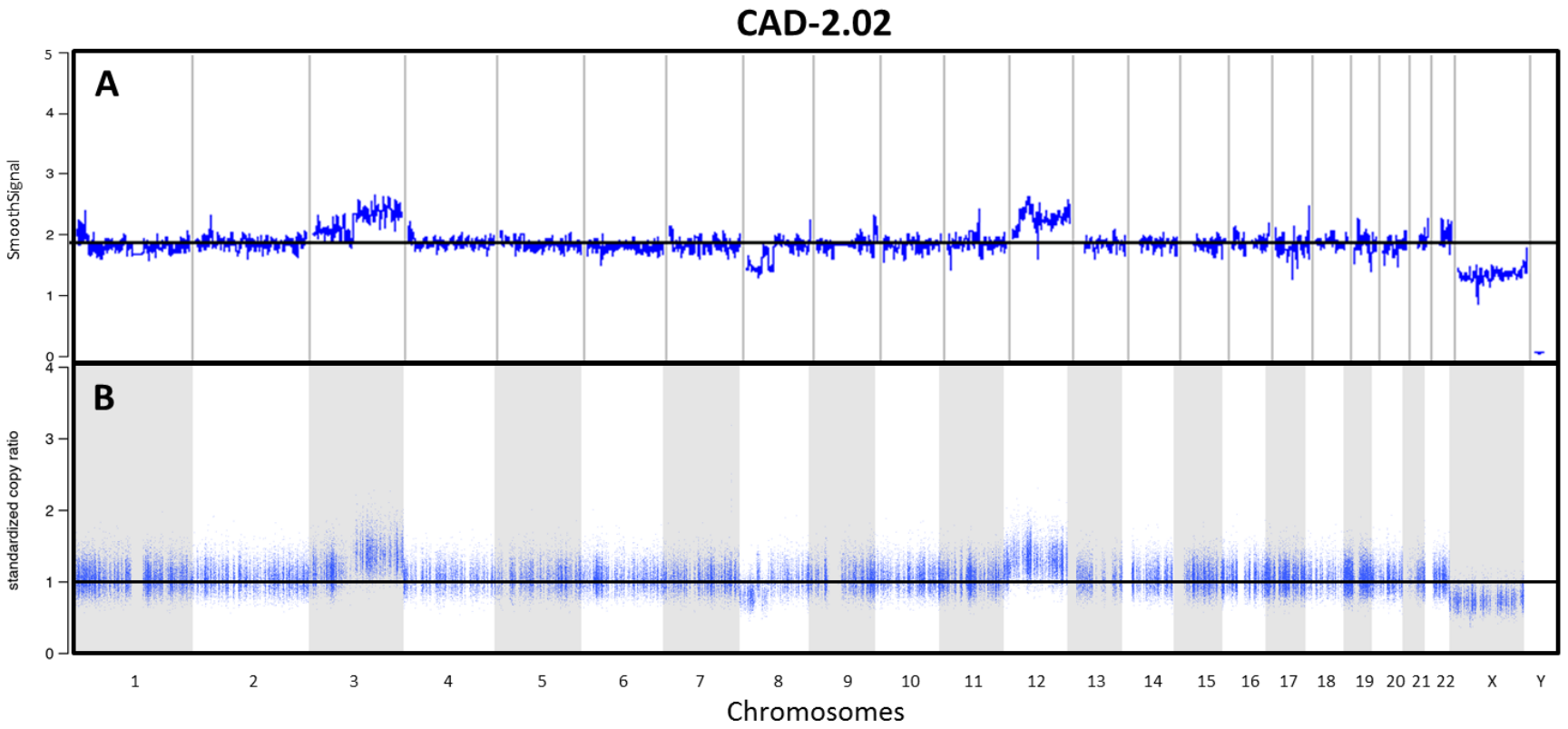

CAD-1.30
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

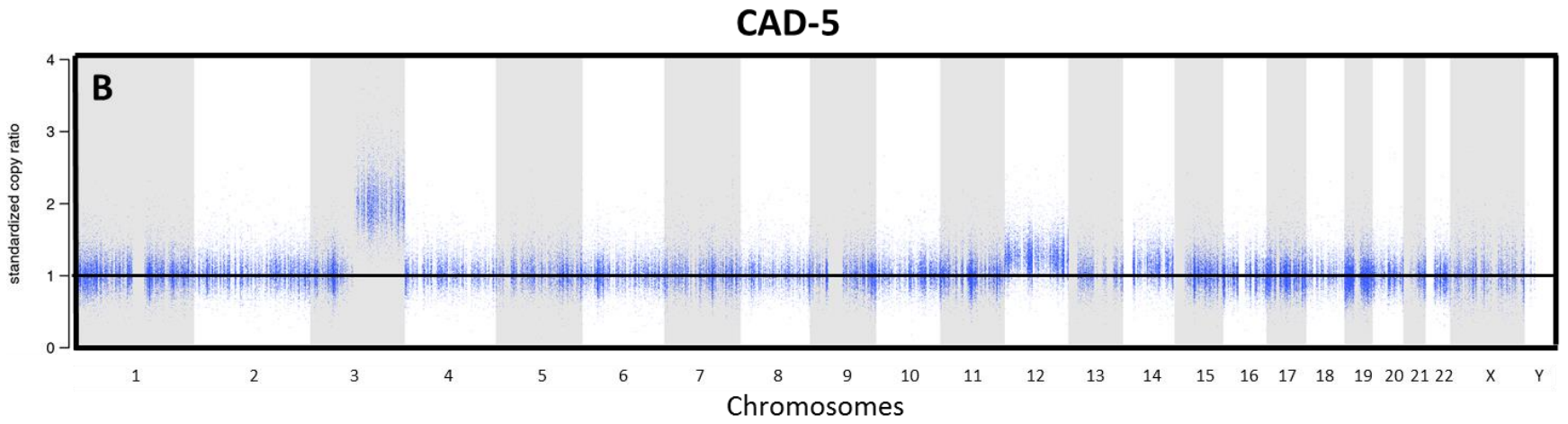

CAD-1.31
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

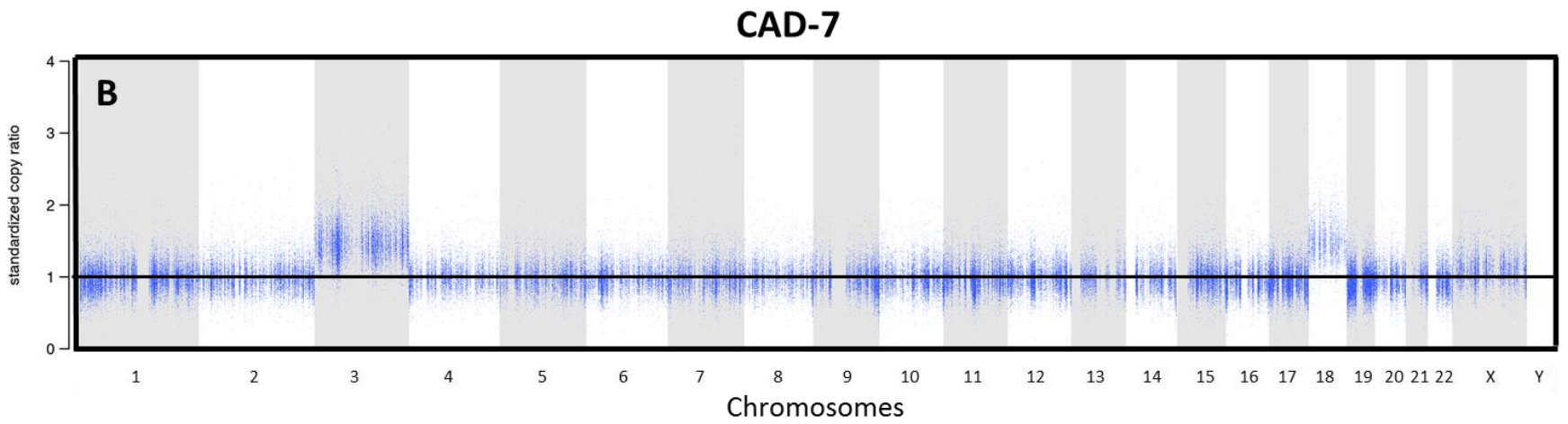

CAD-1.32
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes


CAD-1.34
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes


CAD-1.37
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes

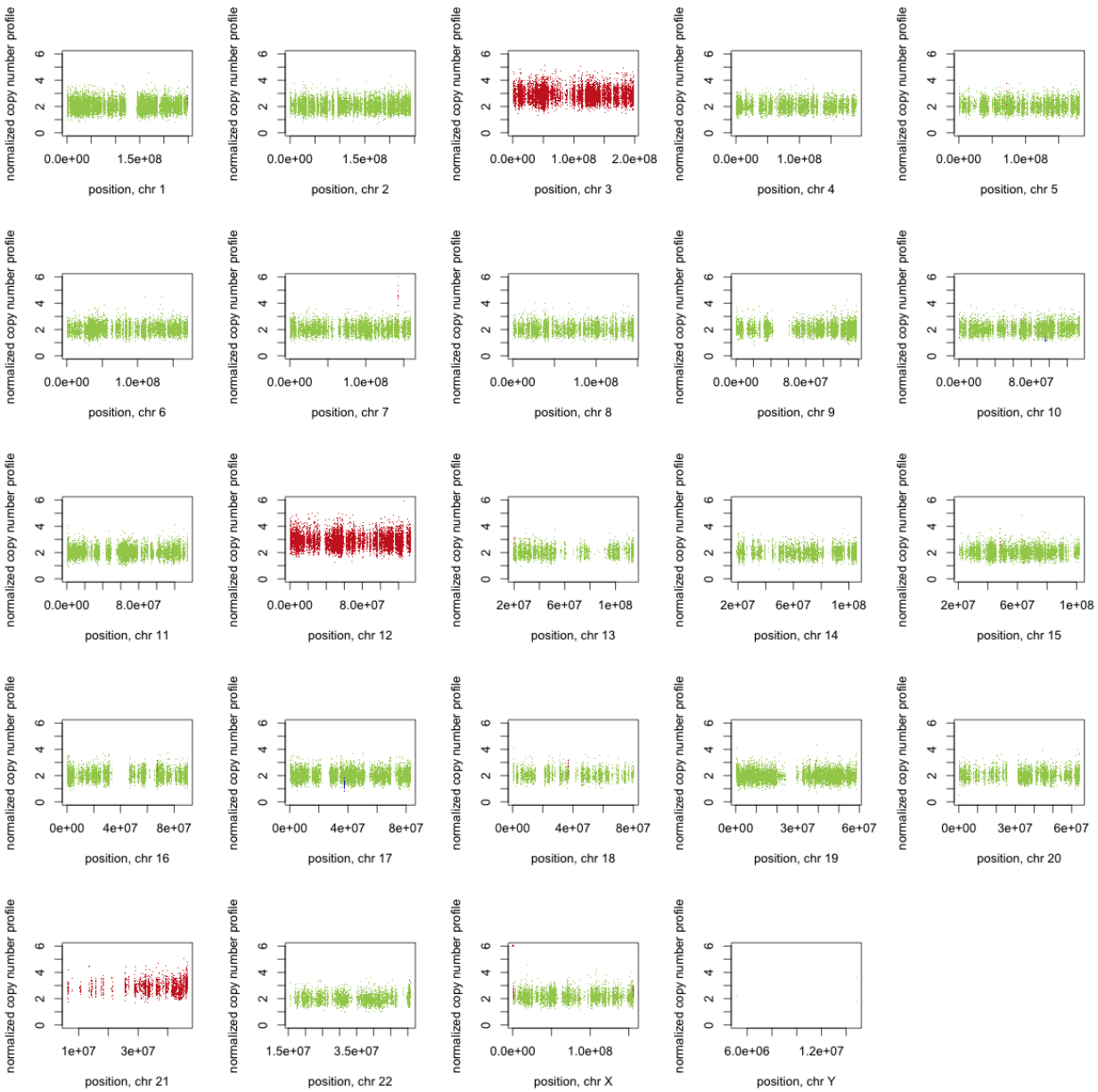

CAD-2.02
A
B
SmoothSignal
standardized copy ratio
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Chromosomes


CAD-5
B
standardized copy ratio
0
1
2
3
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Chromosomes

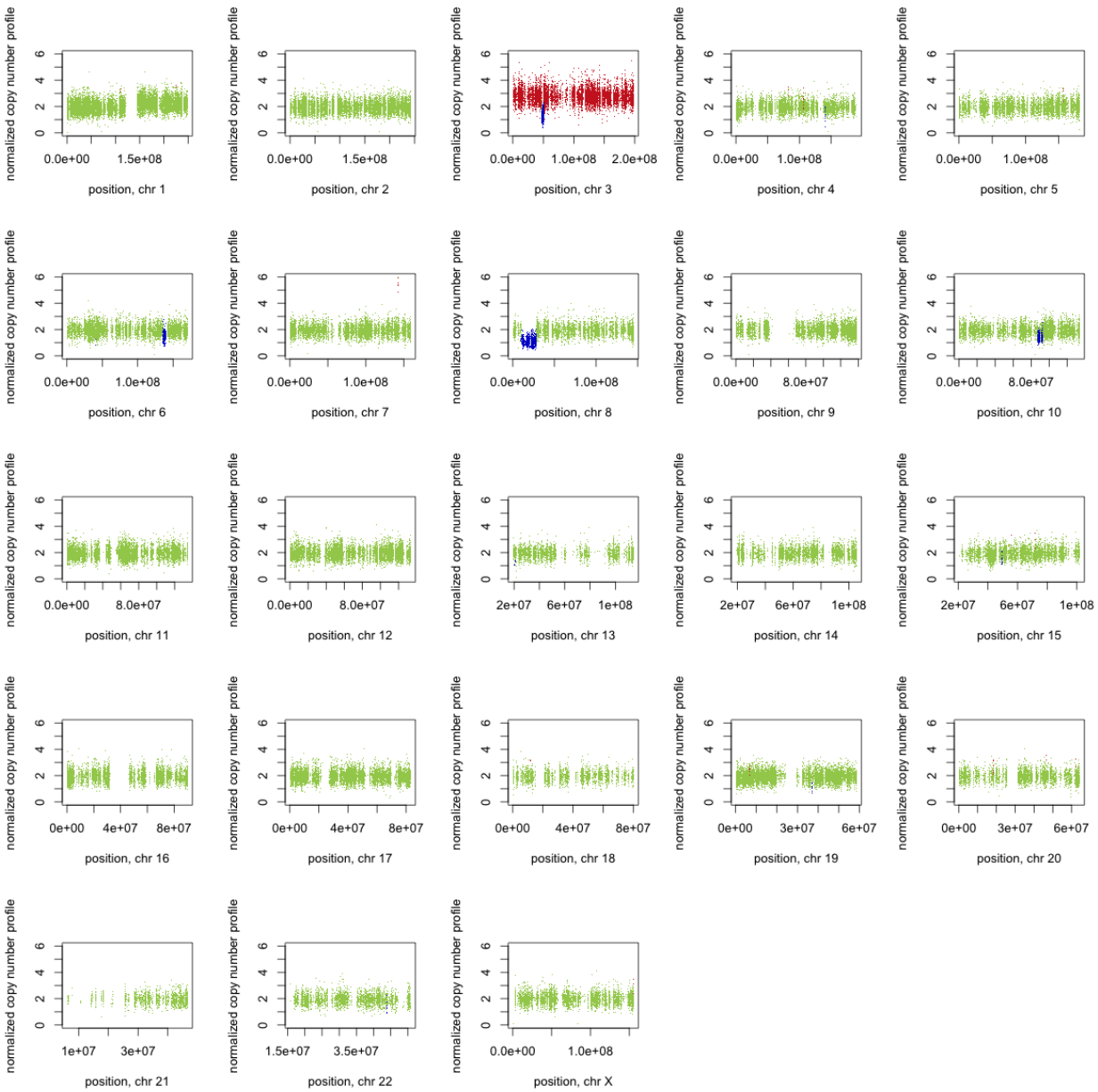

CAD-7
B
standardized copy ratio
0
1
2
3
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Chromosomes

Supplemental Figure 2. Normalized copy number profile of all chromosomes indicating gains (red) and losses (blue) for 4 representative CAD patient samples obtained from exome sequencing CNV analysis using Control-FREEC.

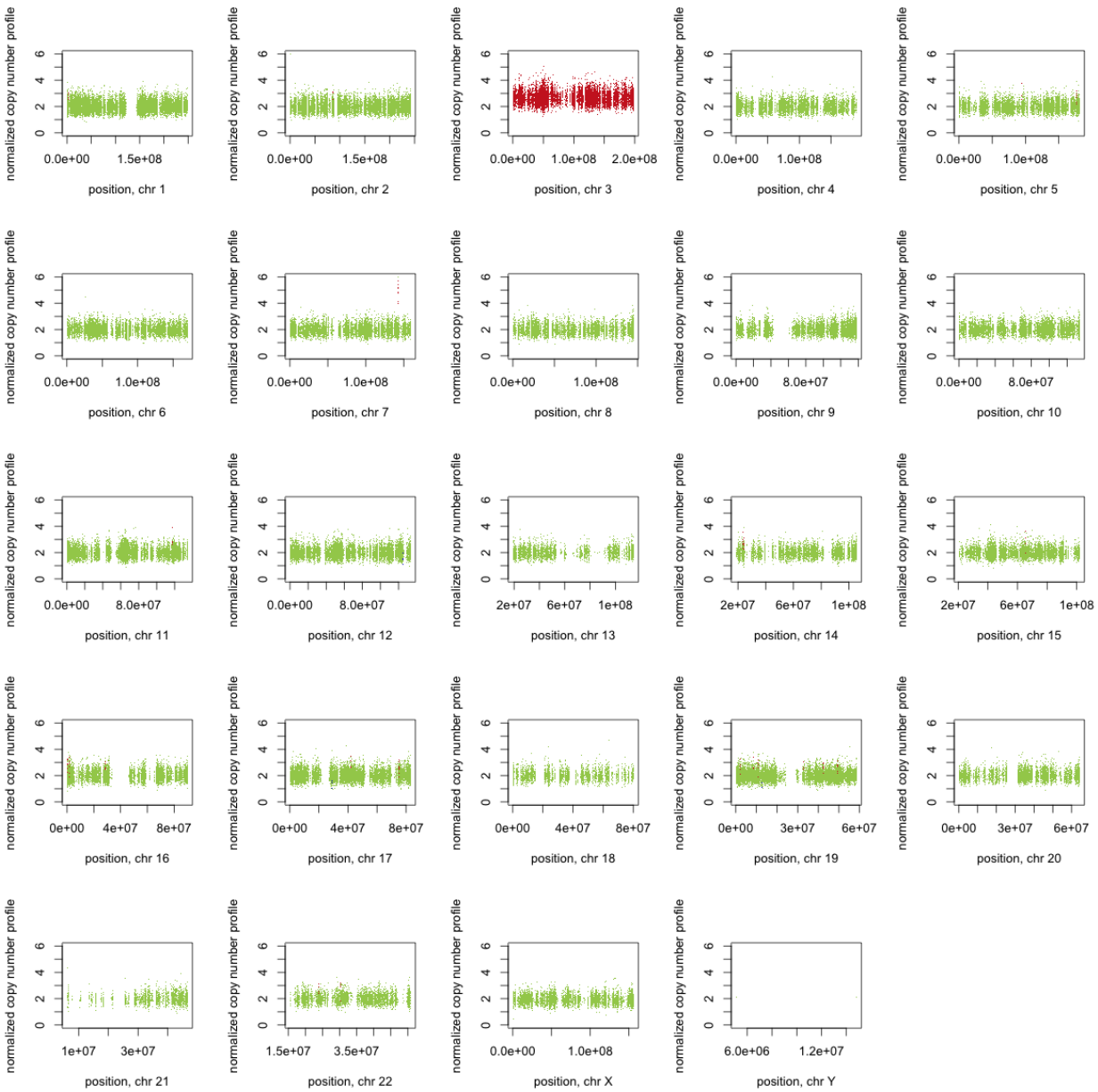
(A) CAD-1.24; (B) CAD-1.25; (C) CAD-1.26; (D) CAD-1.30


normalized copy number profile
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr X
position, chr Y

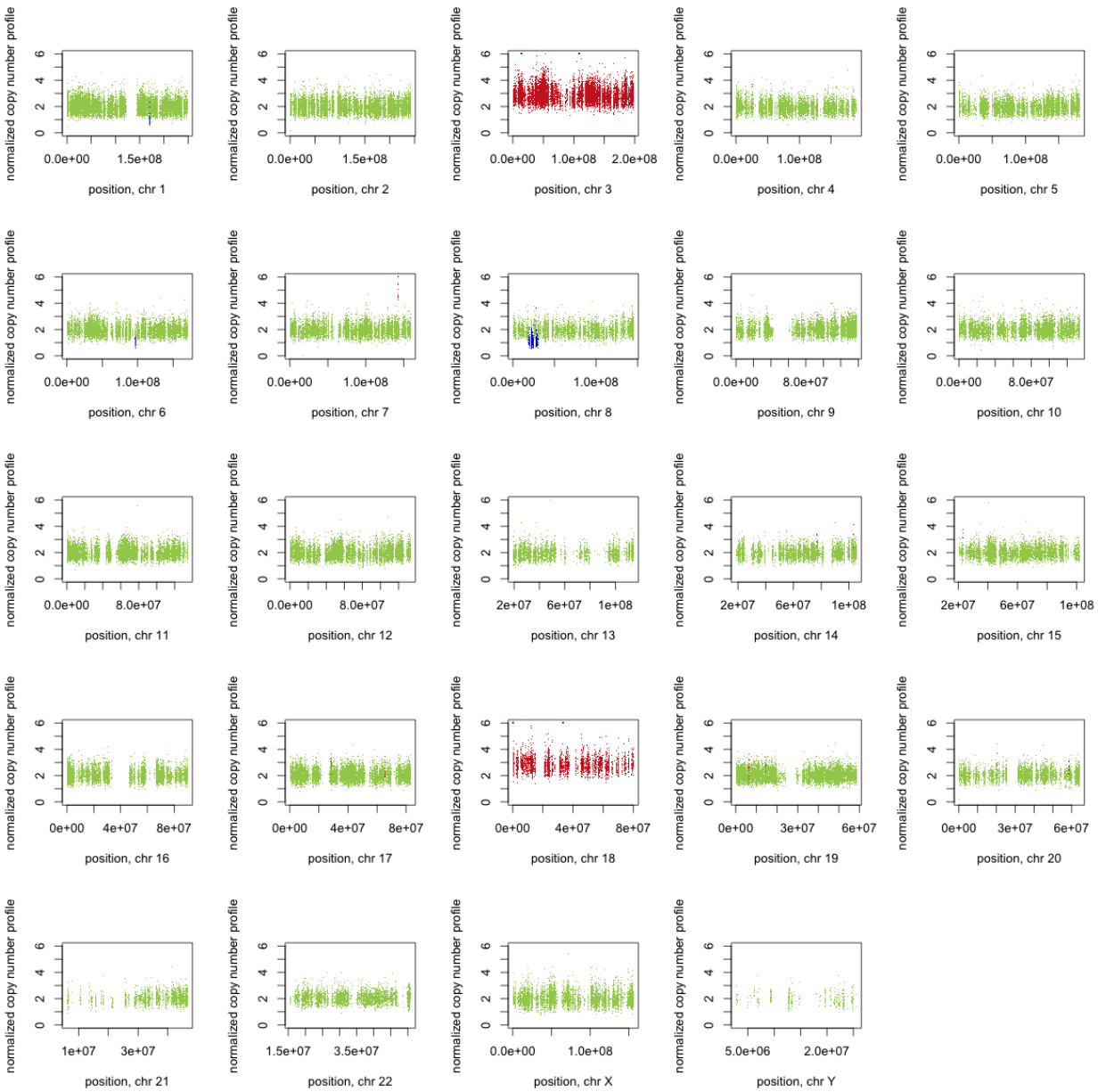
A


normalized copy number profile
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr X

B


normalized copy number profile
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr X
position, chr Y

C


normalized copy number profile
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr X
position, chr Y

D

References

1. Berentsen S, Randen U, Oksman M, Birgens H, Tvedt THA, Dalgaard J, Galteland E, Haukas E, Brudevold R, Sorbo JH, Naess IA, Malecka A, Tjonnfjord GE. Bendamustine plus rituximab for chronic cold agglutinin disease: results of a Nordic prospective multicenter trial. *Blood*. 2017;130(4):537-541.
2. Malecka A, Troen G, Tierens A, Ostlie I, Malecki J, Randen U, Wang J, Berentsen S, Tjonnfjord GE, Delabie JMA. Frequent somatic mutations of KMT2D (MLL2) and CARD11 genes in primary cold agglutinin disease. *Br J Haematol*. 2018;183(5):838-842.
3. Malecka A, Troen G, Tierens A, Ostlie I, Malecki J, Randen U, Berentsen S, Tjonnfjord GE, Delabie JM. Immunoglobulin heavy and light chain gene features are correlated with primary cold agglutinin disease onset and activity. *Haematologica*. 2016;101(9):e361-364.
4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
5. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-498.
6. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
7. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11 10 11-33.
8. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, Janoueix-Lerosey I, Delattre O, Barillot E. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*. 2012;28(3):423-425.
9. Rinaldi A, Mian M, Chigrinova E, Arcaini L, Bhagat G, Novak U, Rancoita PMV, De Campos CP, Forconi F, Gascoyne RD, Facchetti F, Ponzoni M, Govi S, Ferreri AJM, Mollejo M, Piris MA, Baldini L, Soulier J, Thieblemont C, Canzonieri V, Gattei V, Marasca R, Franceschetti S, Gaidano G, Tucci A, Uccella S, Tibiletti MG, Dirnhofer S, Tripodo C, Doglioni C, Dalla Favera R, Cavalli F, Zucca E, Kwee I, Bertoni F. Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact on the clinical outcome. *Blood*. 2011;117(5):1595-1604.
10. Ganapathi KA, Jobanputra V, Iwamoto F, Jain P, Chen J, Cascione L, Nahum O, Levy B, Xie Y, Khattar P, Hoehn D, Bertoni F, Murty VV, Pittaluga S, Jaffe ES, Alobeid B, Mansukhani MM, Bhagat G. The genetic landscape of dural marginal zone lymphomas. *Oncotarget*. 2016;7(28):43052-43061.
11. Braggio E, Fonseca R. Genomic abnormalities of Waldenström macroglobulinemia and related low-grade B-cell lymphomas. *Clin Lymphoma Myeloma Leuk*. 2013;13(2):198-201.
12. Braggio E, Dogan A, Keats JJ, Chng WJ, Huang G, Matthews JM, Maurer MJ, Law ME, Bosler DS, Barrett M, Lossos IS, Witzig TE, Fonseca R. Genomic analysis of marginal zone and lymphoplasmacytic lymphomas identified common and disease-specific abnormalities. *Mod Pathol*. 2012;25(5):651-660.
13. Berglund M, Enblad G, Thunberg U, Amini R-M, Sundström C, Roos G, Erlanson M, Rosenquist R, Larsson C, Lagercrantz S. Genomic imbalances during transformation from follicular lymphoma to diffuse large B-cell lymphoma. *Mod Pathol*. 2007;20(1):63-75.

14. Cheung K-JJ, Shah SP, Steidl C, Johnson N, Relander T, Telenius A, Lai B, Murphy KP, Lam W, Al-Tourah AJ, Connors JM, Ng RT, Gascoyne RD, Horsman DE. Genome-wide profiling of follicular lymphoma by array comparative genomic hybridization reveals prognostically significant DNA copy number imbalances. *Blood*. 2009;113(1):137-148.