nature genetics

SUPPLEMENTARY INFORMATION

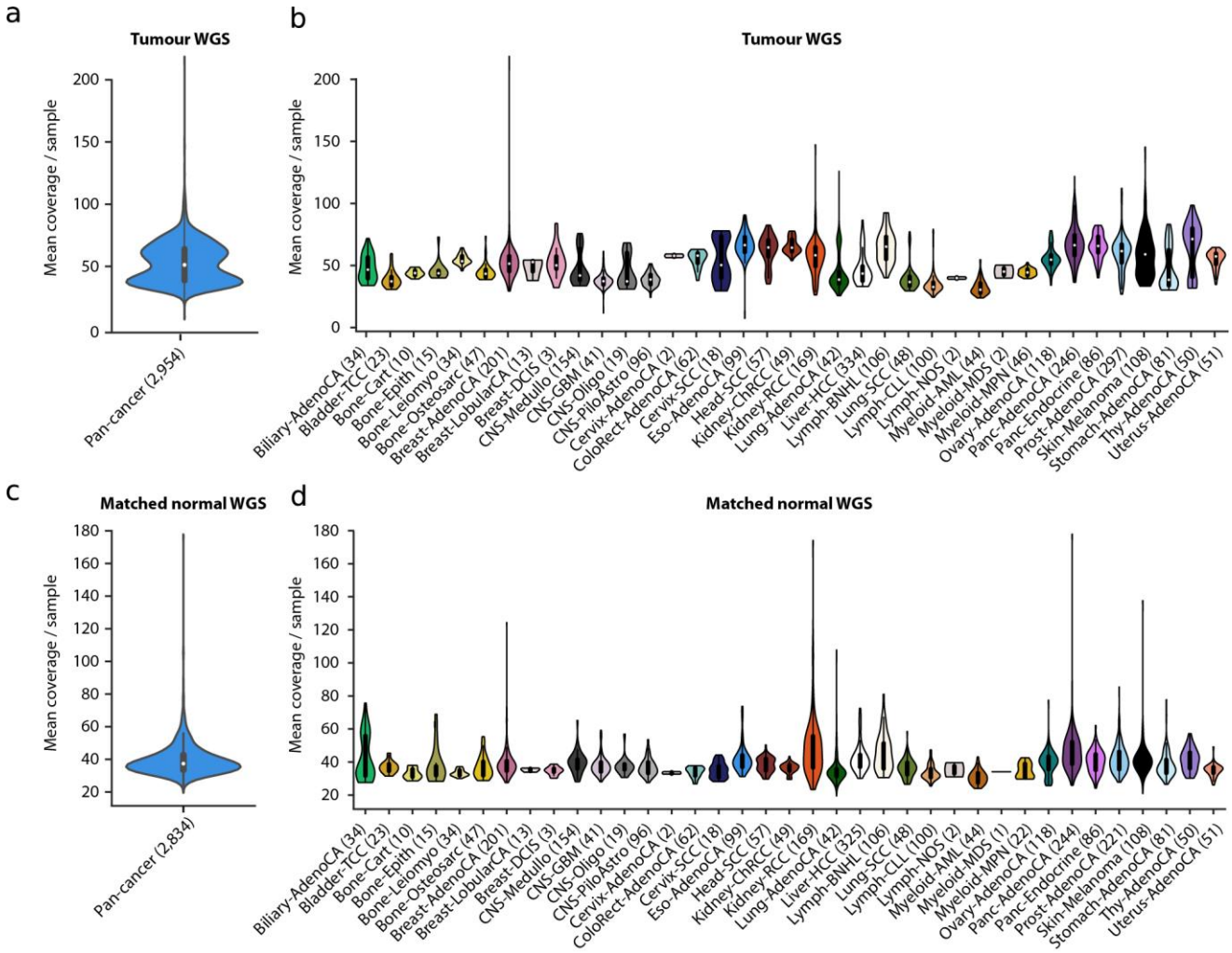
https://doi.org/10.1038/s41588-019-0562-0

In the format provided by the authors and unedited.

Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition

Bernardo Rodriguez-Martin[1,2,3], Eva G. Alvarez[1,2,3,44], Adrian Baez-Ortega[4,44], Jorge Zamora[1,2,44], Fran Supek[5,6,44], Jonas Demeulemeester[7,8], Martin Santamarina[1,2,3], Young Seok Ju[9,10], Javier Temes[1], Daniel Garcia-Souto[1], Harald Detering[3,11,12], Yilong Li[10], Jorge Rodriguez-Castro[1], Ana Dueso-Barroso[13,14], Alicia L. Bruzos[1,2,3], Stefan C. Dentro[7,15,16], Miguel G. Blanco[17,18], Gianmarco Contino[19], Daniel Ardeljan[20], Marta Tojo[11], Nicola D. Roberts[10], Sonia Zumalave[1,2], Paul A. W. Edwards[21,22], Joachim Weischenfeldt[23,24,25], Montserrat Puiggròs[13], Zechen Chong[26,27], Ken Chen[26], Eunjung Alice Lee[28,29], Jeremiah A. Wala[29,30,31], Keiran Raine[10], Adam Butler[10], Sebastian M. Waszak[25], Fabio C. P. Navarro[32,33,34], Steven E. Schumacher[29,30,31], Jean Monlong[35], Francesco Maura[10,36,37], Niccolo Bolli[36,37], Guillaume Bourque[35], Mark Gerstein[32,33], Peter J. Park[38], David C. Wedge[39,10,16], Rameen Beroukhim[29,30,31], David Torrents[13,6], Jan O. Korbel[25], Inigo Martincorena[10], Rebecca C. Fitzgerald[19], Peter Van Loo[7,8], Haig H. Kazazian[20], Kathleen H. Burns[20,40], PCAWG Structural Variation Working Group[41], Peter J. Campbell[10,42,45*], Jose M. C. Tubio[1,2,3,10,45*] and PCAWG Consortium[43]

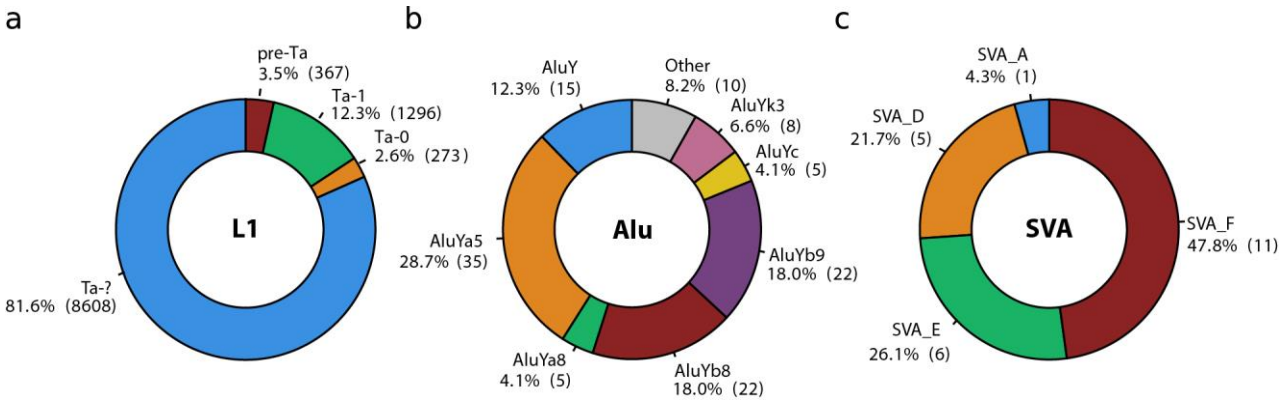
[1]Genomes and Disease, Centre for Research in Molecular Medicine and Chronic Diseases (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [2]Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [3]Biomedical Research Centre (CINBIO), University of Vigo, Vigo, Spain. [4]Transmissible Cancer Group, Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. [5]Genome Data Science, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [6]Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. [7]The Francis Crick Institute, London, UK. [8]Department of Human Genetics, University of Leuven, Leuven, Belgium. [9]Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. [10]Cancer Ageing and Somatic Mutation Programme, Wellcome Sanger Institute, Cambridge, UK. [11]Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain. [12]Galicia Sur Health Research Institute, Vigo, Spain. [13]Barcelona Supercomputing Center (BSC-CNS), Barcelona, Spain. [14]Faculty of Science and Technology, University of Vic—Central University of Catalonia (UVic-UCC), Vic, Spain. [15]Experimental Cancer Genetics, Wellcome Sanger Institute, Cambridge, UK. [16]Oxford Big Data Institute, University of Oxford, Oxford, UK. [17]DNA Repair and Genome Integrity, Centre for Research in Molecular Medicine and Chronic Diseases (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [18]Department of Biochemistry and Molecular Biology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [19]Medical Research Council (MRC) Cancer Unit, University of Cambridge, Cambridge, UK. [20]Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Baltimore, MD, USA. [21]Department of Pathology, University of Cambridge, Cambridge, UK. [22]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK. [23]Biotech Research & Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark. [24]Finsen Laboratory, Rigshospitalet, Copenhagen, Denmark. [25]European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Heidelberg, Germany. [26]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [27]Department of Genetics and Informatics Institute, University of Alabama at Birmingham (UAB) School of Medicine, Birmingham, AL, USA. [28]Division of Genetics and Genomics, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA. [29]The Broad Institute of Harvard and MIT, Cambridge, MA, USA. [30]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA. [31]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [32]Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA. [33]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA. [34]Department of Computer Science, Yale University, New Haven, CT, USA. [35]Department of Human Genetics, McGill University, Montreal, Québec, Canada. [36]Department of Oncology and Onco-Hematology, University of Milan, Milan, Italy. [37]Department of Medical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. [38]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [39]Oxford NIHR Biomedical Research Centre, Oxford, UK. [40]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Baltimore, MD, USA. [41]A full list of members appears at the end of the paper. [42]Department of Haematology, University of Cambridge, Cambridge, UK. [43]A full list of members and their affiliations appears in the Supplementary Note. [44]These authors contributed equally: Eva G. Alvarez, Adrian Baez-Ortega, Jorge Zamora, Fran Supek. [45]These authors jointly supervised this work: Peter J. Campbell, Jose M. C. Tubio. *e-mail: pc8@sanger.ac.uk; jose.mc.tubio@usc.es



Supplementary Figure 1

Whole-genome sequencing coverage in tumours and matched-normal genomes in the PCAWG cohort.

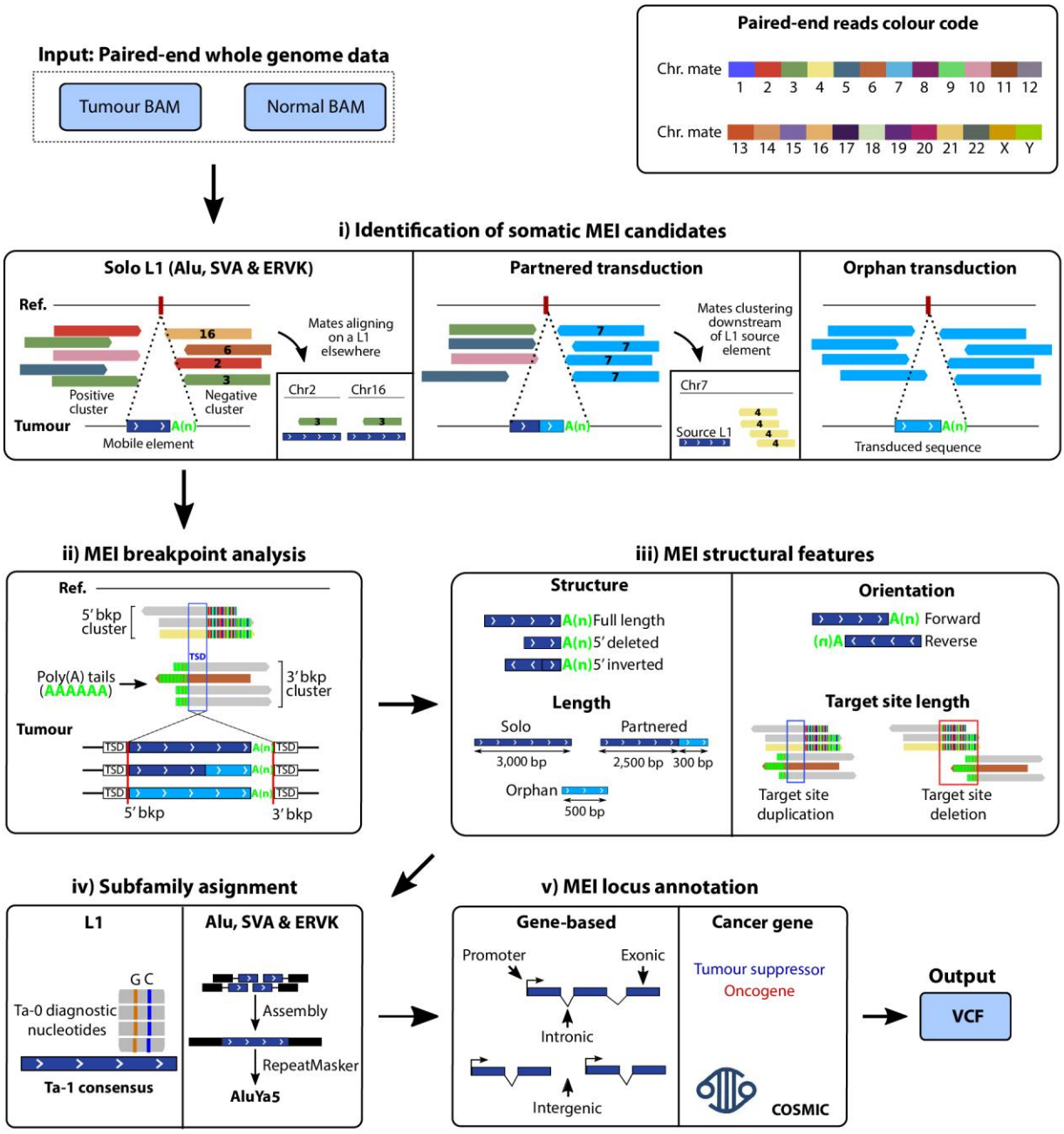
(a) Violin plot for the distribution of the mean coverage from all tumours (N=2,954) analyzed shows a bimodal distribution with maxima at 38 and 60 reads per base-pair. White point, median; box, 25th to 75th percentile (interquartile range, IQR); whiskers, data within 1.5 times the IQR. (b) Tumour samples mean depth of coverage distribution by cancer type. The number of samples per tumour type is shown in parenthesis. Violin plot features are as in 'a'. (c) Violin plot for the distribution of the mean coverage from all PCAWG matched-normal samples (N=2,834) analyzed shows a mean coverage of 30 reads per genome. Violin plot features are as in 'a'. (d) Matched-normal samples mean depth of coverage distribution by cancer type. Number of samples per tumour type is shown in parenthesis. Violin plot features are as in 'a'.



Supplementary Figure 2

Distribution of somatic retrotranspositions according to subfamily.

(a) L1 subfamilies. Ta-1 and Ta-0 elements – the youngest subfamilies of L1 retrotransposons – represent 97.5% of all L1 somatic mobilizations that were characterized to subfamily level, although we also find 367 L1 events bearing the diagnostic hallmarks of pre-Ta elements, which have been shown to retain retrotransposition activity in modern humans[5]. The category “Ta-?” contains those L1-Ta events for which it was not possible to detect the Ta-0 or Ta-1 diagnostic nucleotides. (b) Alu subfamilies. (c) SVA subfamilies.

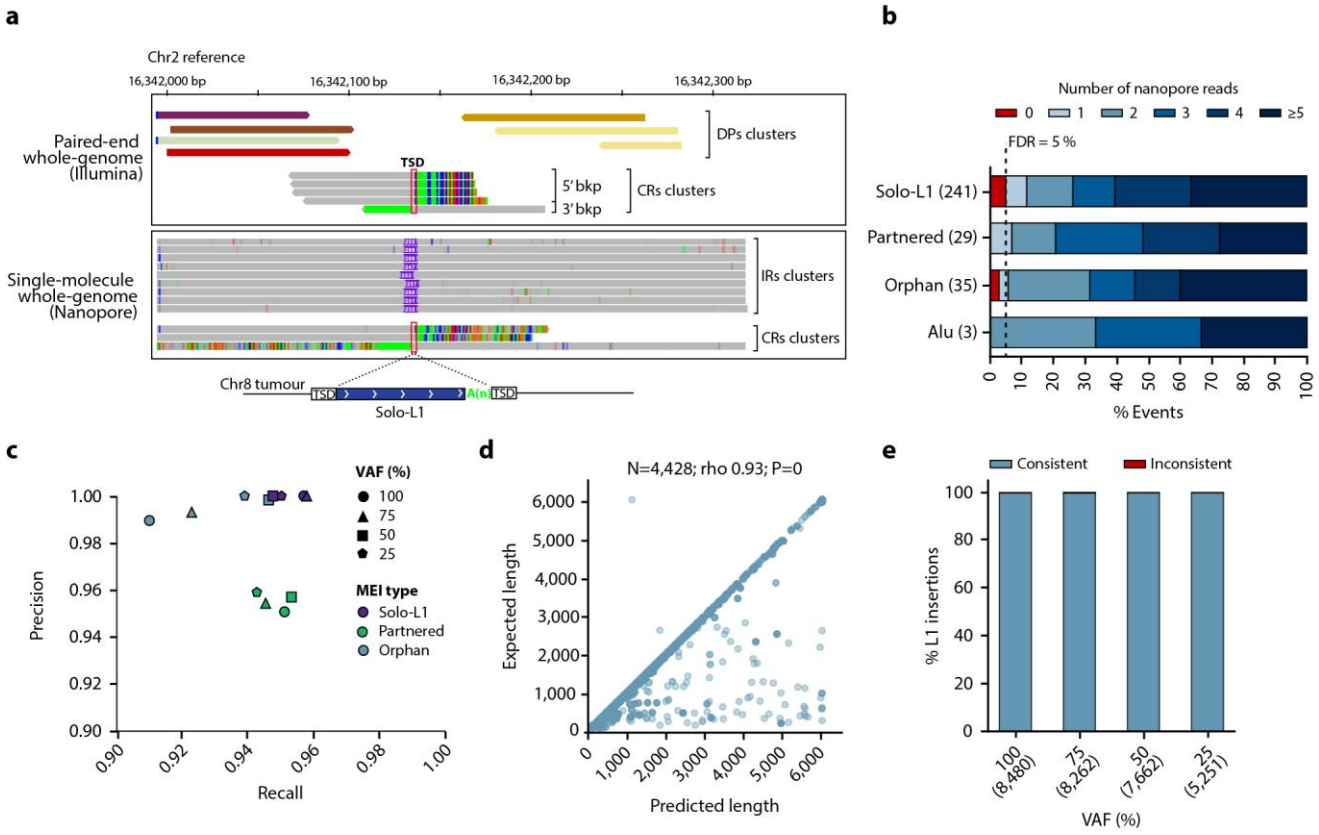


Supplementary Figure 3

Overview of TraFiC-mem.

TraFiC-mem analyzes Illumina paired-end mapping data (see **Supplementary Note**). i) Identification of candidate somatic mobile element insertions (MEIs) by TraFiC-mem. Solo-retrotransposon insertions are detected by the identification of two reciprocal clusters (positive and negative, or head-to-head) of interchromosomal reads whose mates map onto retrotransposons of the same type located elsewhere in the genome. Partnered transductions are detected by the identification of one cluster of interchromosomal reads whose mates map onto L1 retrotransposons of the same family elsewhere in the genome, and one single reciprocal cluster of reads whose mates are clusterered at a unique region adjacent to a donor source L1 element (the example illustrates a transduction from chromosome 7). Orphan transductions are detected by the identification of two reciprocal clusters whose mates map downstream to a source element as described for partnered transductions. ii) MEI breakpoint analysis. TraFiC-mem seeks for two additional clusters (5'

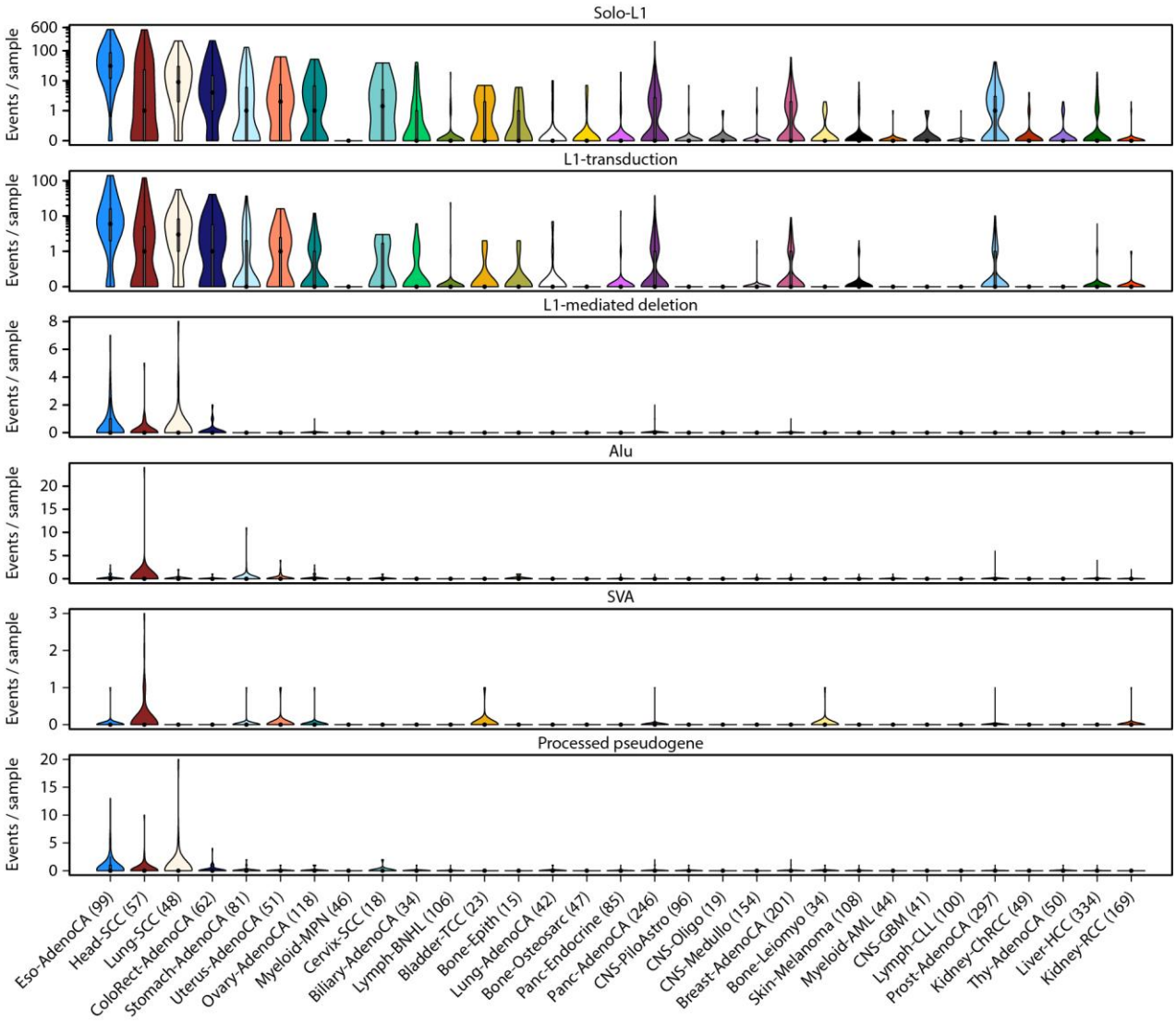
brakpoint cluster and 3’ breakpoint cluster) of clipped reads in the candidate insertion region, in order to reveal the 5’ and 3’ insertion breakpoint coordinates to base-pair resolution. iii) MEI structural features annotation. iv) Subfamily assignment. Subfamily specific diagnostic nucleotides are used to determine the subfamily for L1 events. v) MEI locus annotation: The target genomic region is annotated and MEIs inserted within cancer genes, according to the COSMIC database, are flagged. Output is a VCF file.



Supplementary Figure 4

Validation and evaluation of TraFiC-mem.

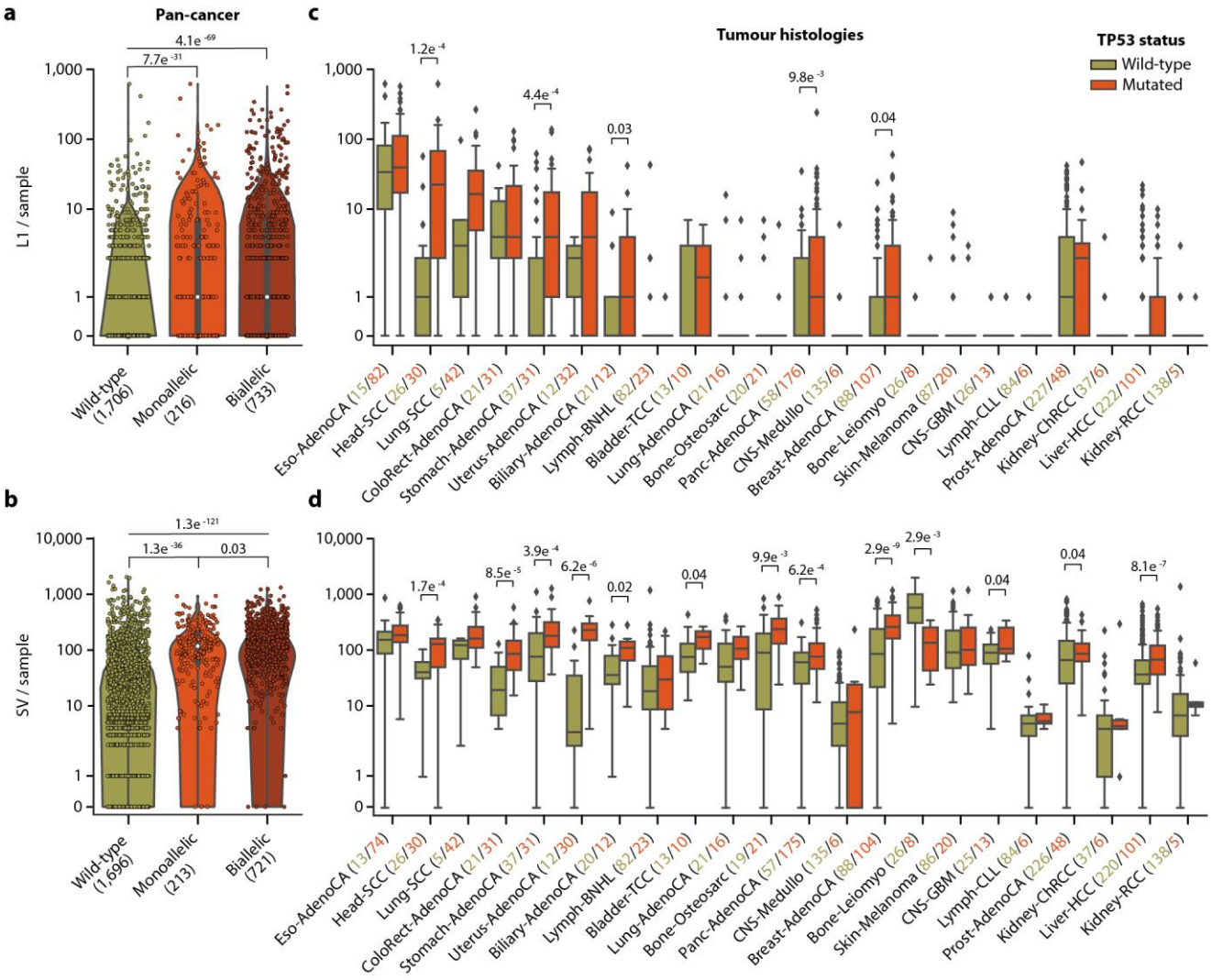
(a) Retrotransposition breakpoint validation approach using long-reads with Oxford Nanopore Technologies (ONT). Illustrative example of a Solo-L1 insertion in cancer cell-line NCI-H2087 detected with short and long-reads. Top, TraFiC-mem relies on the identification of discordant read-pairs (DPs) and clipped reads (CRs) to detect a Solo-L1 insertion using Illumina paired-end data. Bottom, indel reads (IRs) and clipped reads confirmation using ONT. (b) For each type of insertion (solo-L1, partnered transductions, orphan transductions, Alu), the proportion of events that are supported by different counts of long-reads is represented (from zero to more than 5 reads). Events supported by at least one long-read and absent in the matched-normal sample were considered true positive (i.e., somatic), while those not supported by ONT and/or present in the matched-normal sample were considered false positive. The total number of events assessed for each retrotransposition category is shown in parenthesis. (c) Precision and recall of TraFiC-mem after in-silico simulation of 10,000 L1 insertions (Solo-L1, partnered and orphan transductions) in tumors of different clonalities at 25%, 50%, 75% and 100%. (d) Plot showing the correlation between the observed and expected lengths for 8,025 Solo-L1 insertions simulated in-silico. Sample size (N), Spearman's rho and *P-value* are displayed above the panel. (e) Fraction of true positive Solo-L1 events with a predicted orientation consistent (green), and inconsistent (red), with the expected. Orientation consistency was assessed for four clonality levels (25%, 50%, 75%, 100%).



Supplementary Figure 5

Rates of somatic retrotransposition across PCAWG tumor types.

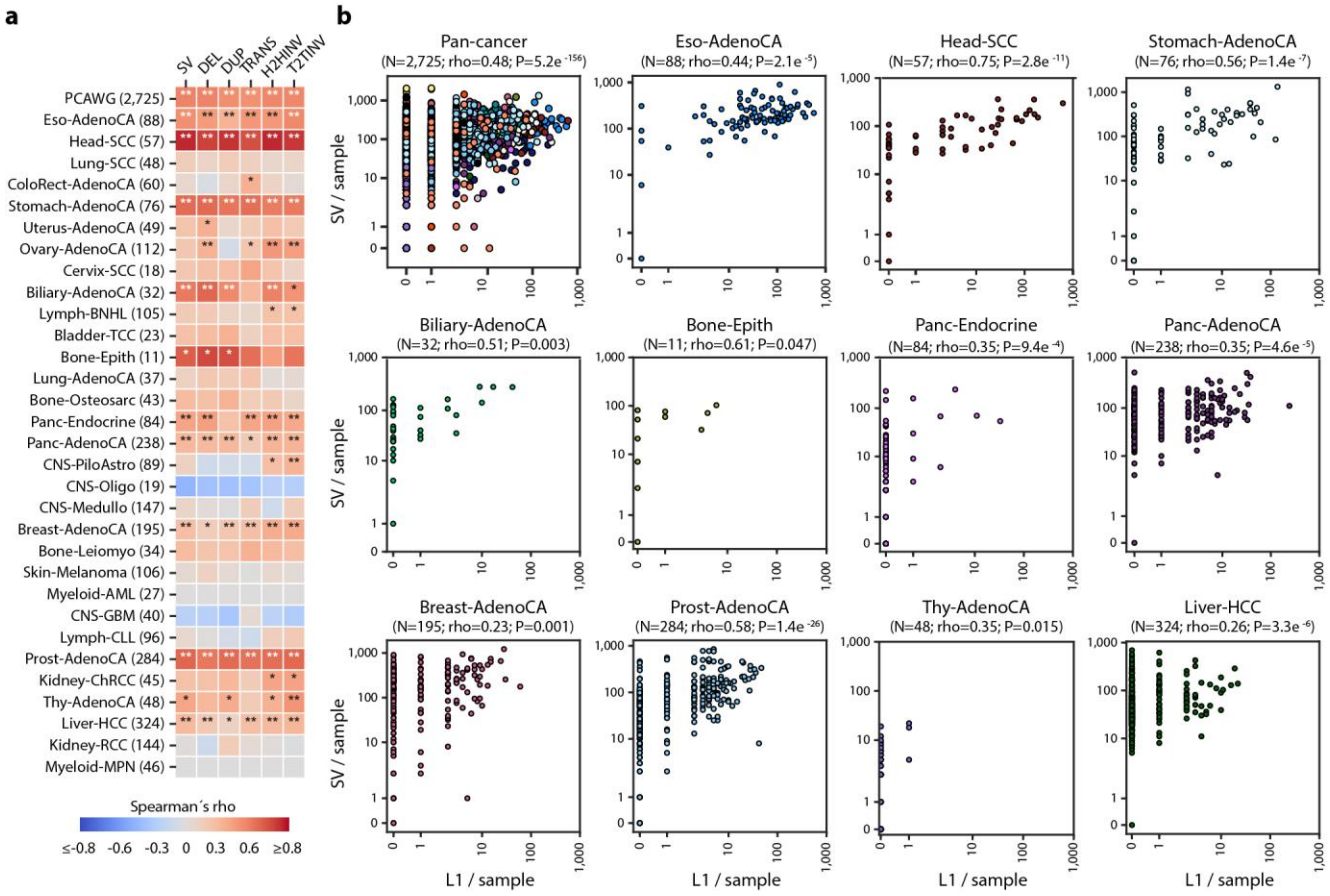
Violin plots showing the distribution of the number of retrotranspositions per sample across cancer types, for the six different categories of retrotranspositions that were analyzed (Solo-L1, L1-transductions, L1-mediated deletions, Alu, SVA and Processed pseudogenes). The number of samples per tumor type is shown in parenthesis. Y-axis is represented in a logarithmic scale. Black points, median; boxes, 25th to 75th percentile (interquartile range, IQR); whiskers, data within 1.5 times the IQR.



Supplementary Figure 6

TP53 mutation is associated with high rates of L1 retrotransposition and structural variants.

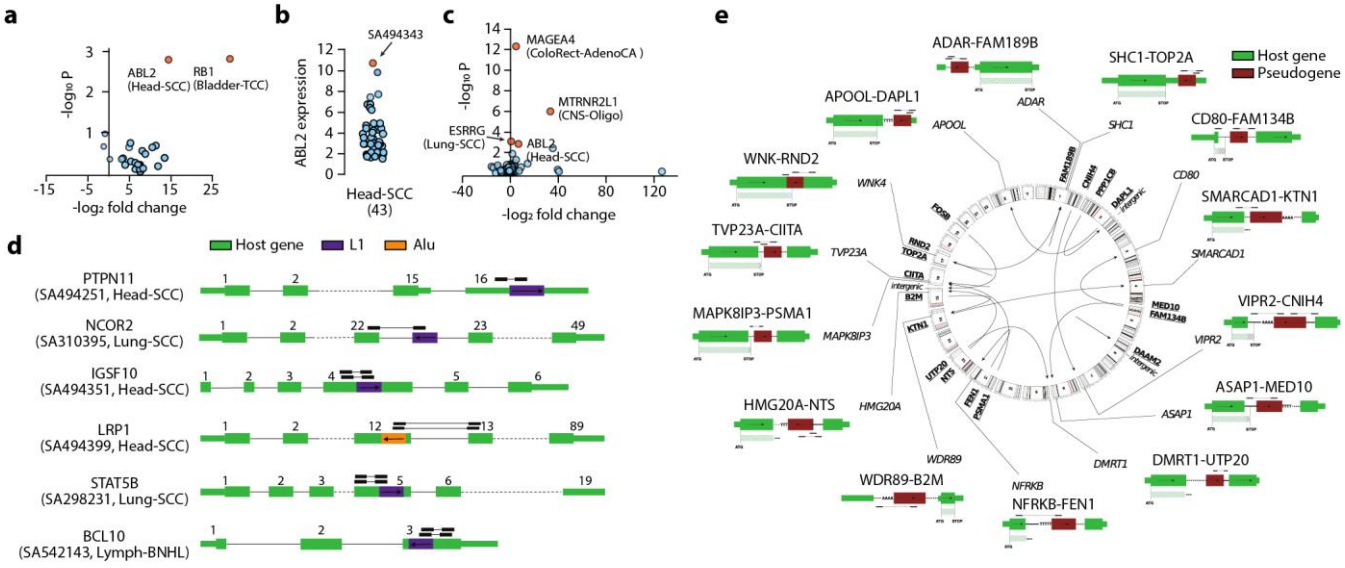
(a) Distribution of L1 counts of three sample groups according to their *TP53* mutational status: wild-type, monoallelic and biallelic driver mutation. The number of samples per group is shown within parenthesis. Point, median; box, 25th to 75th percentile (interquartile range, IQR); whiskers, data within 1.5 times the IQR. *P-value*s indicate significance from a two-tailed Mann–Whitney *U*-test. Y-axis is presented in a logarithmic scale. (b) The same for structural variant (SV) counts. (c) Distribution of L1 counts across tumor types from samples grouped in two categories: *TP53* wild-type and *TP53*-mutated (monoallelic or biallelic). The number of samples per tumor type and TP53 status (green: wild-type; orange: mutated) is shown within parenthesis. Violin plot features are as in 'a'. Outlier values outside 1.5 times the IQR are represented as diamonds. *P-value*s indicate significance from a two-tailed Mann–Whitney *U*-test. Y-axis is presented in a logarithmic scale. (d) The same for structural variant (SV) counts.



Supplementary Figure 7

Correlation between L1 retrotransposition and structural variation burden.

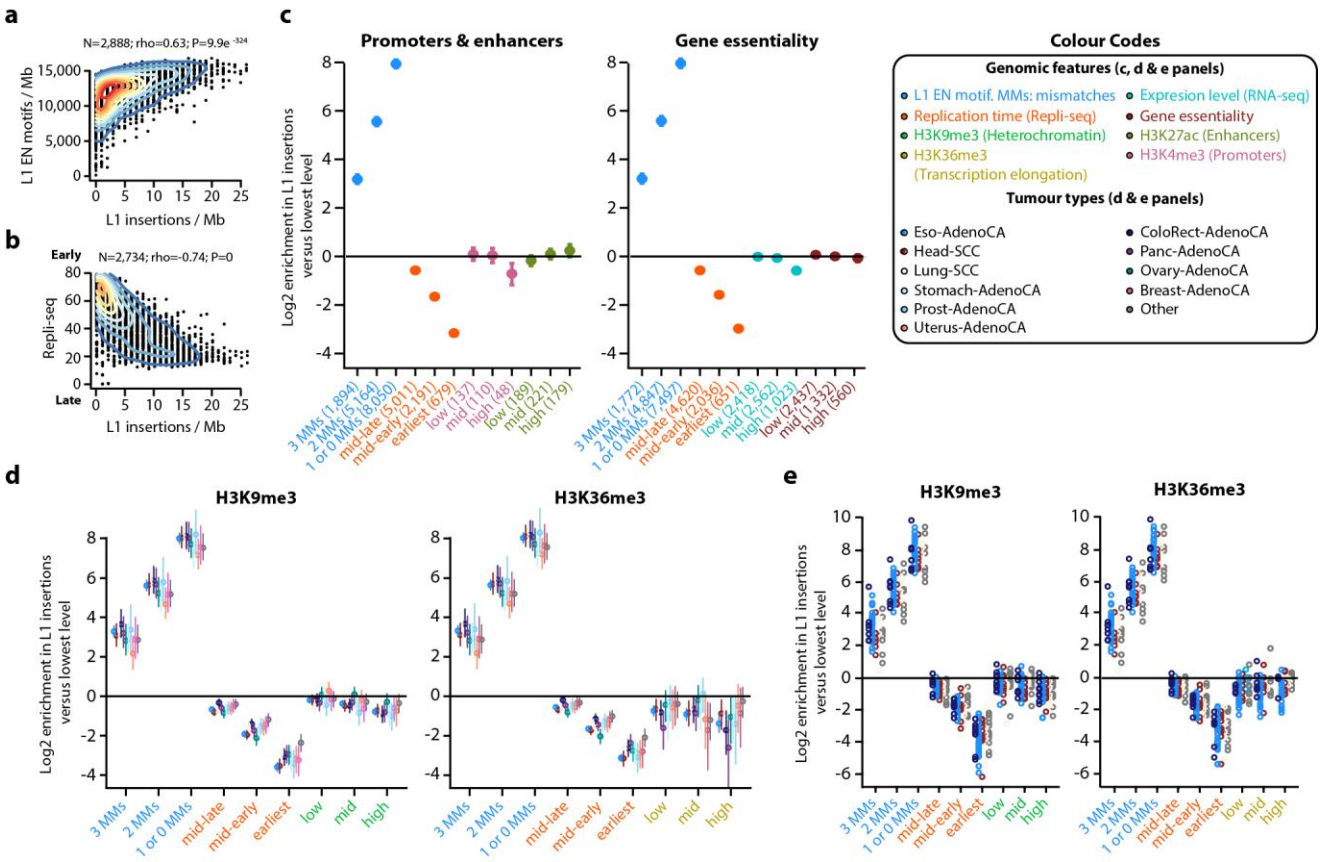
(a) Heatmap showing the correlation between the number of L1 events, the total number of structural variants (SVs) and the number of 5 different types of SVs per sample: deletions (DEL), duplications (DUP), translocations (TRANS), head-2-head inversions (H2HINV) and tail-2-tail inversions (T2TINV). Correlation was assessed both at Pan-Cancer and tumor type levels. Spearman´s correlation coefficients are shown in a blue (negative) to a red (positive) colored gradient. *P-value*s lower than 0.05 and 0.01 are represented as single and double asterisks, respectively. (b) Scatter plots showing correlations between the number of L1 events and the total number of SVs per sample at both Pan-Cancer and tumour type levels, for those comparisons that were significant in panel 'a'. The sample size (N) together with Spearman's rho and *P-value* are displayed above each chart. Both axes are displayed on a symlog scale.



Supplementary Figure 8

Some gene expression effects associated with somatic retrotransposition in PCAWG.

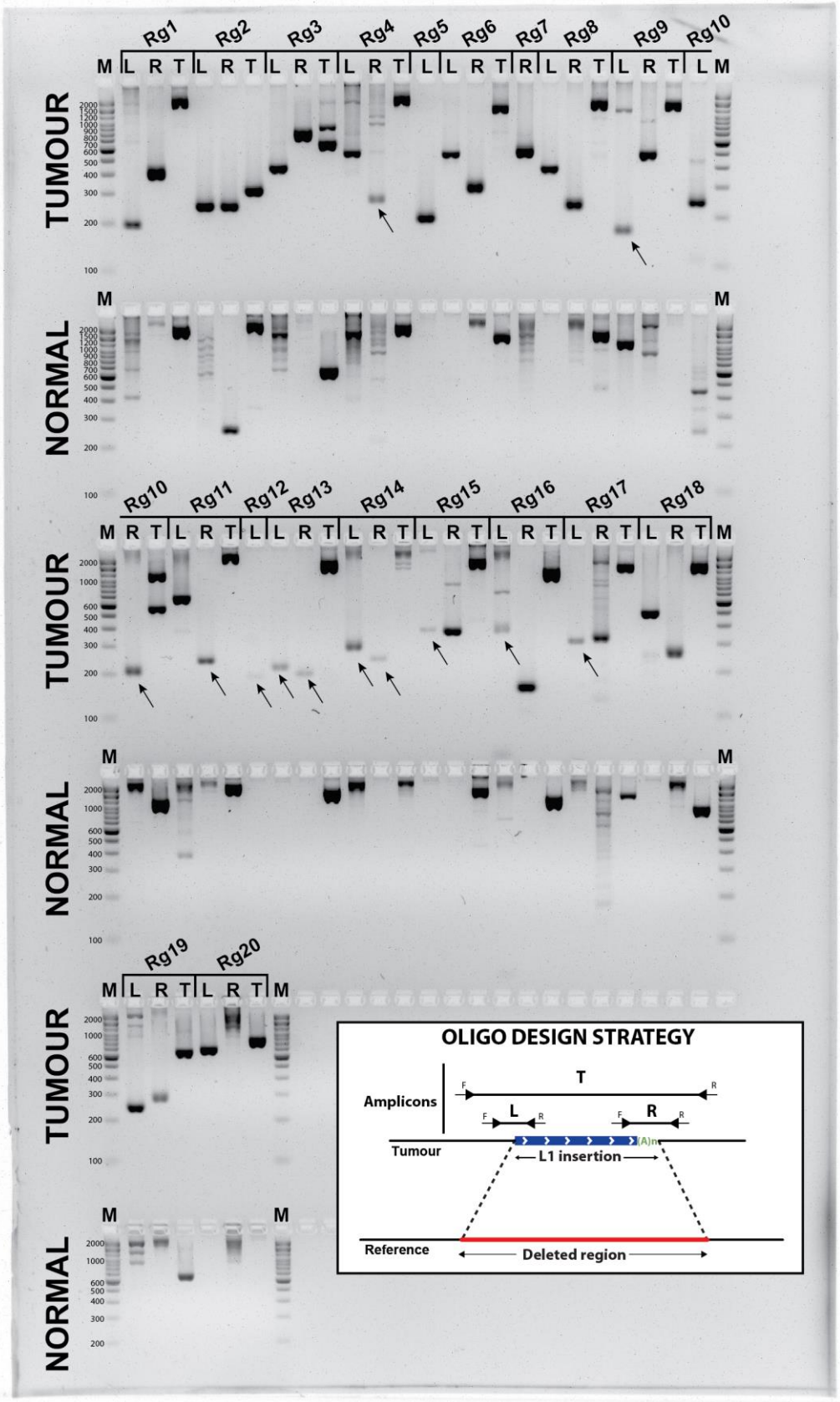
(a) Volcano plot showing the impact of L1 integration in the expression of cancer genes. Gene expression fold-change (x axis) is represented versus inverted significance (y axis). Red dots indicate significant associations under Benjamini–Hochberg adjusted *p-value*s < 0.1 from two-tailed Student's *t*-test. Adjusted *p-value*s for *ABL2* and *RB1* are 0.0017 and 0.0014, respectively. (b) Up-regulation of the *ABL2* oncogene in tumour SA494343, a head-and-neck squamous carcinoma (Head-SCC), relative to the expression of the same oncogene in other Head-SCC samples (blue) from PCAWG. Expression levels measured as Fragments Per Kilobase Million (FPKM). (c) Gene expression differences for genes with L1-retrotranspositions in promoter regions reveals significant upregulation in four genes. Volcano plot features are as in 'a'. Adjusted *p-value*s for *MTRNR2L1*, *ESRRG*, *ABL2* and *MAGEA4* are 1.1294^{e-06}, 0.0009, 0.0017 and 6.0019^{e-13}, respectively. (d) Exonization of somatic retrotranspositions, including cancer genes *PTPN11* and *NCOR2*. Green boxes are exons, thinner green blocks UTRs, lines introns and dotted lines means that a piece of the gene model is not shown for visualization purposes. Purple and orange boxes correspond to L1 and Alu integrations, respectively. Discordant read-pairs supporting fusion transcript expression are shown above the predicted transcript. (e) Host gene and processed pseudogene fusion transcripts. Arcs with arrows within the circos indicate the processed pseudogene events, connecting the source gene (underlined and bold) with the corresponding integration site. Predicted fusion transcript structures are shown in the outermost layer of the figure. Coding potential is shown underneath the fusion transcript representation. Start codon is denoted as ATG, termination codon as STOP, and uncertain termination is represented using dots.



Supplementary Figure 9

L1 integration and genomic features.

(a) Correlation between the number of L1 events and L1 endonuclease (EN) motif instances per 1 Mb-windows. The sample size (N) together with Spearman's rho and *P-value* is displayed above the plot. 2D Kernel density estimate (KDE) is displayed over the data points in a blue to red gradient. (b) Correlation between the number of somatic L1 insertions per Mb and replication timing, which is measured through Repli-seq wavelet-smoothed signal (late to early replication) and averaged per Mb. Plot features are as in 'a'. (c-e) In each panel, enrichment scores are shown, adjusted for multiple covariates and comparing the L1 insertion rate in bins 1-3 for a particular genomic feature versus bin 0 of the same feature, which therefore always has log enrichment=0 by definition and is not shown. The error bars represent 95% confidence intervals. The number of observations per bin is provided between parenthesis whenever possible. For replication time, bin 0 is the latest-replicating quarter of the genome. For essentiality, bin 0 is the non-essential genes. For the L1 motif, bin 0 denotes a non-match (4 or more mismatches). MMs stands for the number of mismatches relative to the consensus L1 EN motif. Additional information in **Supplementary Note**. (d) Association between L1 insertion rate and multiple genomic features for those tumour types with at least 100 L1 events. Data colored according to tumor type. (e) Association between L1 insertion rate and multiple genomic features in samples with at least 100 L1 events. Each data point is colored according tumour type.

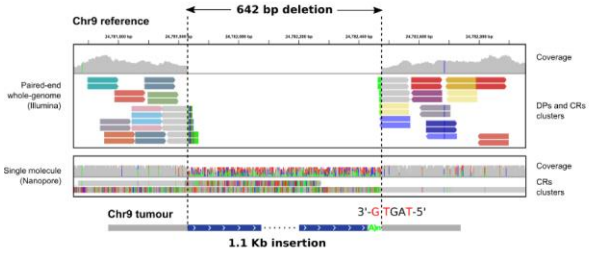

Rg1
Rg2
Rg3
Rg4
Rg5
Rg6
Rg7
Rg8
Rg9
Rg10
M L R T L R T L R T L R T L R T L R T R L R T L R T L M
TUMOUR
NORMAL
Rg10
Rg11
Rg12
Rg13
Rg14
Rg15
Rg16
Rg17
Rg18
M R T L R T L L R T L R T L R T L R T L R T L R T L R T M
Rg19
Rg20
M L R T L R T M
OLIGO DESIGN STRATEGY
Amplicons
T
L
R
F
R
Tumour
(A)n
L1 insertion
Reference
Deleted region

Supplementary Figure 10

PCR validation of somatic L1-mediated rearrangement calls.

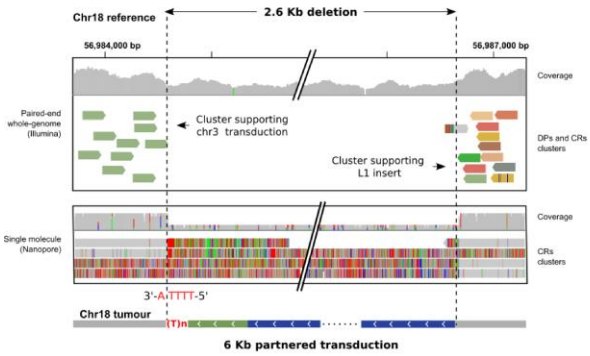
Gel showing PCR results on cancer cell-lines (NCI-H2009 and NCI-H2087) and their matched-normal cell-lines (NCI-BL2009 and NCI-BL2087). We performed validation of 20 L1-mediated rearrangements (for details, see Supplementary Table 7): 16 L1-mediated deletions (Rg1, Rg2, Rg3, Rg4, Rg6, Rg8, Rg9, Rg10, Rg11, Rg13, Rg14, Rg15, Rg16, Rg17, Rg18, Rg19), 1 L1-mediated translocation (Rg20) and 3 independent L1 breakpoints associated with a copy number change from an unknown rearrangement type (Rg5, Rg7, Rg12). For each rearrangement, except those where only one breakpoint is known, at least three regions were amplified in the tumours (see Online Methods): left breakpoint (L), right breakpoint (R), and the target site (T). Arrows are used to highlight the position of some somatic amplicons. Note that the target site could also amplify in the matched-normal sample if the deletion is not too long. “M” denotes the size marker. For illustrative purposes, the oligo design strategy is shown in a panel at the bottom of the figure: amplicons (L, R and T) and oligos – forward (F) and reverese (R) – are represented. This experiment was repeated 3 times with the same result, and results were further confirmed by single-molecule sequencing of the amplicons with ONT (see Online Methods).

a **Rg18 (NCI-H2087)**



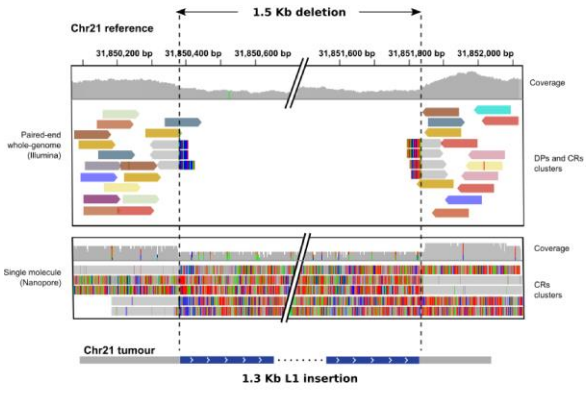

```
TGAGACCAGG ACTCTGAGCT TTAGACTAAT CATTCCCAGC ATTCCTGATT  1350
TCcAgCTCAG TACACAGGAG TTTtGGGAAT CAgCATCtaT ATAGATGCTa  1400
ggagttagCC CTCTGACCcA ACAACAAGTG GGTTCTACAT TCTTGatgag  1450
ccgctatgag actaacaagg aaaaaaggct ttgaagttgc ttatcgaaag  1500
cttaaaattt acaggctgag acaatggggt ttctagatga agcaaatcat  1550
gtcgtctgca acaggattaa ttgacttcct cttttccta attgaatacc  1600
actttatttc cttctcctgc ctgattgctt acaatgggaa ctgcttacta  1650
tgttgaatag ggcagtgaga gagggcatcc tgtcttgtgc cagtttcaaa  1700
gggaatactt ccagtttttgc ccattctggt atgatatttg gctgtgggtt  1750
gtcatagagtg gctattattt tgagatacat cccatccaat acgtcagggc  1800
catctttcta aagagaacat ttatgcagcc aaaaacacat gaagaaatgc  1850
tctcacatcg tcactggcca tcaaatgcaa atcaaaacca ctataaggat  1900
atcatctcac cagttagaat ggcaatcatt aaaagtcaag gaaacaggtg  1950
ctggaaggtt ggagaaatag ggaacacttt tactgttggt gggactgtaa  2000
atagttccat gcattgtgga agtcagtgtg gcgattcctc aagggatcta  2050
gaactaaata ccatttattt gcaagccatc ccattaccag gtatatacta  2100
aaatgagtat aaatctatgc tgctataaga cgcatgcaca cgtatgttat  2150
tcttgacact attctaataa caagacttgg aaccagccca aatgtccaac  2200
aatgatagac tggattaaag aaagatgtgg cacatatgca ccatggaata  2250
ctatgcagcc ataaaaaaatg atgagttcat atctgtaggg acatggatga  2300
agtgaaacca tcattctcag taaactactc atagaacaaa aaccagaagc  2350
acatattctc actcatggag tggggaattg aacaatgaat catggacaca  2400
ggaggggaata tcacctctgg ggactgtggg tgggaatcag gggaggagga  2450
gggatagcat tgggagatat acctgtgcta gatgacacga ttagtgggtg  2500
ccatgccacc agcatggcac atgtatacat atgtaactaa cctgctaatg  2550
tgcatgtacc ctaaaaaact taaattataa taaaataaat aaataaataa  2600
gtaGAAACAT AatGACTACA ATTATTATCC GCTAGTGTGT TTCACAGGTC  2650
ATTATTTTAA CCCTAAAAAT gCTGCTCAGA TaTCCTTCTT TAGAAACAAT  2700
GAAAAAAGTT ATAAATTTGT CATGGAGGCT CAAAACAAat tggtttttaa  2750
```

b **Rg11 (NCI-H2009)**



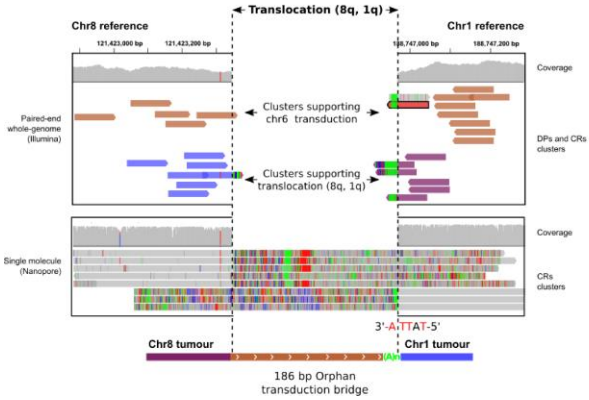

```
ttgccttaag tcttagcaac taatcatttt gctagttgca tttgactgtt  2900
tttatttttT TAAGCAATAT ATTTAcTAAT ATTTAGAAtg ttGTTCTGCA  2950
GACCCAATTA TATTTAGTTA ATTTAATTGA AATGTTTTAA ATATATTGTT  3000
TCATTGCATT GATCACTAAC TTTaAGGATT TATTTTTATT AACTTGCATA  3050
GAAAGGAGTA AAATAAACAA TAaTtATATT TTCCAGTTAA TGAAAAaata  3100
AAACAAGTAT TAGTAAAAAG CTTCTGATCT CTAAGAGTAC ATAATAGAAT  3150
TCTTTTggaG AGAATAACTG CTGTTTTGAC AAATtATATC CTTAAGACAA  3200
ATCAACAACC AAAAGACAAT CTGCCTTTtA TTTCTCAAAT GACTTTGACC  3250
TCAAATTCTG AAACTGACAG GGTGACTTTG CTCAGTGAAT AaatAATATG  3300
ATACAGAACC TTTAGAGGGA ATCGAATATG CtGAATACCC AAGgaTCTCT  3350
GTCTCTCTTc TTTTTTTATT ATAcCTCTAA GTTTTAGGGT ACATGCACAT  3400
TGTGCAGGTT AGTTACATGT ATACATGTta tgttggaagg ctTGCACtCC  3450
ACTAATGTGT CATCTAtATT AGGTggTATC TCCCAATGCT ATCCCTCtct  3500
ctcccctgact tttCCACCAC AGTCCCCAGA GTGTGATATT CCCTTCCTGT  3550
GTCCATGTGA TCTCATTGTT CAATTCCCCA CCcATGAGTG AGAATATGTG  3600
TTTGGTTTTT gtTcCcaagG ATAGTTTACT GAGAATGATG GctctattcT  3650
CATCCATGTC CCTgcCAAAG GATATGAACT CATCATcTTT ATGGCTGCAT  3700
AGcATTCCAT GGTGTATATG TGCCACATTT TCTTAATCCA GTCTATCATT  3750
GTTGGACATT TGGGTTGGTT CCAAGTCTTT GCTgTTGAGA ATAGgtgtca  3800
agAATAAACA TACGGTGCAT GTGTCTTTAT AGCAGCATGA TTTATACTCA  3850
                                   ..........
TGAGGAGGAG AGcgcCTggG CGcTTTAGAG TTTCCAGTTT TCTGTTCTGT  8450
TTTTTCCCCc ATCTTTGTTT TATCTACTTT TGTCTTTGAT GATGGTGATG  8500
TACAGATGGG TTTTCGGTGT AGATGTCCTT TCTGGTTGTT AGTTTTCCTT  8550
CTAAtagagc taggtGGACC CTCAGCTGCA GGTCTGTTGG AATACCtCTG  8600
CCGcGTGAGG TGTCAGTGTG CCCCTGtgct gtGGGTGCCT CAGTTAGGCT  8650
GCTCtcGGGG GTCAgGGAGT CAGGGACCCA CTTGAGGAGG CAGTCTGCCT  8700
GTTCTCAGAT TCCAGCTGCG TGCTGGGAGA ACCACTGCTC TCTTtccCAA  8750
AGCTGTCAGA CAGGGACACT TAAGTCTGCA GAGGTTACTG CTGTCTTTTG  8800
TTTGTCTGTG CCCTGCCCCC AGAGGTGGgA GCCTACAGAG GCAGCAGGtC  8850
TCCTTGAGtc tcGTGGTGGC TCCACCCAGT TCGAGCTTCC TGGCTGCTTT  8900
GTTTACCTgg GCAAGCCTGG GCAATGGCGG CCCTCCCAGC CcCGTTGCtG  8950
CCTTGCAGTT TGATCTCAGA CTGCTGCTAG CAATCAGCGA GACTCCctTT  9000
ttagtaatct attttaaaaa tgctgcttaa aatggtctca atatgaaagg  9050
```

c **Rg13 (NCI-H2009)**




```
TTTTTGtTTC CATTGtTTTA GTTTCTTAAT TTGAAATCAA ACTATAGTAT  12050
AGgacAAAAG GTAACACTTT TcgCTATGGA AATTATTTAT AAATAGTGTT  12100
CtTCTTTTAG ATTTACAAta tctccccaat gctatccctc cccccgaccc  12150
caccacagtc cccagagtgt gatattcctc ctgtgtccat gattcatttg  12200
ttcaattcca gagtgaggaa tatgtcattt tcatgatagt ttcatttcta  12250
gtattcacaa ttcacttcta tccatagtgt ctcctacaag gatatgaact  12300
catcattttt atggctgcat agtattccat ggtgtatatg tgctactttt  12350
aatccaaagt ctatcattat taagcatttg ggttggttcc aagtctttgc  12400
tattgtgaat agtgcgaata aaacatatgt cgtgcatgtg ttttatagca  12450
gcatgattta tactcatttg ggtatatacc cagtaatggg atggctgggt  12500
caaatggtat ttctagttct agatccctga ggaatcgcca atactgggac  12550
ttccacaatg gttgaactag ctttacagtc ccaccaacag tgtaaaagtg  12600
ttcctatttc tccgcatcct tccagcacct gttgtttcct gacttttaa  12650
tgattgctac ctttaactgg tgtggtgatg atatctcata gtggttttga  12700
tttgcatttc tctgatggcc gcgtgatgat gagcatcatt tcttcatgtt  12750
ttggctgcat aaatgtcttc tttgagaagt gtctgttgtt catgtccttt  12800
tcgcccactt ttgatgggtt gtttgttttt cttgtaaatt tgtttgagtt  12850
cattgtagat tctggatagc tacatagccc tttgtcagat gagtaggttg  12900
caaaaatttt ctcccatgtt gtaggttgcc tgttcactct tgatggtagt  12950
ttttgctgtg cagaagctct ttagtttaat tagatcccat ttgtcaattt  13000
tgtctttgtt gctatattgt tggtgttttg acatgaagtc cttgccccac  13050
gcctatgtcc tgaatggtaa tgctcaggtt ttctagggtt ttatggtttt  13100
aggtttaacg tttaaatctt taatccatct tgaattgatt tttgtataag  13150
gtacaaggaa tccagtttca gctttctaca tggctagcca gtttttccag  13200
caccatttat taaaatagggg aatccttttcc ccattgcttg ttcttttctag  13250
gttttgtcaaa gatcagatag ttgtgatatg ctgcattatt tctggagggc  13300
ctgttctgtt ccattgatct aTATCTCTTG TTttggtacc agtaacatgc  13350
tgttttggtt actgtagcct tgtagtatag tttgaagtca ggtatgtgat  13400
gcctccagcc tttgttcttt ggttaggatt tttatctctc tttgttAATA  13450
TGATATCTTt ATTTtacCTC AAGCTTCCGT TAACACAGTA AACTAGAAAC  13500
ATTaacaagA TTCCACCATG tggACAGATC TCACTGCCcc CATAAGTAAA  13550
```

d **Rg20 (NCI-H2087)**

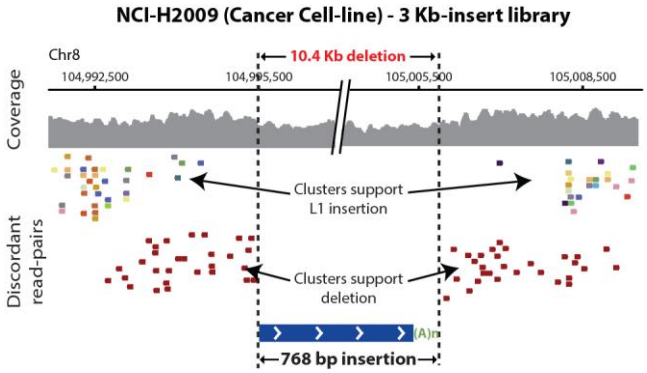



```
CGCTAGTTTT TTGTATTTTA TTAGAGATGG GTTTCACtga taaGGCCAGG  5750
tTGGTCTTGA ACTACCAACC TTAGGTGATC CACtccCCAG GCTTCCTCTC  5800
CCTCTCCTCC CTCTCCCACT CCCTCCCCTT CTCTtgGACA AGTCTTGCTC  5850
TGTTGTTCAG GCTGagaggc agtgatattc ctagaactta gctgggattg  5900
ctcatttcaa tgctactcaa agtatctgct actggcaact ctttgcctga  5950
caaatggtgt ccatgtgatc tgtgttattt tctgttctgt gcatgtgtac  6000
ctgtaacaag aatcctctct ccatgcttct tgaccaaaaa aaaaaaaaaa  6050
ttaatcactg gctcttaatt catatatgtt tatccaaaat tgaatctcaa  6100
aaatattgga gtattaggtg gggatttgct tttattacat tatcaaactg  6150
gattttacac ttacagttaa aattgtgtgc tatatctagt tgtatttcag  6200
```

Supplementary Figure 11

Single-molecule sequencing validation of somatic L1-mediated rearrangement calls.

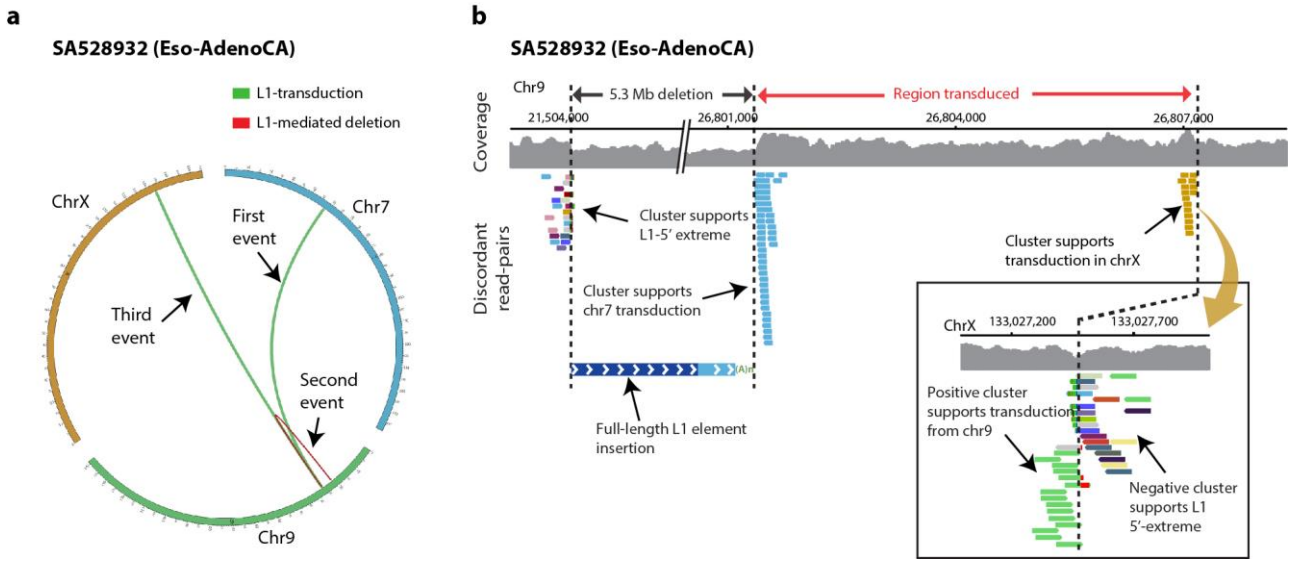
We sequenced to high-coverage (>1,000x) the PCR amplicons shown in Supplementary Fig. 10 using single-molecule sequencing with a MinION sequencer (Oxford Nanopore Technologies). We also carried out whole-genome single-molecule sequencing to low coverage of the same two tumor cell-lines (NCI-H2009 and NCI-H2087) subjected to PCR validation. For illustrative purposes, this figure only shows the validation of four representative rearrangements (Rg18, Rg11, Rg13, Rg20). The sequences of the remaining PCR amplicons can be found in Supplementary Table 7. On the left side of each panel, paired-end and Oxford Nanopore reads supporting a given rearrangement are displayed over a virtual reconstruction of the rearrangement breakpoints. On the right side of each panel, nucleotide sequence obtained by single-molecule sequencing validating each event shown in left. Nucleotide colors match those in the virtual reconstruction of the rearrangement (blue for L1, bright-green for poly-A, grey for target region, light-green for transduction). (a) Solo-L1 insertion mediating a 642 bp deletion. (b) Partnered transduction promoting a 2.6 Kb long deletion. (c) A 1.5 Kb deletion generated through an endonuclease independent L1 integration. Long reads confirm the truncation of the L1 element at its 5′ and 3′ ends. (d) Translocation between 1q31.1 and 8q24.12 mediated by an orphan transduction (same rearrangement as in **Fig. 6b**). Nanopore reads validate the orphan transduction bridge between both chromosomes.



Supplementary Figure 12

Validation of L1-mediated rearrangements in cancer cell lines by mate-pairs sequencing.

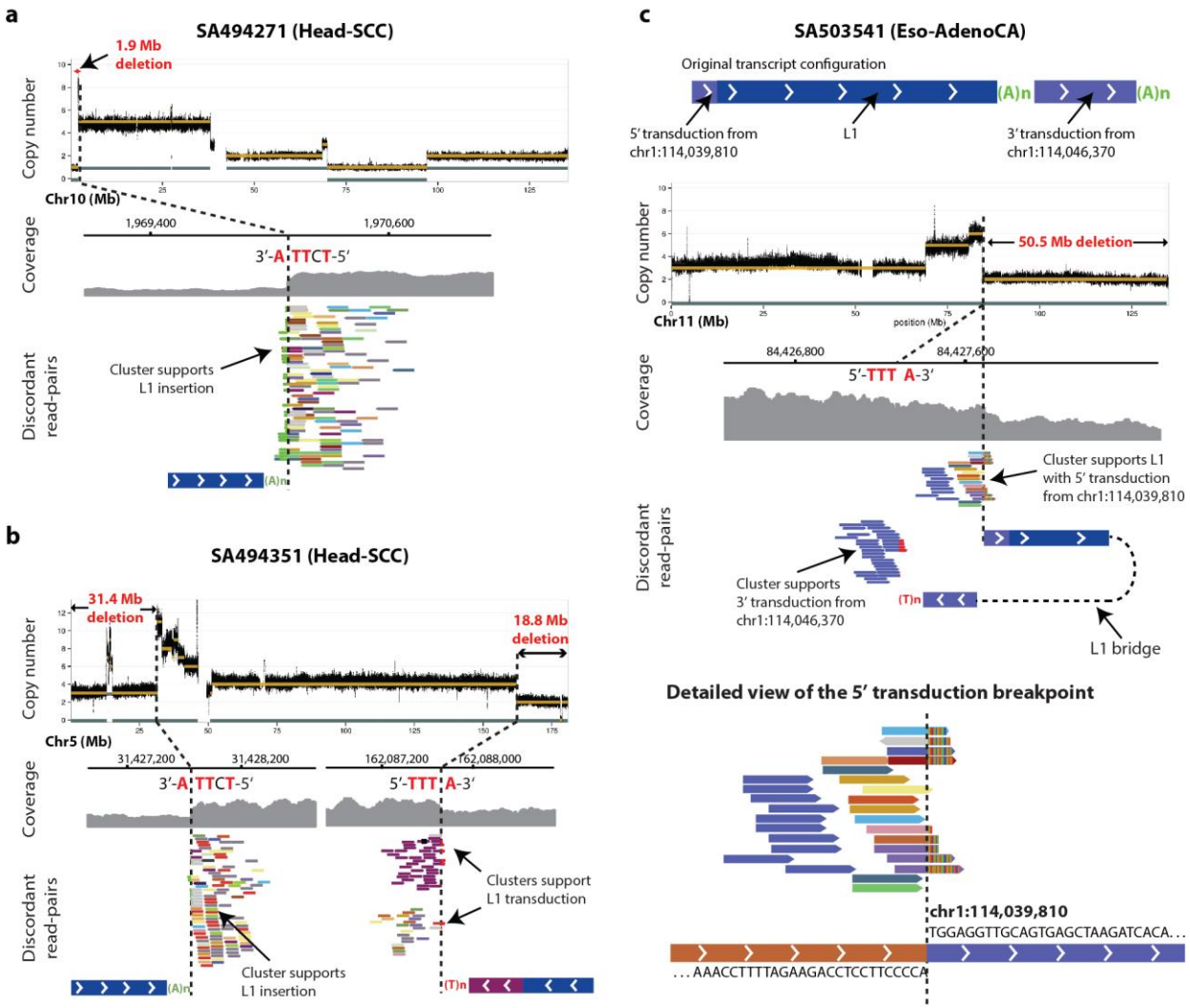
In order to further validate L1-mediated deletions, we performed mate-pair sequencing of long-inserts libraries (3 kb and 10 kb) on two cancer cell-lines with high-retrotransposition rates. In these samples, our algorithms confirmed 16 events with the hallmarks of L1-mediated deletions, in which the mate-pair data confirmed a single L1-derived (i.e., solo-L1 or L1-transduction) retrotransposition as the cause of the copy number loss, and identified the sizes of the deletion and the associated insertion. For illustrative purposes, here it is shown the validation of a 10.4 kb long deletion promoted by integration of a 768 bp L1 insertion in the cancer cell-line NCI-H2009. The L1 element inserted within the deletion breakpoints is too long to be characterized using standard paired-end sequencing libraries, but the mate-pairs successfully span the breakpoints of the deletion and confirm a single L1 insertion associated with the rearrangement.



Supplementary Figure 13

Some L1-mediated deletions are transduction-competent.

(a) Circos plot summarizing the three concatenated retrotransposition events shown in the panel b. First event, an L1 transduction mobilized from chromosome 7 is integrated into chromosome 9. Second event, this insertion concomitantly causes a 5.3 Mb deletion in the acceptor chromosome 9. Third event, the L1 element causing the deletion is subsequently able to promote a transduction that integrates into chromosome X. (b) Discordant read-pairs in chromosome 9 supports a 5.3 Mb deletion generated by the integration of a transduction from chromosome 7, and reveals an L1-event with full-length structure. Five kilobases downstream, a positive cluster of reads supports a transduction from this L1-retrotransposition event into chromosome X.



Supplementary Figure 14

Somatic integration of L1 and telomere loss.

The PCAWG-11 consensus total copy number and the copy number of the minor allele are plotted as gold and gray bands, respectively. (a) In a head-and-neck tumor, SA494271, deletion of 1.9 Mb at the short arm of chromosome 10, which involves the telomeric region, is associated with the somatic integration of an L1 retrotransposon. (b) In another head-and-neck tumor, SA494351, two independent L1 events promote deletion of both ends of chromosome 5. (c) In a Lung squamous carcinoma, SA503541, the aberrant integration of an L1 event bearing 5' and 3' transductions causes a complex rearrangement with loss of 50.5 Mb from the long arm of chromosome 11 that includes the telomere. Only the two clusters supporting both extremes of a putative L1-mediated fold-back inversion are shown. Below, a detailed view of the 5'-transduction breakpoint.

SUPPLEMENTARY NOTES

Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition

Bernardo Rodriguez-Martin[1,2,3], Eva G. Alvarez[1,2,3§], Adrian Baez-Ortega[4,§], Jorge Zamora[1,2,§], Fran Supek[5,36,§], Jonas Demeulemeester[6,7], Martin Santamarina[1,2,3], Young Seok Ju[8,11], Javier Temes[1], Daniel Garcia-Souto[1], Harald Detering[3,9,10], Yilong Li[11], Jorge Rodriguez-Castro[1], Ana Dueso-Barroso[12,13], Alicia L. Bruzos[1,2,3], Stefan C. Dentro[14,6,15], Miguel G. Blanco[16,17], Gianmarco Contino[18], Daniel Ardeljan[19], Marta Tojo[9], Nicola D. Roberts[11], Sonia Zumalave[1,2], Paul A. W. Edwards[20,21], Joachim Weischenfeldt[22,23], Montserrat Puiggròs[12], Zechen Chong[24,25], Ken Chen[24], Eunjung Alice Lee[26,27], Jeremiah A. Wala[27,28], Keiran Raine[11], Adam Butler[11], Sebastian M. Waszak[23], Fabio C. P. Navarro[29,30,31], Steven E. Schumacher[27,28], Jean Monlong[32], Francesco Maura[33,34,10], Niccolo Bolli[33,34], Guillaume Bourque[32], Mark Gerstein[29,30], Peter J. Park[35], David C. Wedge[15,11,36], Rameen Beroukhim[27,28], David Torrents[12,37], Jan O. Korbel[23], Inigo Martincorena[11], Rebecca C. Fitzgerald[18], Peter Van Loo[6,7], Haig H. Kazazian[19], Kathleen H. Burns[38,19], PCAWG Structural Variation Working Group[^], Peter J. Campbell[11,39,*] & Jose M. C. Tubio[1,2,3,11,*], PCAWG Consortium

TABLE OF CONTENTS

1. ANALYSIS OF SOMATIC RETROTRANSPOSITION

1.1. Detection of mobile element insertions using TraFiC-mem

BAM files from tumour and matched-normal pairs were processed with TraFiC-mem v1.1.0 (https://gitlab.com/mobilegenomes/TraFiC) to identify somatic mobile element insertions (MEIs) including solo-L1, L1-mediated transductions, Alu, SVA and ERV-K, using Illumina paired-end mapping data. In donors where multiple samples of primary, metastatic and/or recurrent tumours were available, each sample was independently processed and a list of non-redundant MEI calls for each donor was generated by merging the MEI calls from multiple samples as follows: those MEIs within a breakpoint offset of ±15 bp were clustered together, and the call supported by the highest number of discordant read-pairs in each cluster was selected as representative in the merging process.

TraFiC-mem starts by identifying candidate somatic MEIs via the analysis of discordant read-pairs. Contrary to previous version of the algorithm [1] , the new pipeline uses BWA-mem instead of RepeatMasker as search engine for the identification of retrotransposon-like sequences in the sequencing reads. Calls obtained at this step are preliminary, in which MEI features are outlined and insertion coordinates represent ranges surrounding the breakpoints. Then, a new module of TraFiC-mem, called MEIBA (from Mobile Element Insertion Breakpoint Analyzer) (https://github.com/brguez/MEIBA/tree/master/src/python), is used to identify the integration breakpoints to base-pair resolution, and to perform a detailed characterization of MEI features, including structure, subfamily assignment, and insertion site annotation. TraFiC-mem is illustrated in **Supplementary Fig. 3**.

1.1.1. Identification of MEI candidates via discordant reads analysis

- Identification of Solo-element events: TraFic-mem identifies reads from BWA-mem mapping that are likely to provide information pertaining to mobile elements site inclusion. Two different read-pair types are considered for the identification of insertions, named INTER_CHROM (each end of the pair are mapped to different chromosomes, where the end with the highest mapping quality (MAPQ) is considered the anchor in the reference genome), and ABERRANT (both ends of the pair are impropertly mapped to the same chromosome, where the end with the highest

MAPQ is considered the anchor). In all cases the anchor end must have a MAPQ higher than zero. The pair is excluded if any of the reads is not a primary alignment, fails platform/vendor quality checks, or is PCR or optical duplicate. Then, all non-anchor reads with an anchored mate MAPQ > 0 are interrogated for the existence of mobile element-like sequences. At this step, non-anchor reads are realigned with BWA-mem v0.7.17[1] into a database containing a set of human full-length mobile element consensus sequences, including L1, Alu, SVA and ERV-K. After BWA-mem search, all anchored reads with mates containing sequences from the same mobile element type are clustered together if (a) they have the same orientation – positive or negative – and (b) the distance relative to the nearest mapped read of the same cluster is equal or less than the average read size. Two main cluster categories are considered, namely positive and negative clusters (i.e., anchor reads mapped onto the reference genome with positive and negative orientation, respectively). Initially, a range of genome coordinates is associated with each single cluster (breakpoint coordinates are refined in a later step – see "Reconstruction of MEI breakpoints via clipped-reads analysis"). Ranges are conformed by a lower coordinate (P_L_POS and N_L_POS, respectively for positive and negative clusters) and an upper coordinate (P_R_POS and N_R_POS, for positive and negative). One positive and one negative cluster are reciprocal if P_R_POS ≥ N_L_POS and abs(N_L_POS – P_R_POS) ≤ 2(average read size). Only positive and negative clusters conformed of 4 or more reads are employed in the assignment of reciprocal clusters. Each reciprocal cluster identifies a candidate mobile element insertion (**Supplementary Fig. 3a**).

- Identification of L1 transductions: Two types of L1-mediated transductions were identified [1], namely partnered transductions, in which an L1 and downstream nonrepetitive sequence are retrotransposed together, and orphan transductions, in which only the unique sequence downstream of an active L1 is retrotransposed without the cognate L1. As above, two different read-pair types, INTERCHROM and ABERRANT are considered. In all cases the anchor read and the corresponding mate must have a MAPQ ≥ 37. Then, anchored reads are clustered together if (a) they share the same orientation, (b) the distance relative to the nearest mapped read of the cluster is equal or less than the average read size, and (c) their mates are also clustered together. Two main cluster categories are considered, namely positive and negative clusters, as above. The integration breakpoint of a potential partnered transduction is defined by reciprocal

clusters conformed of one-single cluster of (a) INTER_CHROM and/or ABERRANT reads that support an L1 and (b) one-single cluster of INTER_CHROM and/or ABERRANT reads that supports the integration of a unique DNA region from elsewhere in the genome that is located downstream to an L1 source element locus. One L1 cluster and one INTER_CHROM and/or ABERRANT cluster are reciprocal if they are in opposite orientation and P_R_POS ≥ N_L_POS and abs(N_L_POS – P_R_POS) ≤ 2(average read size). Only positive and negative clusters conformed of 4 or more reads are employed in the assignment of reciprocal clusters. Reciprocal clusters represent preliminary transduction calls that must pass the filters described below to be finally selected. The integration breakpoint of a potential orphan transduction is defined by two reciprocal clusters conformed of INTER_CHROM and/or ABERRANT reads. In this case, two clusters of INTER_CHROM and/or ABERRANT reads are reciprocal if (a) they are in opposite orientation, (b) P_R_POS ≥ N_L_POS and abs(N_L_POS – P_R_POS) ≤ 2(average read size), and (c) the two single clusters that constitute their mates are mapped within a distance of 10 kb to each other (**Supplementary Fig. 3a**).

TraFiC-mem performs the actions described above in both, the tumour and the matched normal genomes. In order to remove potential germline calls, MEI candidates are filtered out from the tumour sample if: (a) they are located within 200 bp of a cluster from the same retrotransposon family in the matched normal sample that is supported by at least 3 reads; and/or (b) there is a polymorphic insertion from the same retrotransposon family within a range of 200 bp that is present in 'TraFiC-ip db' [1], dbRIP [2], the 1,000 Genomes Project Phase 3 callset [3,4] or the dataset of germline events identified by our group across PCAWG[5]. Finally, we noticed the existence of mapping artefacts leading to quite frequent false positive insertion calls (particularly in Alu and SVA calls), located within or between repeats of the same family in the reference genome. So, an additional filter is applied to remove those insertions located within a range of ±150 bp of an element from the same family that shows ≥85% of nucleotide identity relative to the consensus sequence of the family.

1.1.2. Reconstruction of MEI breakpoints via clipped-reads analysis

A new module of TraFiC-mem, called MEIBA, is used to identify the breakpoints of an insertion to base-pair resolution. The algorithm uses two classes of reads mapped

with BWA-mem, namely soft and hard clipped reads, which overlap a putative insertion breakpoint. These reads consist of two segments, one that aligns onto the insertion target region and a second that aligns onto a mobile element elsewhere in the reference genome. Thus, once MEI candidates have been identified, TraFic-mem seeks for two additional clusters of clipped reads (CRs) that would indicate the exact insertion breakpoint coordinates (each individual insertion has two breakpoints, namely 5' and 3' breakpoints). Soft and hard CRs are extracted within a range of ±50 bp to the positive cluster P_R_POS cordinate identified in step "i" of the pipeline via discordant read-pairs. Reads marked as duplicates and reads clipped both at the beginning and ending extremes are filtered out, as they usually constitute mapping artefacts. The same approach is applied to the negative cluster N_L_POS coordinate of the same putative MEI, and both sets of reads, belonging to positive and negative clusters, are merged into a non-redundant dataset as follows: CRs are organized into clusters supporting the same breakpoint position using a maximum breakpoint offset of 3 bp. Those clusters in the tumour that are also detected in the matched-normal genome, and/or those clusters with an overrepresentation of CRs overlapping the breakpoint (we used a cut-off of more than 500 CRs), are excluded. Then, for each breakpoint cluster, supporting CRs are submitted to multiple sequence alignment using MUSCLE v3.8.31 [6], and a consensus sequence spanning the insertion breakpoints is constructed with "Cons" from the EMBOSS suit v6.6.0 [7]. The consensus sequences obtained are processed to assess if they span the target genome region and mobile element breakpoint junction (5' breakpoint), or the target region and poly(A) tail junction (3' breakpoint). If more than one 5' and/or 3' breakpoints are generated, the one supported by the highest number of CRs is selected. Finally, insertion breakpoints are required to be consistetly supported by at least two CRs, and candidate MEIs are filtered out if they do not have at least one of the two insertion breakpoints characterized to base pair resolution.

1.1.3. MEI structural features annotation

MEI structural features including insertion length, structure condition (full-length, partial, inverted), orientation, and size of target site duplication (TSD) or target site deletion, are determined for the insertions with both breakpoints successfully reconstructed. In order to compute the insertion size, the consensus sequence spanning the 5' breakpoint is realigned to the corresponding L1, Alu, SVA or ERV-K reference

sequence using Blat v34.0 [8]. Next, as retrotransposons usually only get truncated at their 5'-extreme, the insertion length is computed as the distance between the beginning of the alignment and the end of the reference sequence. Insertions spanning less than 98% of the consensus sequence for each family of retrotransposons are considered 5'-truncated, and/or if the sequence aligns in opposite orientation than the insertion DNA strand are considered 5'-inverted; otherwise, insertions are catalogued as full-length. MEIs with positive orientation are supported by clusters at the 5' and 3' breakpoints whose CRs are clipped at their ending (end-clipped) and their beginning (beg-clipped), respectively; while MEIs with negative orientation show the opposite clipping pattern. TSD and target-site deletion sizes are estimated as the distance between the two insertion breakpoints. Insertions with TSD show a breakpoint coordinate supported by the end-clipped cluster that is higher than the breakpoint coordinate supported by the beg-clipped cluster. Target site deletions show the opposite pattern.

1.1.4. MEI subfamily assignment

Two different strategies are applied to infer the subfamily of the inserted L1, Alu, SVA and ERV-K element. For L1 insertions, discordant read-pairs from supporting reciprocal clusters are realigned onto an L1 consensus sequence (GenBank identifier: L19088.1) using BWA-mem v0.7.17. The resulting SAM is converted into a binary sorted BAM file using samtools v1.7 [9]. Then, genotype likelihoods at each genomic position are computed with samtools mpileup, reference and variant sites are called with bcftools v1.7 consensus caller [10] and filtered requesting a quality score higher than 20 and a minimum read depth of 2. Finally, subfamily inference is done based on the identification of subfamily diagnostic nucleotide positions [11]: L1 integrations bearing the diagnostic "ACG" or "ACA" triplet at 5,929-5,931 position are classified as "pre-Ta" and "Ta", respectively. Ta elements are subclassified into "Ta-0" or "Ta-1" according to diagnostic bases at 5,535 and 5,538 positions (Ta-0: G and C; Ta-1: T and G). When sequencing reads do not cover the diagnostic nucleotides, subfamily cannot be inferred. For Alu, SVA and ERV-K, discordant read-pairs from reciprocal clusters supporting the insertions are assembled with velvet v1.2.10 [12], using a k-mer length of 21 bp, and the resulting contig is processed with RepeatMasker v4.0.7 to determine the subfamily. If multiple RepeatMasker hits are obtained, the one with the highest Smith-Waterman score is selected as representative. MEIs will be discarded if preliminary family assigment in "i" and subfamily assigment are not consistent.

1.1.5. MEI locus annotation

The target genomic region is annotated using the software ANNOVAR v2016-02-01 [13] and GENCODE v19 basic annotation [14]. MEIs inserted within cancer genes, according to the Cancer Gene Census COSMIC database v77 [15], are flagged.

1.1.6. TraFiC-mem output

The primary TraFiC-mem output is a standard Variant Call Format (VCF) v4.2 file containing all somatic MEI calls coordiantes with annotation features, including family, subfamily, insertion length, structural condition, orientation, size of TSD or deletion, gene annotation, number of supporting reads, and consensus sequences spanning the breakpoint junctions. Additional information is provided for L1-mediated transductions, which includes the transduced sequence length, the genomic position of the source element, and source element. MEI candidates that were filtered out are also reported together with filtering reasons.

1.1.7. TraFiC-mem availability and distribution

TraFiC-mem is implemented using Snakemake [16], a flexible Python-based workflow language, that allows to execute the pipeline from single-core workstations to computing clusters, without the need to modify the workflow. In order to enhance reproducible research, TraFiC-mem and its third party dependencies are also distributed as a Docker image (https://hub.docker.com/r/mobilegenomes/trafic). TraFiC-mem is distributed together with complete documentation and tutorial (https://gitlab.com/mobilegenomes/TraFiC).

1.2. Identification of L1-mediated deletions

Independent read clusters, identified with TraFic-mem, supporting an L1 event (i.e., clusters of discordant read-pairs with no apparent reciprocal cluster within the proximal 500 bp, and whose mates support a somatic L1 retrotransposition event) were interrogated for the presence of an associated copy number change in its proximity. Briefly, we looked for copy number loss calls from PCAWG-11 (see “Copy number dataset” above) for which the following conditions were fulfilled: (i) the upstream breakpoint matches an independent L1 cluster in positive orientation, (ii) the

corresponding downstream breakpoint, if any, from the same copy number change matches an independent L1 cluster in negative orientation, and (iii) the reconstruction of the structure of the putative insertion causing the deletion is compatible with one-single retrotransposition event. We used MEIBA – described above – to reconstruct the insertion breakpoint junctions in order to confirm the ends of the events and identify hallmarks of retrotransposition, including the poly(A) tract and duplication of the target site.

Further to the strategy described above, an additional strategy was adopted to identify L1-mediated deletions shorter than 100 kb, as follows. Coverage drops in the proximity of each independent cluster were detected by, first, normalizing read depth on each side of the cluster using the matched normal sample as a reference. Then, the ratio between the normalized read depth on both sides of the cluster was computed. This calculation was performed for window sizes ranging between 200 and 5,000 bp, with the windows on both sides of the cluster always having the same length. An immediately adjacent 'buffer' region of 300 bp was defined on each side of the cluster, and reads within these regions were omitted in read depth calculations, in order to prevent false positives due to sequence repeats at the cluster location. Subsequently, pairs of independent reciprocal (positive–negative) clusters were selected for which (i) the two clusters were located less than 100 kb apart, (ii) a potential drop in read depth ratio was identified, extending from the positive cluster to the negative cluster (statistical significance of read depth ratios was estimated non-parametrically, as described below), and (iii) the reconstruction of the structure of the putative insertion causing the deletion was compatible with a single L1 event. For each selected cluster pair, the continuity and reliability of the copy number drop was assessed by measuring the normalized read depth ratio between non-overlapping 500 bp windows spanning the region between the positive and negative clusters (i.e. within the putative deletion) and windows located upstream and downstream of the positive and negative cluster (i.e. outside the putative deletion), respectively. The significance of each drop in read depth ratio was estimated non-parametrically using a null distribution of normalized read depth ratios. This distribution was obtained for each tumour sample by randomly sampling 100,000 genomic locations, drawn from the copy number segments with the predominant copy number in that particular sample (If the predominant copy number was 1, then segments with a copy number of 2 were used instead to avoid extreme read depth ratios that could

arise from potentially undetected deletions). Specifically, read depth ratios were calculated from this sample of locations by comparing the normalized read depth between two 2,500-bp windows located immediately upstream and downstream of each location. Non-parametric *p*-values were calculated by comparing the observed read depth ratios with the ones in this null distribution, and adjusted via Benjamini–Hochberg (BH) multiple-testing correction. Three groups of output clusters were produced, corresponding to decreasing significance of the candidate deletions: first, pairs of reciprocal clusters where both clusters present an adjusted p-value below 0.1 ('Tier 1' candidates); second, pairs of reciprocal clusters where only one cluster presents an adjusted p-value below 0.1 ('Tier 2' candidates); and third, any individual clusters presenting an adjusted p-value below 0.1 ('Tier 3' candidates). The resulting L1-mediated deletion candidates (Tiers 1 and 2) were subsequently confirmed via visual inspection using the Integrative Genomics Viewer (igv) [17].

1.3. Analysis of the association between L1 insertion rate and genomic features

L1 insertion rate was calculated as the total number of somatic L1 insertions, identified across the complete PCAWG cohort, per 1-Mb window. L1 endonuclease motif density was computed as the number of canonical endonuclease motifs, here defined as TTTT|R (where R is A or G) or Y|AAAA (where Y is C or T), per 1-Mb. Bivariate correlations between L1 insertion rate, endonuclease motif density and replication timing were assessed using Spearman's rank.

To study the association of L1 insertion rate with multiple predictor variables at single-nucleotide resolution we used a statistical framework based on negative binomial regression, as described in detail previously [18]. This method was adapted herein such that originally the regression adjusted for content of trinucleotides in each genomic bin, while in this case we instead adjusted for the content of the L1 endonuclease motif. More specifically, we stratified the genome into four bins (0-3) by the closeness of match to the canonical L1 motif, here defined as TTTT|R (where R is A or G). The bin 0 contains dissimilar DNA motifs, which have 4 or more (out of 5) mismatches (MMs), encompassing 1149.7 Mb of the genome. Bin 1, 2 and 3 contain genome segments with exactly 3, exactly 2 and at most 1 MM, encompassing 749.4 Mb, 380.2 Mb and 114.1 Mb of the GRCh37 assembly, respectively. The closest match of either of the two DNA strands was considered.

Histone mark data (ChIP-Seq for H3K9me3, H3K4me3, H3K36me3, H3K27ac) and DNase hypersensitivity (DHS) data for the regional analyses was collected from Roadmap Epigenomics Consortium by averaging fold-enrichment signal over 8 cell types (E017, E114, E117, E118, E119, E122, E125 and E127) and processed by stratifying into four genomic bins, as described previously [18]. For histone marks and DHS, bin 0 are the areas of the genome with below-baseline signal (Roadmap fold-enrichment compared to input < 1), while bins 1-3 are approximately equal-sized bins covering the remaining parts of the genome with above-average fold-enrichment score. In particular, DHS bins 1-3 encompass 122.8-123.0 Mb each; for H3K36me3 129.1-136.0 Mb each; for H3K4me3 43.2-43.7 Mb each; for H3K27ac 73.6-75.1 Mb each. RNA-Seq data was also collected from Roadmap and processed as previously [18] by averaging over 8 cell types (E071, E096, E114, E117, E118, E119, E122, E127): bin 0 consisted of non-expressed genes (FPKM=0) and intergenic DNA that was not explicitly listed as expressed (total 1076.6 Mb), while bin 1 (up to 0.59 FPKM), 2 (up to 5.68 FPKM) and 3 (above 5.68 FPKM) spanned 389.9, 462.1 and 473.8 Mb of the genome, respectively. Replication time (RT) data was processed similarly as histone marks, but collected from ENCODE and processed by averaging the wavelet-smoothed signal over 8 cell types (HeLa S3, HEP G2, HUVEC, NHEK, BJ, IMR-90, MCF-7 and SK-N-SH) and then dividing into four equal-sized genomic bins (quartiles), where bin 0 is the latest-replicating and bin 3 is the earliest replicating. Essential genes were determined by CERES score based on CRISPR essentiality screens, ordering by median score across all 342 cell lines tested [19] and then stratifying genes into equal-frequency bins, from less negative to more negative median CERES score (implying commonly essential genes). For the purposes of finding L1 rates in CERES essential genes an additional 1 kb flanking the transcript was also considered together with the gene. All enrichment scores shown in plots compare bins 1-3 for a particular feature (RT, histone marks, gene expression, L1 motif) versus bin 0 of the same feature, which therefore always has log enrichment=0 by definition and is not shown on enrichment plots. The regional analyses are restricted to parts of the genome with perfect mappability scores, according to the CRG Alignability 75 track of the UCSC browser.

Further to what is reported in the main text, our analyses confirm fewer L1 events at active promoters (1.63-fold), here detected by the H3K4me3 histone mark [20], yet there

is no decrease at active enhancers, marked by H3K27ac; and we detect no significant association between gene essentiality and L1 rates (1.03-fold decrease in essential genes), suggesting that only a minor fraction of the somatic L1 events may be under negative selection (**Supplementary Fig. 9c**). Different cancer types and different samples appear remarkably consistent in the biases of L1 events towards later-replicating DNA and towards other epigenomic features examined (**Supplementary Fig. 9d-e**)

1.4. Validation of TraFiC-mem calls using single-molecule sequencing

Due to the unavailability of Pan-Cancer DNA specimens, in order to evaluate our algorithm for the identification of retrotransposon integrations, we performed validation of 308 putative somatic retrotranspositions identified with TraFiC-mem in one cancer cell-line (NCI-H2087) with high retrotransposition rate, and absent in its matched normal cell-line (NCI-BL2087) derived from blood, by single-molecule sequencing using Oxford Nanopore technology. Genomic DNA was sheared to 10 kb fragments using Covaris g-TUBEs (Covaris), cleaned with 0.4x Ampure XP Beads (Beckman Coulter Inc). After end-repairing and dA-tailing using the NEBNext End Repair/dA-tailing module (NEB), whole-genome libraries were constructed with the Oxford Nanopore Sequencing 1D ligation library prep kit (SQK-LSK108, Oxford Nanopore Technologies Ltd). We obtained four and five libraries for NCI-H2087 and NCI-BL2087, respectively. Genomic libraries were loaded on MinION R9.4 flowcells (FLO-MIN106, Oxford Nanopore Technologies Ltd), and sequencing runs were controlled using the Oxford Nanopore MinKNOW software v18.01.6. We used the Oxford Nanopore basecaller Albacore v2.0.1 to identify DNA sequences directly from raw data and generate fatsq files. Files with quality score values below 7 were excluded at this point. Minion adapter sequences were trimmed using Porechop v0.2.3 (https://github.com/rrwick/Porechop). Then, we used minimap2 v2.10-r764-dirty[21] to map sequencing reads onto the hs37d5 human reference genome, and the SAM files were converted to BAM format, sorted and indexed with Samtools v1.7 for each one of sequencing runs. BAM files were merged, sorted and indexed. After this process, sequencing coverage were 8.2x (NCI-BL2087) and 9.17X (NCI-H2087), and average read size of mapped reads were ~4.5 kb (NCI-BL2087) and ~11 kb (NCI-H2087).

Once having the whole-genome BAM files, for each one of the 308 putative somatic retrotransposition call identified with TraFic-mem, we interrogated the long-read tumour BAM file to seek for reads validating the event. Two types of MEI supporting clusters of sequencing reads were catalogued (**Supplementary Fig. 4a**), namely (i) "indel-read clusters", composed of Nanopore-reads completely spanning the insertion, so they can be identified as a standard insertion on the reference, and (ii) "clipped-read clusters", composed of Nanopore-reads spanning only one of the inserted element extremes, so they get clipped during the alignment in the reference. We observe that short MEI insertions are predominantly supported by indel-read clusters, while longer MEI insertions are mainly supported by clipped-read clusters. MEIs supported by at least one Nanopore-read in the tumour and absent in the matched-normal sample were considered true positive (TP) somatic events, while MEIs not supported by long-reads in the tumour and/or present in the matched-normal were considered false positive (FP) calls. Overall, we find 4.22% (13/308) false positive events, which showed to be particularly frequent in regions with low sequencing coverage. However, we cannot not rule out the possibility that these are true positive events, as they were not found in the matched-normal sample. False discovery rate (FDR) was estimated as follows: FDR = FP / (TP + FP).

1.5. Validation of L1-mediated rearrangements with PCR and single-molecule sequencing

Due to the unavailability of Pan-Cancer DNA specimens, we performed validation of 20 somatic L1-mediated rearrangements, mostly deletions, identified in two cancer cell-lines with high retrotransposition rates (NCI-H2009 and NCI-H2087). We carried out PCR followed by single-molecule sequencing of amplicons from the two tumour cell-lines and their matched normal samples (NCI-BL2009 and NCI-BL2087), using a Minion sequencer from Oxford Nanopore. PCR primers were designed with Primer3 v0.4.0 [22] , to amplify three regions from each event (namely, 5'-extreme, 3'-extreme and target site) as follows. For the amplification of the 3'-extreme of the event (the one that contains the poly(A) tract), we designed one forward oligo to hybridize the 3' extreme of the MEI, and a reverse oligo that hybridizes the DNA downstream. In the case of Solo-L1s, we used an L1Hs specific forward oligo matching the 3'-UTR: 5'-GGGAGATATACCTAATGCTAGATGACAC-3' [23], or an alternative forward oligo that matches other region at the 3'-extreme of the element. For the amplification of the

target site in the tumour and matched normal, we designed primers to amplify the DNA sequence between both breakpoints (5' and 3') of the rearrangement. For the amplification of the 5'-extreme of a MEI in a tumour, we designed one forward oligo to match the non-repetitive region immediately adjacent to the 5'-extreme of the element, and a reverse oligo that hybridizes the 5' extreme of the MEI.

Each PCR mixture contains 10ng of DNA, 5pmol of each primer, 5U Taq DNA polymerase (Sigma-Aldrich, catalog number D1806) with 1x Buffer containing $MgCl_2$, 0.2mM of each dNTPs, and water to a final volume of 25µl. PCR conditions were as follows: initial denaturation at 95ºC for 7 minutes; then 30-35 cycles of 95ºC for 30 seconds, 60ºC for 30 seconds, 72ºC for 45 seconds; and a final extension of 72ºC for 7 minutes. In some cases, when amplification was tricky, we used Platinum Taq High-fidelity, with a 94ºC denaturation and a 68ºC extension.

PCR amplicons were sequenced with single-molecule sequencing using a MinION from Oxford Nanopore. Amplicons were pooled and total DNA was cleaned with 0.4x AMPure XP Beads (Beckman Coulter Inc). After end-repairing and dA-tailing using the NEBNext End Repair/dA-tailing module (NEB), the sequencing library was constructed with the Oxford Nanopore Sequencing 1D ligation library prep kit (SQK-LSK108, Oxford Nanopore Technologies Ltd) and loaded on a MinION R9.4 flowcell (FLO-MIN106, Oxford Nanopore Technologies Ltd). Mapping to human reference genome was performed as described above, with minor modifications.

2. INTERPRETATION OF THE PAIRED-END MAPPING DATA FROM FIGURES

TraFiC-mem analyzes Illumina paired-end whole-genome sequencing data, aligned with BWA-mem, from a pair of tumour and matched-normal samples, to identify somatically acquired mobile element insertions (MEIs). All figures in this paper use default read colours by Integrative Genomics Viewer (igv)[17], where paired-end reads are coloured by the chromosome on which their mates can be found. Thus, retrotransposon-specific clusters (i.e., clusters of reads supporting the integration of a retrotransposon) are conformed of multicoloured reads, because mates can map

ambiguously elsewhere in the reference genome where a retrotransposon of the same family is present.

Supplementary Fig. 3 illustrates how TraFiC-mem identifies different types of candidate somatic MEIs. Briefly, the integration of a Solo-retrotransposon is detected by the identification of two reciprocal clusters (positive and negative) of interchromosomal (multicoloured) reads whose mates map onto retrotransposon of the same type located elsewhere in the genome. The integration of a partnered transduction from chromosome 7 is detected by the identification of two different types of reciprocal clusters: one cluster of multicoloured interchromosomal reads whose mates map onto L1 retrotransposons of the same family elsewhere in the genome, and one single-coloured cluster of reads whose mates are clustered at a unique region adjacent to a donor source L1 element – for illustrative purposes, this last cluster identifies a transduction from chromosome 7, and reads from this cluster are homogeneously coloured in light-blue because that is the default colour in igv for chromosome 7 –. The integration of an orphan transduction from chromosome 7 is detected by the identification of two reciprocal clusters of the same type, but this time both clusters are single-coloured because mates are clustered on the same chromosome.

3. REFERENCES

1. Tubio, J.M. *et al.* Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* **345**, 1251343 (2014).
2. Wang, J. *et al.* dbRIP: a highly integrated database of retrotransposon insertion polymorphisms in humans. *Hum Mutat* **27**, 323-9 (2006).
3. Stewart, C. *et al.* A comprehensive map of mobile element insertion polymorphisms in humans. *PLoS Genet* **7**, e1002236 (2011).
4. Sudmant, P.H. *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature* **526**, 75-81 (2015).
5. The, I.T.P.-C.A.o.W.G.N. Pan-cancer analysis of whole genomes. *Nature* (2019).
6. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-7 (2004).
7. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-7 (2000).
8. Kent, W.J. BLAT--the BLAST-like alignment tool. *Genome Res* **12**, 656-64 (2002).

9. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).
10. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-93 (2011).
11. Boissinot, S., Chevret, P. & Furano, A.V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol Biol Evol* **17**, 915-28 (2000).
12. Zerbino, D.R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821-9 (2008).
13. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
14. Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**, 1760-74 (2012).
15. Futreal, P.A. *et al.* A census of human cancer genes. *Nat Rev Cancer* **4**, 177-83 (2004).
16. Koster, J. & Rahmann, S. Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520-2 (2012).
17. Thorvaldsdottir, H., Robinson, J.T. & Mesirov, J.P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178-92 (2013).
18. Supek, F. & Lehner, B. Clustered Mutation Signatures Reveal that Error-Prone DNA Repair Targets Mutations to Active Genes. *Cell* **170**, 534-547 e23 (2017).
19. Meyers, R.M. *et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**, 1779-1784 (2017).
20. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-37 (2007).
21. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* (2018).
22. Untergasser, A. *et al.* Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**, e115 (2012).
23. Ewing, A.D. & Kazazian, H.H., Jr. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res* **20**, 1262-70 (2010).

The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes (PCAWG) Consortium Working Groups

PCAWG Steering committee

Peter J Campbell#[1,2], **Gad Getz#**[3,4,5,6], **Jan O Korbel#**[7,8], **Lincoln D Stein#**[9,10] and **Joshua M Stuart#**[11,12]

PCAWG Head of project management

Jennifer L Jennings[13]

PCAWG Executive committee

Sultan T Al-Sedairy[14], Axel Aretz[15], Cindy Bell[16], Miguel Betancourt[17], Christiane Buchholz[18], Fabien Calvo[19], Christine Chomienne[20], Michael Dunn[21], Stuart Edmonds[22], Eric Green[23], Shailja Gupta[24], Carolyn M Hutter[23], Karine Jegalian[25], Jennifer L Jennings[13], Nic Jones[26], Hyung-Lae Kim[27], Youyong Lu[28,29,30], Hitoshi Nakagama[31], Gerd Nettekoven[32], Laura Planko[32], David Scott[26], Tatsuhiro Shibata[33,34], Kiyo Shimizu[35], **Lincoln D Stein#**[9,10], Michael R Stratton[1], Takashi Yugawa[35], Giampaolo Tortora[36,37], K VijayRaghavan[24], Huanming Yang[38] and Jean C Zenklusen[39]

PCAWG Ethics and Legal Working Group

Don Chalmers#[40], Yann Joly[41], **Bartha M Knoppers#**[41], Fruzsina Molnár-Gábor[42], Mark Phillips[41], Adrian Thorogood[41] and David Townend[43]

PCAWG Technical Working Group

Brice Aminou[44], Javier Bartolome[45], Keith A Boroevich[46,47], Rich Boyce[7], Alvis Brazma[7], Angela N Brooks[3,11,12,48], Alex Buchanan[49], Ivo Buchhalter[50,51,52], Adam P Butler[1], Niall J Byrne[44], Andy Cafferkey[7], Peter J Campbell[1,2], Zhaohong Chen[53], Sunghoon Cho[54], Wan Choi[55], Peter Clapham[1], Brandi N Davis-Dusenbery[56], Francisco M De La Vega[57,58,59], Jonas Demeulemeester[60,61], Michelle T Dow[53], Lewis Jonathan Dursi[9,62], Juergen Eils[63,64], Roland Eils[50,52,63,64], Kyle Ellrott[49], Claudiu Farcas[53], Nodirjon Fayzullaev[44], Vincent Ferretti[44,65], Paul Flicek[7], Nuno A Fonseca[7,66], Josep L L Gelpi[45,67], Gad Getz[3,4,5,6], Robert L Grossman[68], Olivier Harismendy[69,70], Allison P Heath[71], Michael C Heinold[50,52], Julian M Hess[3,72], Oliver Hofmann[73], Jongwhi H Hong[74], Thomas J Hudson[75,76], Barbara Hutter[77,78,79], Carolyn M Hutter[23], Daniel Hübschmann[52,63,80,81,82], Seiya Imoto[83], Sinisa Ivkovic[56], Seung-Hyup Jeon[55], Wei Jiao[9], Jongsun Jung[84], Rolf Kabbe[50], Andre Kahles[85,86,87,88,89], Jules NA Kerssemakers[50], Hyung-Lae Kim[27], Hyunghwan Kim[55], Jihoon Kim[90], Youngwook Kim[91,92],

Kortine Kleinheinz[50,52], Jan O Korbel[7,8], Michael Koscher[93], Antonios Koures[53], Milena Kovacevic[56], Chris Lawerenz[64], Ignaty Leshchiner[3], Jia Liu[94], Dimitri Livitz[3], George L Mihaiescu[44], Sanja Mijalkovic[56], Ana Mijalkovic Mijalkovic-Lazic[56], Satoru Miyano[83], Naoki Miyoshi[83], Hardeep K Nahal-Bose[44], Hidewaki Nakagawa[47], Mia Nastic[56], Steven J Newhouse[7], Jonathan Nicholson[1], **Brian D O'Connor#**[44,95], David Ocana[7], Kazuhiro Ohi[83], Lucila Ohno-Machado[53], Larsson Omberg[96], BF Francis Ouellette[44,97], Nagarajan Paramasivam[50,78], Marc D Perry[44,98], Todd D Pihl[99], Manuel Prinz[50], Montserrat Puiggròs[45], Petar Radovic[56], Keiran M Raine[1], Esther Rheinbay[3,6,100], Mara Rosenberg[3,100], Romina Royo[45], Gunnar Rätsch[85,86,87,88,89,101], Gordon Saksena[3], Matthias Schlesner[50,102], Solomon I Shorser[9], Charles Short[7], Heidi J Sofia[23], Jonathan Spring[103], **Lincoln D Stein#**[9,10], Adam J Struck[104], Grace Tiao[3], Nebojsa Tijanic[56], David Torrents[45,105], Peter Van Loo[60,61], Miguel Vazquez[45,106], David Vicente[45], Jeremiah A Wala[3,6,48], Zhining Wang[39], Sebastian M Waszak[8], Joachim Weischenfeldt[8,107,108], Johannes Werner[50,109], Ashley Williams[53], Youngchoon Woo[55], Adam J Wright[9], Qian Xiang[44], **Sergei Yakneen#**[8], Liming Yang[39], Denis Yuen[9], **Christina K Yung#**[44] and **Junjun Zhang#**[44]

PCAWG Reference Annotations Working Group

Angela N Brooks[3,11,12,48], Ivo Buchhalter[50,51,52], Peter J Campbell[1,2], Priyanka Dhingra[110,111], Lars Feuerbach[112], Mark Gerstein[113,114,115], Gad Getz[3,4,5,6], Mark P Hamilton[116], Henrik Hornshøj[117], Todd A Johnson[46], Andre Kahles[85,86,87,88,89], Abdullah Kahraman[118,119,120], Manolis Kellis[3,121], **Ekta Khurana#**[110,111,122,123], Jan O Korbel[7,8], Morten Muhlig Nielsen[117], Jakob Skou Pedersen[117,124], Paz Polak[3,4,6], Jüri Reimand[9,125], Esther Rheinbay[3,6,100], Nicola D Roberts[1], Gunnar Rätsch[85,86,87,88,89,101], Richard Sallari[3], Nasa Sinnott-Armstrong[3,59], Alfonso Valencia[45,105], Miguel Vazquez[45,106], Sebastian M Waszak[8], Joachim Weischenfeldt[8,107,108] and Christian von Mering[120,126]

PCAWG Quality-Control Working Group

Sergi Beltran[127,128], Ivo Buchhalter[50,51,52], Peter J Campbell[1,2], Roland Eils[50,52,63,64], Daniela S Gerhard[129], Gad Getz[3,4,5,6], **Ivo G Gut#**[127,128], Marta Gut[127,128], Barbara Hutter[77,78,79], Daniel Hübschmann[52,63,80,81,82], Kortine Kleinheinz[50,52], Jan O Korbel[7,8], Dimitri Livitz[3], Marc D Perry[44,98], Keiran M Raine[1], Esther Rheinbay[3,6,100], Mara Rosenberg[3,100], Gordon Saksena[3], Matthias Schlesner[50,102], Miranda D Stobbe[127,128], Jean-Rémi Trotta[127], Johannes Werner[50,109] and Justin P Whalley[127]

PCAWG SNV Calling Working Group

Matthew H Bailey[130,131], Beifang Niu[132], Matthias Bieg[78,133], Paul C Boutros[9,125,134,135], Ivo Buchhalter[50,51,52], Adam P Butler[1], Ken Chen[136], Zechen Chong[137], **Li Ding#**[130,131,138], Oliver Drechsel[128,139], Lewis Jonathan Dursi[9,62], Roland Eils[50,52,63,64], Kyle Ellrott[49], Shadrielle MG Espiritu[9], Yu Fan[140], Robert S Fulton[130,131,138], Shengjie Gao[38], Josep L L Gelpi[45,67], Mark Gerstein[113,114,115], Gad Getz[3,4,5,6], Santiago Gonzalez[7,8], Ivo G Gut[127,128], Faraz Hach[141,142], Michael C Heinold[50,52], Julian M Hess[3,72], Jonathan Hinton[1], Taobo Hu[143], Vincent Huang[9], Yi Huang[144,145], Barbara Hutter[77,78,79],

David R Jones[1], Jongsun Jung[84], Natalie Jäger[50], Hyung-Lae Kim[27], Kortine Kleinheinz[50,52], Sushant Kumar[114,115], Yogesh Kumar[143], Christopher M Lalansingh[9], Ignaty Leshchiner[3], Ivica Letunic[146], Dimitri Livitz[3], Eric Z Ma[143], Yosef E Maruvka[3,72,100], R Jay Mashl[131,147], Michael D McLellan[130,131,138], Andrew Menzies[1], Ana Milovanovic[45], Morten Muhlig Nielsen[117], Stephan Ossowski[128,139,148], Nagarajan Paramasivam[50,78], Jakob Skou Pedersen[117,124], Marc D Perry[44,98], Montserrat Puiggròs[45], Keiran M Raine[1], Esther Rheinbay[3,6,100], Romina Royo[45], S Cenk Sahinalp[142,149,150], Gordon Saksena[3], Iman Sarrafi[142,150], Matthias Schlesner[50,102], **Jared T Simpson#**[9,151], Lucy Stebbings[1], Chip Stewart[3], Miranda D Stobbe[127,128], Jon W Teague[1], Grace Tiao[3], David Torrents[45,105], Jeremiah A Wala[3,6,48], Jiayin Wang[131,145,152], Wenyi Wang[140], Sebastian M Waszak[8], Joachim Weischenfeldt[8,107,108], Michael C Wendl[131,138,153], Johannes Werner[50,109], David A Wheeler[154,155], Zhenggang Wu[143], Hong Xue[143], Sergei Yakneen[8], Takafumi N Yamaguchi[9], Kai Ye[152,156], Venkata D Yellapantula[138,157], Christina K Yung[44] and Junjun Zhang[44]

PCAWG Drivers and Functional Interpretation Working Group

Federico Abascal[1], Samirkumar B Amin[158,159,160], Gary D Bader[10], Pratiti Bandopadhayay[3,161,162], Jonathan Barenboim[9], Rameen Beroukhim[3,6,163], Johanna Bertl[117,164], Keith A Boroevich[46,47], Søren Brunak[165,166], Peter J Campbell[1,2], Joana Carlevaro-Fita[167,168,169], Dimple Chakravarty[170], Calvin Wing Yiu Chan[50,171], Ken Chen[136], Jung Kyoon Choi[172], Jordi Deu-Pons[173,174], Priyanka Dhingra[110,111], Klev Diamanti[175], Lars Feuerbach[112], J Lynn Fink[45,176], Nuno A Fonseca[7,66], Joan Frigola[173], Carlo Gambacorti-Passerini[177], Dale W Garsed[178], **Mark Gerstein#**[113,114,115], **Gad Getz#**[3,4,5,6], Qianyun Guo[124], Ivo G Gut[127,128], David Haan[11], Mark P Hamilton[116], Nicholas J Haradhvala[3,100], Arif O Harmanci[115,179], Mohamed Helmy[180], Carl Herrmann[50,52,181], Julian M Hess[3,72], Asger Hobolth[124,164], Ermin Hodzic[150], Chen Hong[112,171], Henrik Hornshøj[117], Keren Isaev[9,125], Jose MG Izarzugaza[182], Rory Johnson[168,183], Todd A Johnson[46], Malene Juul[117], Randi Istrup Juul[117], Andre Kahles[85,86,87,88,89], Abdullah Kahraman[118,119,120], Manolis Kellis[3,121], Ekta Khurana[110,111,122,123], Jaegil Kim[3], Jong K Kim[184], Youngwook Kim[91,92], Jan Komorowski[175,185], Jan O Korbel[7,8], Sushant Kumar[114,115], Andrés Lanzós[168,169,183], Erik Larsson[85], **Michael S Lawrence#**[3,46,100], Donghoon Lee[115], Kjong-Van Lehmann[85,86,87,88,89], Shantao Li[115], Xiaotong Li[115], Ziao Lin[3,186], Eric Minwei Liu[110,111,187], Lucas Lochovsky[114,115,160], Shaoke Lou[114,115], Tobias Madsen[117], Kathleen Marchal[188,189], Iñigo Martincorena[1], Alexander Martinez-Fundichely[110,111,123], Yosef E Maruvka[3,72,100], Patrick D McGillivray[114], William Meyerson[115,190], Ferran Muiños[174,191], Loris Mularoni[174,191], Hidewaki Nakagawa[47], Morten Muhlig Nielsen[117], Marta Paczkowska[9], Keunchil Park[192,193], Kiejung Park[194], **Jakob Skou Pedersen#**[117,124], Tirso Pons[195], Sergio Pulido-Tamayo[188,189], **Benjamin J Raphael#**[196], Jüri Reimand[9,125], Iker Reyes-Salazar[191], Matthew A Reyna[196], Esther Rheinbay[3,6,100], Mark A Rubin[183,197,198,199,200], Carlota Rubio-Perez[174,191,201], S Cenk Sahinalp[142,149,150], Gordon Saksena[3], Leonidas Salichos[114,115], Chris Sander[85,202,203], Steven E Schumacher[3,204], Mark Shackleton[178], Ofer Shapira[3,205], Ciyue Shen[203,206], Raunak Shrestha[142], Shimin Shuai[9,10], Nikos Sidiropoulos[108], Lina Sieverling[112,171], Nasa Sinnott-Armstrong[3,59], Lincoln D Stein[9,10], **Joshua M Stuart#**[11,12], David Tamborero[174,191], Grace Tiao[3], Tatsuhiko Tsunoda[46,207,208,209], Husen M Umer[175,210], Liis Uusküla-Reimand[211,212], Alfonso Valencia[45,105], Miguel Vazquez[45,106], Lieven PC Verbeke[189,213], Claes Wadelius[214], Lina Wadi[9], Jiayin Wang[131,145,152], Jonathan Warrell[114,115], Sebastian M Waszak[8], Joachim Weischenfeldt[8,107,108], **David A Wheeler#**[154,155], Guanming Wu[215], Jun Yu[216], Jing Zhang[115], Xuanping Zhang[145,217], Yan Zhang[115,218,219], Zhongming Zhao[220], Lihua Zou[221] and Christian von

Mering[120,126]

PCAWG Transcriptome Working Group

Samirkumar B Amin[158,159,160], Philip Awadalla[9,10], Peter J Bailey[222], **Alvis Brazma#**[7], **Angela N Brooks#**[3,11,12,48], Claudia Calabrese[7,8], Aurélien Chateigner[44], Isidro Cortés-Ciriano[223,224,225], Brian Craft[12], David Craft[3,226], Chad J Creighton[227], Natalie R Davidson[85,86,87,88,101], Deniz Demircioğlu[228,229], Serap Erkek[8], Nuno A Fonseca[7,66], Milana Frenkel-Morgenstern[230], Mary J Goldman[12], Liliana Greger[7], Jonathan Göke[228,231], Yao He[232], Katherine A Hoadley[233,234], Yong Hou[38,235], Matthew R Huska[236], Andre Kahles[85,86,87,88,89], Ekta Khurana[110,111,122,123], Helena Kilpinen[237], Jan O Korbel[7,8], Fabien C Lamaze[9], Kjong-Van Lehmann[85,86,87,88,89], Chang Li[38,235], Siliang Li[38,235], Xiaobo Li[38,235], Xinyue Li[38], Dongbing Liu[38,235], Fenglin Liu[232,238], Xingmin Liu[38,235], Maximillian G Marin[11], Julia Markowski[236], Matthew Meyerson[3,6,48], Tannistha Nandi[239], Morten Muhlig Nielsen[117], Akinyemi I Ojesina[240,241,242], BF Francis Ouellette[44,97], Qiang Pan-Hammarström[38,243], Peter J Park[223,225], Chandra Sekhar Pedamallu[3,6,163], Jakob Skou Pedersen[117,124], Marc D Perry[44,98], **Gunnar Rätsch#**[85,86,87,88,89,101], Roland F Schwarz[7,81,236,244], Yuichi Shiraishi[83], Reiner Siebert[245,246], Cameron M Soulette[11], Stefan G Stark[86,88,247,248], Oliver Stegle[7,8,249], Hong Su[38,235], Patrick Tan[239,250,251,252], Bin Tean Teh[250,251,252,253,254], Lara Urban[7,8], Jian Wang[38], Sebastian M Waszak[8], Kui Wu[38,235], Qian Xiang[44], Heng Xiong[38,235], Sergei Yakneen[8], Huanming Yang[38], Chen Ye[38,235], Christina K Yung[44], Fan Zhang[232], Junjun Zhang[44], Xiuqing Zhang[38], Zemin Zhang[232,255], Liangtao Zheng[232], Jingchun Zhu[12] and Shida Zhu[38,235]

PCAWG Epigenome Working Group

Hiroyuki Aburatani[256], **Benjamin P Berman#**[257,258,259], Hans Binder[260,261], **Benedikt Brors#**[79,112,262], Huy Q Dinh[257], Lars Feuerbach[112], Shengjie Gao[38], Ivo G Gut[127,128], Simon C Heath[127,128], Steve Hoffmann[260,261,263,264], Charles David Imbusch[112], Ekta Khurana[110,111,122,123], Helene Kretzmer[261,264], Peter W Laird[265], Jose I Martin-Subero[105,266], Genta Nagae[256,267], **Christoph Plass#**[268], Paz Polak[3,4,6], Hui Shen[269], Reiner Siebert[245,246], Nasa Sinnott-Armstrong[3,59], Miranda D Stobbe[127,128], Qi Wang[93], Dieter Weichenhan[268], Sergei Yakneen[8] and Wanding Zhou[269]

PCAWG Structural Variation Working Group

Kadir C Akdemir[136], Eva G Alvarez[270,271,272], Adrian Baez-Ortega[273], **Rameen Beroukhim#**[3,6,163], Paul C Boutros[9,125,134,135], David D L Bowtell[178], Benedikt Brors[79,112,262], Kathleen H Burns[274,275], John Busanovich[3,276], **Peter J Campbell#**[1,2], Kin Chan[277], Ken Chen[136], Isidro Cortés-Ciriano[223,224,225], Ana Dueso-Barroso[45], Andrew J Dunford[3], Paul A Edwards[278,279], Xavier Estivill[139,280], Dariush Etemadmoghadam[178], Lars Feuerbach[112], J Lynn Fink[45,176], Milana Frenkel-Morgenstern[230], Dale W Garsed[178], Mark Gerstein[113,114,115], Dmitry A Gordenin[281], David Haan[11], James E Haber[282], Julian M Hess[3,72], Barbara Hutter[77,78,79], Marcin Imielinski[283,284], David TW Jones[285,286], Young Seok Ju[1,172], Marat D Kazanov[287,288,289], Leszek J Klimczak[290], Youngil Koh[291,292], Jan O Korbel[7,8], Kiran Kumar[3], Eunjung Alice Lee[293], Jake June-Koo Lee[223,225], Yilong Li[1], Andy G Lynch[278,279,294], Geoff

Macintyre[278], Florian Markowetz[278,279], Iñigo Martincorena[1], Alexander Martinez-Fundichely[110,111,123], Satoru Miyano[83], Hidewaki Nakagawa[47], Fabio CP Navarro[114], Stephan Ossowski[128,139,148], Peter J Park[223,225], John V Pearson[295,296], Montserrat Puiggròs[45], Karsten Rippe[81], Nicola D Roberts[1], Steven A Roberts[297], Bernardo Rodriguez-Martin[270,271,272], Steven E Schumacher[3,204], Ralph Scully[298], Mark Shackleton[178], Nikos Sidiropoulos[108], Lina Sieverling[112,171], Chip Stewart[3], David Torrents[45,105], Jose MC Tubio[270,271,272], Izar Villasante[45], Nicola Waddell[295,296], Jeremiah A Wala[3,6,48], Joachim Weischenfeldt[8,107,108], Lixing Yang[299], Xiaotong Yao[284,300], Sung-Soo Yoon[292], Jorge Zamora[1,270,271,272] and Cheng-Zhong Zhang[3,301]

PCAWG Mutational Signatures Working Group

Ludmil B Alexandrov[1,70,302], Erik N Bergstrom[70,303], Arnoud Boot[251,304], Paul C Boutros[9,125,134,135], Kin Chan[277], Kyle Covington[155], Akihiro Fujimoto[47], Gad Getz[3,4,5,6], Dmitry A Gordenin[281], Nicholas J Haradhvala[3,100], Mi Ni Huang[251,304], S. M. Ashiqul Islam[53], Marat D Kazanov[287,288,289], Jaegil Kim[3], Leszek J Klimczak[290], Michael S Lawrence[3,46,100], Iñigo Martincorena[1], John R McPherson[251,304], Sandro Morganella[1], Ville Mustonen[305,306,307], Hidewaki Nakagawa[47], Alvin Wei Tian Ng[308], Serena Nik-Zainal[1,309,310,311], Paz Polak[3,4,6], Stephenie D Prokopec[9], Steven A Roberts[297], **Steven G Rozen#**[251,252,304], Radhakrishnan Sabarinathan[174,191,312], Natalie Saini[281], Tatsuhiro Shibata[33,34], Yuichi Shiraishi[83], **Michael R Stratton#**[1], **Bin Tean Teh#**[250,251,252,253,254], Ignacio Vázquez-García[1,157,313,314], Yang Wu[251,304], Fouad Yousif[9] and Willie Yu[315]

PCAWG Germline Cancer Genome Working Group

Ludmil B Alexandrov[1,70,302], Eva G Alvarez[270,271,272], Adrian Baez-Ortega[273], Matthew H Bailey[130,131], Mattia Bosio[45,128,139], G Steven Bova[316], Alvis Brazma[7], Alicia L Bruzos[270,271,272], Ivo Buchhalter[50,51,52], Carlos D Bustamante[58,59], Atul J Butte[317], Andy Cafferkey[7], Claudia Calabrese[7,8], Peter J Campbell[1,2], Stephen J Chanock[318], Nilanjan Chatterjee[319,320], Jieming Chen[115,321], Francisco M De La Vega[57,58,59], Olivier Delaneau[322,323,324], German M Demidov[128,139,148], Anthony DiBiase[325], Li Ding[130,131,138], Oliver Drechsel[128,139], Lewis Jonathan Dursi[9,62], Douglas F Easton[326,327], Serap Erkek[8], Georgia Escaramis[139,328,329], **Xavier Estivill#**[139,280], Erik Garrison[1], Mark Gerstein[113,114,115], Gad Getz[3,4,5,6], Dmitry A Gordenin[281], Nina Habermann[8], Olivier Harismendy[69,70], Eoghan Harrington[330], Shuto Hayashi[83], Seong Gu Heo[331], José María Heredia-Genestar[332], Aliaksei Z Holik[139], Eun Pyo Hong[331], Xing Hua[318], Kuan-lin Huang[131,333], Seiya Imoto[83], Sissel Juul[330], Ekta Khurana[110,111,122,123], Hyung-Lae Kim[27], Youngwook Kim[91,92], Leszek J Klimczak[290], **Jan O Korbel#**[7,8], Roelof Koster[334], Sushant Kumar[114,115], Ivica Letunic[146], Yilong Li[1], Tomas Marques-Bonet[105,127,332,335], R Jay Mashl[131,147], Simon Mayes[336], Michael D McLellan[130,131,138], Lisa Mirabello[318], Francesc Muyas[128,139,148], Hidewaki Nakagawa[47], Arcadi Navarro[105,127,332], Steven J Newhouse[7], Stephan Ossowski[128,139,148], Ji Wan Park[331], Esa Pitkänen[8], Aparna Prasad[128], Raquel Rabionet[128,139,337], Benjamin Raeder[8], Tobias Rausch[8], Steven A Roberts[297], Bernardo Rodriguez-Martin[270,271,272], Vasilisa A Rudneva[8], Gunnar Rätsch[85,86,87,88,89,101], Natalie Saini[281], Matthias Schlesner[50,102], Roland F Schwarz[7,81,236,244], Ayellet V Segre[3,338], Tal Shmaya[57], Suyash S Shringarpure[59], Nikos Sidiropoulos[108], Reiner Siebert[245,246], Jared T Simpson[9,151], Lei Song[318], Oliver Stegle[7,8,249], Hana Susak[128,139], Tomas J Tanskanen[339], Grace Tiao[3], Marta Tojo[272], Jose MC

Tubio[270,271,272], Daniel J Turner[336], Lara Urban[7,8], Sebastian M Waszak[8], David C Wedge[1,340,341], Joachim Weischenfeldt[8,107,108], David A Wheeler[154,155], Mark H Wright[59], Dai-Ying Wu[57], Tian Xia[342], Sergei Yakneen[8], Kai Ye[152,156], Venkata D Yellapantula[138,157], Jorge Zamora[1,270,271,272] and Bin Zhu[318]

PCAWG Pathology and Clinical Correlates Working Group

Fatima Al-Shahrour[343], Gurnit Atwal[9,10,344], Peter J Bailey[222], **Andrew V Biankin#**[222,345,346,347], Paul C Boutros[9,125,134,135], Peter J Campbell[1,2], David K Chang[222,345], Susanna L Cooke[222], Vikram Deshpande[100], Bishoy M Faltas[101], William C Faquin[100], **Levi Garraway#**[48], Gad Getz[3,4,5,6], **Sean M Grimmond#**[348], Syed Haider[9], **Katherine A Hoadley#**[233,234], Wei Jiao[9], Vera B Kaiser[349], Rosa Karlić[350], Mamoru Kato[351], Kirsten Kübler[3,6,100], Alexander J Lazar[158,352], Constance H Li[9,125], David N Louis[100], Adam A Margolin[104], Sancha Martin[1,353], Hardeep K Nahal-Bose[44], G Petur Nielsen[100], Serena Nik-Zainal[1,309,310,311], Larsson Omberg[96], Christine P'ng[9], Marc D Perry[44,98], Paz Polak[3,4,6], Esther Rheinbay[3,6,100], Mark A Rubin[183,197,198,199,200], Colin A Semple[349], Dennis C Sgroi[100], Tatsuhiro Shibata[33,34], Reiner Siebert[245,246], Jaclyn Smith[354], **Lincoln D Stein#**[9,10], Miranda D Stobbe[127,128], Ren X Sun[9], Kevin Thai[44], Derek W Wright[222,355], Chin-Lee Wu[100], Ke Yuan[278,353,356] and Junjun Zhang[44]

PCAWG Evolution & Heterogeneity Working Group

David J Adams[1], Pavana Anur[357], Rameen Beroukhim[3,6,163], Paul C Boutros[9,125,134,135], David D L Bowtell[178], Peter J Campbell[1,2], Shaolong Cao[140], Elizabeth L Christie[178], Marek Cmero[358,359,360], Yupeng Cun[361], Kevin J Dawson[1], Jonas Demeulemeester[60,61], Stefan C Dentro[1,60,340], Amit G Deshwar[362], Nilgun Donmez[142,150], Ruben M Drews[278], Roland Eils[50,52,63,64], Yu Fan[140], Matthew W Fittall[60], Dale W Garsed[178], Moritz Gerstung[7,8], Gad Getz[3,4,5,6], Santiago Gonzalez[7,8], Gavin Ha[3], Kerstin Haase[60], Marcin Imielinski[283,284], Lara Jerman[8,363], Yuan Ji[364,365], Clemency Jolly[60], Kortine Kleinheinz[50,52], Juhee Lee[366], Henry Lee-Six[1], Ignaty Leshchiner[3], Dimitri Livitz[3], Geoff Macintyre[278], Salem Malikic[142,150], Florian Markowetz[278,279], Iñigo Martincorena[1], Thomas J Mitchell[1,279,367], Quaid D Morris[344,368], Ville Mustonen[305,306,307], Layla Oesper[369], Martin Peifer[361], Myron Peto[357], Benjamin J Raphael[196], Daniel Rosebrock[3], Yulia Rubanova[151,344], S Cenk Sahinalp[142,149,150], Adriana Salcedo[9], Matthias Schlesner[50,102], Steven E Schumacher[3,204], Subhajit Sengupta[370], Ruian Shi[368], Seung Jun Shin[248], **Paul T Spellman#**[357], Oliver Spiro[3], Lincoln D Stein[9,10], Maxime Tarabichi[1,60], **Peter Van Loo#**[60,61], Shankar Vembu[368,371], Ignacio Vázquez-García[1,157,313,314], Wenyi Wang[140], **David C Wedge#**[1,340,341], David A Wheeler[154,155], Jeffrey A Wintersinger[151,180,344], Tsun-Po Yang[361], Xiaotong Yao[284,300], Kaixian Yu[372], Ke Yuan[278,353,356] and Hongtu Zhu[372,373]

PCAWG Portals and Visualization Working Group

Fatima Al-Shahrour[343], Elisabet Barrera[7], Wojciech Bazant[7], Alvis Brazma[7], Isidro Cortés-Ciriano[223,224,225], Brian Craft[12], David Craft[3,226], Vincent Ferretti[44,65], Nuno A Fonseca[7,66],

Anja Füllgrabe[7], Mary J Goldman[12], **David Haussler#**[12,374], Wolfgang Huber[8], Maria Keays[7], Alfonso Muñoz[7], Brian D O'Connor[44,95], Irene Papatheodorou[7], Robert Petryszak[7], Elena Piñeiro-Yáñez[343], Alfonso Valencia[45,105], **Miguel Vazquez#**[45,106], John N Weinstein[375,376], Qian Xiang[44], Junjun Zhang[44] and **Jingchun Zhu#**[12]

PCAWG Mitochondrial Genome and Immunogenomics Working Group

Peter J Campbell[1,2], Yiwen Chen[140], Chad J Creighton[227], Li Ding[130,131,138], Akihiro Fujimoto[47], Masashi Fujita[47], Gad Getz[3,4,5,6], Leng Han[217], Takanori Hasegawa[83], Shuto Hayashi[83], Seiya Imoto[83], Young Seok Ju[1,172], Hyung-Lae Kim[27], Youngwook Kim[91,92], Youngil Koh[291,292], Mitsuhiro Komura[83], Jun Li[140], **Han Liang#**[140], Iñigo Martincorena[1], Satoru Miyano[83], Shinichi Mizuno[377], **Hidewaki Nakagawa#**[47], Keunchil Park[192,193], Eigo Shimizu[83], Yumeng Wang[140], John N Weinstein[375,376], Yanxun Xu[378], Rui Yamaguchi[83], Fan Yang[368], Yang Yang[217], Christopher J Yoon[172], Sung-Soo Yoon[292], Yuan Yuan[140], Fan Zhang[232] and Zemin Zhang[232,255]

PCAWG Pathogens Working Group

Malik Alawi[379,380], Ivan Borozan[9], Daniel S Brewer[381,382], Colin S Cooper[382,383,384], Nikita Desai[44], Roland Eils[50,52,63,64], Vincent Ferretti[44,65], Adam Grundhoff[380,385], Murat Iskar[386], Kortine Kleinheinz[50,52], **Peter Lichter#**[77,386], Hidewaki Nakagawa[47], Akinyemi I Ojesina[240,241,242], Chandra Sekhar Pedamallu[3,6,163], Matthias Schlesner[50,102], Xiaoping Su[352] and **Marc Zapatka#**[386]

Providers of tumour-sequencing data

PCAWG Tumour-specific providers (ovarian cancer) in Australia

Kathryn Alsop[178], Australian Ovarian Cancer Study Group[295,387,388], **David D L Bowtell#**[178], Timothy JC Bruxner[176], Angelika N Christ[176], Elizabeth L Christie[178], Stephen M Cordner[389], Prue A Cowin[387], Ronny Drapkin[390], Dariush Etemadmoghadam[178], Sian Fereday[178], Dale W Garsed[178], Joshy George[160], Sean M Grimmond[348], Anne Hamilton[387], Oliver Holmes[295,296], Jillian A Hung[391], Karin S Kassahn[176,392], Stephen H Kazakoff[295,296], Catherine J Kennedy[391,393], Conrad R Leonard[295,296], Linda Mileshkin[178], David K Miller[176,345], Gisela Mir Arnau[387], Chris Mitchell[178], Felicity Newell[295,296], Katia Nones[295,296], Ann-Marie Patch[295,296], John V Pearson[295,296], Michael C Quinn[295,296], Mark Shackleton[178], Darrin F Taylor[176], Heather Thorne[178], Nadia Traficante[178], Ravikiran Vedururu[387], Nick M Waddell[296], Nicola Waddell[295,296], Paul M Waring[394], Scott Wood[295,296], Qinying Xu[295,296] and Anna deFazio[391,393,395]

PCAWG Tumour-specific providers (pancreatic cancer) in Australia

Matthew J Anderson[176], Davide Antonello[396], Andrew P Barbour[397,398], Claudio Bassi[396], Samantha Bersani[399], **Andrew V Biankin#**[222,345,346,347], Timothy JC Bruxner[176], Ivana Cataldo[399,400], David K Chang[222,345], Lorraine A Chantrill[345,401], Yoke-Eng Chiew[391,393], Angela Chou[345,393], Angelika N Christ[176], Sara Cingarlini[36], Nicole Cloonan[402], Vincenzo Corbo[400,403], Maria Vittoria Davi[404], Fraser R Duthie[222,405], J Lynn Fink[45,176], Anthony J Gill[345,406], Janet S Graham[222,407], **Sean M Grimmond#**[348], Ivon Harliwong[176], Oliver Holmes[295,296], Nigel B Jamieson[222,347,408], Amber L Johns[345], Karin S Kassahn[176,392], Stephen H Kazakoff[295,296], James G Kench[345,406,409], Luca Landoni[396], Rita T Lawlor[400], Conrad R Leonard[295,296], Andrea Mafficini[400], Neil D Merrett[396,410], David K Miller[176,345], Marco Miotto[396], Elizabeth A Musgrove[222], Adnan M Nagrial[345], Felicity Newell[295,296], Katia Nones[295,296], Karin A Oien[394,411], Marina Pajic[345], Ann-Marie Patch[295,296], John V Pearson[295,296], Mark Pinese[345], Andreia V Pinho[412], Michael C Quinn[295,296], Alan J Robertson[176], Ilse Rooman[345], Borislav C Rusev[400], Jaswinder S Samra[396,413], Maria Scardoni[399], Christopher J Scarlett[345,414], Aldo Scarpa[400], Elisabetta Sereni[396], Katarzyna O Sikora[400], Michele Simbolo[403], Morgan L Taschuk[44], Christopher W Toon[345], Giampaolo Tortora[36,37], Caterina Vicentini[400], Nick M Waddell[296], Nicola Waddell[295,296], Scott Wood[295,296], Jianmin Wu[345], Qinying Xu[295,296] and Nikolajs Zeps[415,416]

PCAWG Tumour-specific providers (skin cancer) in Australia

Lauri A Aaltonen[417], Andreas Behren[418], Hazel Burke[419], Jonathan Cebon[418], Rebecca A Dagg[420], Ricardo De Paoli-Iseppi[419], Ken Dutton-Regester[295], Matthew A Field[421], Anna Fitzgerald[422], Sean M Grimmond[348], **Nicholas K Hayward#**[295,419], Peter Hersey[419], Oliver Holmes[295,296], Valerie Jakrot[419], Peter A Johansson[295], Hojabr Kakavand[419], Stephen H Kazakoff[295,296], Richard F Kefford[423], Loretta MS Lau[424], Conrad R Leonard[295,296], Georgina V Long[419], **Graham J Mann#**[393,419,419], Felicity Newell[295,296], Katia Nones[295,296], Ann-Marie Patch[295,296], John V Pearson[295,296], Hilda A Pickett[424], Antonia L Pritchard[295], Gulietta M Pupo[393], Robyn PM Saw[419], Sarah-Jane Schramm[393], **Richard A Scolyer#**[409,413,419], Mark Shackleton[178], Catherine A Shang[422], Ping Shang[419], Andrew J Spillane[419], Jonathan R Stretch[419], Varsha Tembe[393], John F Thompson[419], Ricardo E Vilain[425], Nick M Waddell[296], Nicola Waddell[295,296], James S Wilmott[419], Scott Wood[295,296], Qinying Xu[295,296] and Jean Y Yang[426]

PCAWG Tumour-specific providers (pancreatic cancer) in Canada

John Bartlett[427,428], Prashant Bavi[429], Ivan Borozan[9], Dianne E Chadwick[430], Michelle Chan-Seng-Yue[429], Sean Cleary[429,431], Ashton A Connor[431,432], Karolina Czajka[76], Robert E Denroche[429], Neesha C Dhani[433], Jenna Eagles[76], Vincent Ferretti[44,65], Steven Gallinger[429,431,432], Robert C Grant[429,432], David Hedley[433], Michael A Hollingsworth[434], **Thomas J Hudson#**[75,76], Gun Ho Jang[429], Jeremy Johns[76], Sangeetha Kalimuthu[429], Sheng-Ben Liang[430], Ilinca Lungu[429,435], Xuemei Luo[9], Faridah Mbabaali[76], **John D McPherson#**[76,429,436], Treasa A McPherson[432], Jessica K Miller[76], Malcolm J Moore[433], Faiyaz Notta[429,437], Danielle Pasternack[76], Gloria M Petersen[438], Michael H A Roehrl[125,429,430,439,440,441], Michelle Sam[76], Iris Selander[432], Stefano Serra[394], Sagedeh Shahabi[430], **Lincoln D Stein#**[9,10], Morgan L Taschuk[44], Sarah P Thayer[434], Lee E Timms[76], Gavin W Wilson[9,429], Julie M Wilson[429] and Bradly G Wouters[125]

PCAWG Tumour-specific providers (prostate cancer) in Canada

Timothy A Beck[44,442], Vinayak Bhandari[9], **Paul C Boutros#**[9,125,134,135], **Robert G Bristow#**[125,443,444,445,446], Colin C Collins[142], Shadrielle MG Espiritu[9], Neil E Fleshner[447], Natalie S Fox[9], Michael Fraser[9], Syed Haider[9], Lawrence E Heisler[44], Vincent Huang[9], Emilie Lalonde[9], Julie Livingstone[9], John D McPherson[76,429,436], Alice Meng[448], Veronica Y Sabelnykova[9], Adriana Salcedo[9], Yu-Jia Shiah[9], Theodorus Van der Kwast[441] and Takafumi N Yamaguchi[9]

PCAWG Tumour-specific providers (gastric cancer) in China

Shuai Ding[449], Daiming Fan[450], Yong Hou[38,235], Yi Huang[144,145], Lin Li[38], Siliang Li[38,235], Dongbing Liu[38,235], Xingmin Liu[38,235], **Youyong Lu#**[28,29,30], Yongzhan Nie[450,451], Hong Su[38,235], Jian Wang[38], Kui Wu[38,235], Xiao Xiao[145], Rui Xing[30], **Huanming Yang#**[38], Shanlin Yang[449], Yingyan Yu[452], Xiuqing Zhang[38], Yong Zhou[38] and Shida Zhu[38,235]

PCAWG Tumour-specific providers (renal cancer) in the EU & France

Rosamonde E Banks[453], Guillaume Bourque[454,455], Alvis Brazma[7], Paul Brennan[456], **Mark Lathrop#**[455], Louis Letourneau[457], Yasser Riazalhosseini[455], Ghislaine Scelo[456], **Jörg Tost#**[458], Naveen Vasudev[459] and Juris Viksna[460]

PCAWG Tumour-specific providers (breast cancer) in the EU & United Kingdom

Sung-Min Ahn[461], Ludmil B Alexandrov[1,70,302], Samuel Aparicio[462], Laurent Arnould[463], MR Aure[464], Shriram G Bhosle[1], Ewan Birney[7], Ake Borg[465], Sandrine Boyault[466], Arie B Brinkman[467], Jane E Brock[468], Annegien Broeks[469], Adam P Butler[1], Anne-Lise Børresen-Dale[464,470], Carlos Caldas[278,471], Peter J Campbell[1,2], Suet-Feung Chin[278,471], Helen Davies[1,309,310], Christine Desmedt[472,473], Luc Dirix[474], Serge Serge[1], Anna Ehinger[475], Jorunn E Eyfjord[476], Aquila Fatima[204], John A Foekens[477], P Andrew Futreal[478], Øystein Garred[479,480], Moritz Gerstung[7,8], Dilip D Giri[481], Dominik Glodzik[1], Dorthe Grabau[482], Holmfridur Hilmarsdottir[476], Gerrit K Hooijer[483], Jocelyne Jacquemier[484], Se Jin Jang[485], Jon G Jonasson[476], Jos Jonkers[486], Hyung-Yong Kim[484], Tari A King[487,488,489], Stian Knappskog[1,490], Gu Kong[484], Savitri Krishnamurthy[352,491], Sunil R Lakhani[492], Anita Langerød[464], Denis Larsimont[493], Hee Jin Lee[485], Jeong-Yeon Lee[494], Ming Ta Michael Lee[478], Yilong Li[1], Ole Christian Lingjærde[495], Gaetan MacGrogan[496], John WM Martens[477], Sancha Martin[1,353], Iñigo Martincorena[1], Andrew Menzies[1], Sandro Morganella[1], Ville Mustonen[305,306,307], Serena Nik-Zainal[1,309,310,311], Sarah O'Meara[1], Iris Pauporté[20], Sarah Pinder[497], Xavier Pivot[498], Elena Provenzano[499], Colin A Purdie[500], Keiran M Raine[1], Manasa Ramakrishna[1], Kamna Ramakrishnan[1], Jorge Reis-Filho[481], Andrea L Richardson[204], Markus Ringnér[501], Javier Bartolomé

Rodriguez[45], F Germán Rodríguez-González[502], Gilles Romieu[503], Roberto Salgado[394], Torill Sauer[495], Rebecca Shepherd[1], Anieta M Sieuwerts[477], Peter T Simpson[492], Marcel Smid[477], Christos Sotiriou[53], Paul N Span[504], Lucy Stebbings[1], Ólafur Andri Stefánsson[505], Alasdair Stenhouse[506], **Michael R Stratton#**[1], Henk G Stunnenberg[235,507], Fred Sweep[508], Benita Kiat Tee Tan[509], Jon W Teague[1], Gilles Thomas[510], Alastair M Thompson[506], Stefania Tommasi[511], Isabelle Treilleux[512,513], Andrew Tutt[204], Naoto T Ueno[514], Steven Van Laere[474], Peter Van Loo[60,61], Gert G Van den Eynden[474], Peter Vermeulen[474], Alain Viari[400], Anne Vincent-Salomon[507], David C Wedge[1,340,341], Bernice H Wong[515], Lucy Yates[1], Xueqing Zou[1], Carolien HM van Deurzen[516], Marc J van de Vijver[394] and L van't Veer[517]

PCAWG Tumour-specific providers (malignant lymphoma) in Germany

Ole Ammerpohl[518,519], Sietse Aukema[520,521], Anke K Bergmann[522], Stephan H Bernhart[260,261,264], Hans Binder[260,261], Arndt Borkhardt[523], Christoph Borst[524], Benedikt Brors[79,112,262], Birgit Burkhardt[525], Alexander Claviez[526], Roland Eils[50,52,63,64], Maria Elisabeth Goebler[527], Andrea Haake[518], Siegfried Haas[524], Martin Hansmann[528], Jessica I Hoell[523], Steve Hoffmann[260,261,263,264], Michael Hummel[529], Daniel Hübschmann[52,63,80,81,82], Dennis Karsch[530], Wolfram Klapper[520], Kortine Kleinheinz[50,52], Michael Kneba[530], Jan O Korbel[7,8], Helene Kretzmer[261,264], Markus Kreuz[531], Dieter Kube[532], Ralf Küppers[533], Chris Lawerenz[64], Dido Lenze[529], Peter Lichter[77,386], Markus Loeffler[531], Cristina López[246,518], Luisa Mantovani-Löffler[534], Peter Möller[535], German Ott[536], Bernhard Radlwimmer[386], Julia Richter[518,520], Marius Rohde[537], Philip C Rosenstiel[538], Andreas Rosenwald[539], Markus B Schilhabel[538], Matthias Schlesner[50,102], Stefan Schreiber[540], **Reiner Siebert#**[245,246], Peter F Stadler[260,261,264], Peter Staib[541], Stephan Stilgenbauer[542], Stephanie Sungalee[8], Monika Szczepanowski[520], Umut H Toprak[52,543], Lorenz HP Trümper[532], Rabea Wagener[246,518] and Thorsten Zenz[79]

PCAWG Tumour-specific providers (paediatric brain cancer) in Germany

Ivo Buchhalter[50,51,52], Juergen Eils[63,64], Roland Eils[50,52,63,64], Volker Hovestadt[386], Barbara Hutter[77,78,79], David TW Jones[285,286], Natalie Jäger[50], Christof von Kalle[81], Marcel Kool[93,285], Jan O Korbel[7,8], Andrey Korshunov[93], Pablo Landgraf[544,545], Chris Lawerenz[64], Hans Lehrach[546], **Peter Lichter#**[77,386], Paul A Northcott[547], Stefan M Pfister[93,285,548], Bernhard Radlwimmer[386], Guido Reifenberger[545], Matthias Schlesner[50,102], Hans-Jörg Warnatz[546], Joachim Weischenfeldt[8,107,108], Stephan Wolf[549], Marie-Laure Yaspo[546] and Marc Zapatka[386]

PCAWG Tumour-specific providers (prostate cancer) in Germany

Yassen Assenov[550], Benedikt Brors[79,112,262], Juergen Eils[63,64], Roland Eils[50,52,63,64], Lars Feuerbach[112], Clarissa Gerhauser[268], Jan O Korbel[7,8], Chris Lawerenz[64], Hans Lehrach[546], Sarah Minner[551],

Christoph Plass[268], **Guido Sauter#**[552], Thorsten Schlomm[107,553], Nikos Sidiropoulos[108], Ronald Simon[552], **Holger Sültmann#**[79,554], Hans-Jörg Warnatz[546], Dieter Weichenhan[268], Joachim Weischenfeldt[8,107,108] and Marie-Laure Yaspo[546]

PCAWG Tumour-specific providers (oral cancer) in India

Nidhan K Biswas[555], Luca Landoni[396], Arindam Maitra[555], **Partha P Majumder#**[555] and **Rajiv Sarin#**[556]

PCAWG Tumour-specific providers (pancreatic cancer) in Italy

Davide Antonello[396], Stefano Barbi[403], Claudio Bassi[396], Samantha Bersani[399], Giada Bonizzato[400], Cinzia Cantù[400], Ivana Cataldo[399,400], Sara Cingarlini[36], Vincenzo Corbo[400,403], Maria Vittoria Davi[404], Angelo P Dei Tos[557], Matteo Fassan[558], Sonia Grimaldi[400], Luca Landoni[396], Rita T Lawlor[400], Claudio Luchini[399], Andrea Mafficini[400], Giuseppe Malleo[396], Giovanni Marchegiani[396], Michele Milella[36], Marco Miotto[396], Salvatore Paiella[396], Antonio Pea[396], Paolo Pederzoli[396], Borislav C Rusev[400], Andrea Ruzzenente[396], Roberto Salvia[396], Maria Scardoni[399], **Aldo Scarpa#**[400], Elisabetta Sereni[396], Michele Simbolo[403], Nicola Sperandio[400], Giampaolo Tortora[36,37] and Caterina Vicentini[400]

PCAWG Tumour-specific providers (biliary tract cancer) in Japan

Yasuhito Arai[33], Natsuko Hama[33], Nobuyoshi Hiraoka[559], Fumie Hosoda[33], Mamoru Kato[351], Hiromi Nakamura[33], Hidenori Ojima[560], Takuji Okusaka[561], **Tatsuhiro Shibata#**[33,34], Yasushi Totoki[33] and Tomoko Urushidate[34]

PCAWG Tumour-specific providers (gastric cancer) in Japan

Hiroyuki Aburatani#[256], Yasuhito Arai[33], Masashi Fukayama[562], Natsuko Hama[33], Fumie Hosoda[33], Shumpei Ishikawa[563], Hitoshi Katai[564], Mamoru Kato[351], Hiroto Katoh[563], Daisuke Komura[563], Genta Nagae[256,267], Hiromi Nakamura[33], Hirofumi Rokutan[351], Mihoko Saito-Adachi[351], **Tatsuhiro Shibata#**[33,34], Akihiro Suzuki[256,565], Hirokazu Taniguchi[34], Kenji Tatsuno[256], Yasushi Totoki[33], Tetsuo Ushiku[562], Shinichi Yachida[33,566] and Shogo Yamamoto[256]

PCAWG Tumour-specific providers (liver cancer) in Japan

Hiroyuki Aburatani[256], Hiroshi Aikata[567], Koji Arihiro[567], Shun-ichi Ariizumi[568], Keith A Boroevich[46,47], Kazuaki Chayama[567], Akihiro Fujimoto[47], Masashi Fujita[47], Mayuko Furuta[47], Kunihito Gotoh[569], Natsuko Hama[33], Takanori Hasegawa[83], Shinya Hayami[570], Shuto Hayashi[83], Satoshi Hirano[571], Seiya Imoto[83], Mamoru Kato[351], Yoshiiku Kawakami[567], Kazuhiro Maejima[47],

Satoru Miyano[83], Genta Nagae[256,267], **Hidewaki Nakagawa#**[47], Hiromi Nakamura[33], Toru Nakamura[571], Kaoru Nakano[47], Hideki Ohdan[567], Yasushi Rino[572], Aya Sasaki-Oku[47], **Tatsuhiro Shibata#**[33,34], Yuichi Shiraishi[83], Hiroko Tanaka[83], Yasushi Totoki[33], Tatsuhiko Tsunoda[46,207,208,209], Masaki Ueno[570], Rui Yamaguchi[83], Masakazu Yamamoto[568] and Hiroki Yamaue[570]

PCAWG Tumour-specific providers (biliary tract cancer) in Singapore

Su Pin Choo[573], Ioana Cutcutache[251,304], Narong Khuntikeo[396,574], John R McPherson[251,304], Choon Kiat Ong[575], Chawalit Pairojkul[394], Irinel Popescu[576], **Steven G Rozen#**[251,252,304], **Patrick Tan#**[239,250,251,252] and **Bin Tean Teh#**[250,251,252,253,254]

PCAWG Tumour-specific providers (blood cancer) in South Korea

Keun Soo Ahn[577], Hyung-Lae Kim[27], Youngil Koh[291,292] and **Sung-Soo Yoon#**[292]

PCAWG Tumour-specific providers (chronic lymphocytic leukaemia) in Spain

Marta Aymerich[578], **Elias Campo#**[579,580], Josep L L Gelpi[45,67], Ivo G Gut[127,128], Marta Gut[127,128], Armando Lopez-Guillermo[581], Carlos López-Otín[582], Xose S Puente[582], Romina Royo[45] and David Torrents[45,105]

PCAWG Tumour-specific providers (bone cancer) in the United Kingdom

Fernanda Amary[583], Daniel Baumhoer[584], Sam Behjati[1], Bodil Bjerkehagen[584,585], **Peter J Campbell#**[1,2], **Adrienne M Flanagan#**[586], P Andrew Futreal[478], Ola Myklebost[490], Nischalan Pillay[587], Patrick Tarpey[588], Roberto Tirabosco[589] and Olga Zaikova[590]

PCAWG Tumour-specific providers (chronic myeloid disorders) in the United Kingdom

Jacqueline Boultwood[591], David T Bowen[1], Adam P Butler[1], **Peter J Campbell#**[1,2], Mario Cazzola[592], Carlo Gambacorti-Passerini[177], Anthony R Green[279], Eva Hellstrom-Lindberg[593], Luca Malcovati[592], Sancha Martin[1,353], Jyoti Nangalia[1], Elli Papaemmanuil[2] and Paresh Vyas[295,594]

PCAWG Tumour-specific providers (oesophageal cancer) in the

United Kingdom

Yeng Ang[595], Hugh Barr[596], Duncan Beardsmore[597], Matthew Eldridge[278], **Rebecca C Fitzgerald#**[310], James Gossage[598], Nicola Grehan[310], George B Hanna[599], Stephen J Hayes[600,601], Ted R Hupp[602], David Khoo[603], Jesper Lagergren[593,604], Laurence B Lovat[237], Shona MacRae[375], Maria O'Donovan[310], J Robert O'Neill[605], Simon L Parsons[606], Shaun R Preston[607], Sonia Puig[608], Tom Roques[609], Grant Sanders[234], Sharmila Sothi[610], Simon Tavaré[278], Olga Tucker[611], Richard Turkington[612], Timothy J Underwood[613] and Ian Welch[614]

PCAWG Tumour-specific providers (prostate cancer) in the United Kingdom

Daniel M Berney[615], Johann S De Bono[383], G Steven Bova[316], Daniel S Brewer[381,382], Adam P Butler[1], Declan Cahill[616], Niedzica Camacho[383], **Colin S Cooper#**[382,383,384], Nening M Dennis[616], Tim Dudderidge[616,617], Sandra E Edwards[383], **Rosalind A Eeles#**[383,616], Cyril Fisher[616], Christopher S Foster[618,619], Mohammed Ghori[1], Pelvender Gill[594], Vincent J Gnanapragasam[367,620], Gunes Gundem[187], Freddie C Hamdy[594], Steve Hawkins[278], Steven Hazell[616], William Howat[367], William B Isaacs[621], Katalin Karaszi[594], Jonathan D Kay[237], Vincent Khoo[616], Zsofia Kote-Jarai[383], Barbara Kremeyer[1], Pardeep Kumar[616], Adam Lambert[594], Daniel A Leongamornlert[1,383], Naomi Livni[616], Yong-Jie Lu[615,622], Hayley J Luxton[237], Andy G Lynch[278,279,294], Luke Marsden[594], Charlie E Massie[278], Lucy Matthews[383], Erik Mayer[616,623], Ultan McDermott[1], Sue Merson[383], Thomas J Mitchell[1,279,367], David E Neal[278,367], Anthony Ng[624], David Nicol[616], Christopher Ogden[616], Edward W Rowe[616], Nimish C Shah[367], Jon W Teague[1], Sarah Thomas[616], Alan Thompson[616], Peter Van Loo[60,61], Clare Verrill[594,625], Tapio Visakorpi[316], Anne Y Warren[367,626], David C Wedge[1,340,341], Hayley C Whitaker[237], Jorge Zamora[1,270,271,272], Hongwei Zhang[622] and Nicholas van As[616]

PCAWG Tumour-specific providers (TCGA) in the United States

Adam Abeshouse[187], Nishant Agrawal[627], Rehan Akbani[310,376], Hikmat Al-Ahmadie[187], Monique Albert[428], Kenneth Aldape[352,628], Adrian Ally[629], Yeng Ang[595], Elizabeth L Appelbaum[131,237], Joshua Armenia[630], Sylvia Asa[441,606], J Todd Auman[631], Matthew H Bailey[130,131], Miruna Balasundaram[629], Saianand Balu[234], Jill Barnholtz-Sloan[632,633], Hugh Barr[596], John Bartlett[427,428], Oliver F Bathe[634,635], Stephen B Baylin[320,617], Duncan Beardsmore[597], Christopher Benz[636], Andrew Berchuck[637], Benjamin P Berman[257,258,259], Rameen Beroukhim[3,6,163], Mario Berrios[638], Darell Bigner[639], Michael Birrer[100], Tom Bodenheimer[234], Lori Boice[608], Moiz S Bootwalla[638], Marcus Bosenberg[640], Reanne Bowlby[629], Jeffrey Boyd[641], Russell R Broaddus[352], Malcolm Brock[642], Denise Brooks[629], Susan Bullman[3,163], Samantha J Caesar-Johnson[39], Thomas E Carey[643], Rebecca Carlsen[629], Robert Cerfolio[644], Vishal S Chandan[645], Hsiao-Wei Chen[595,630], Andrew D Cherniack[3,48,163], Jeremy Chien[646], Juok Cho[3], Eric Chuah[629], Carrie Cibulskis[3], Kristian Cibulskis[3], Leslie Cope[320], Matthew G Cordes[131,609], Kyle Covington[155], Erin Curley[647], Bogdan Czerniak[352,603], Ludmila Danilova[320], Ian J Davis[648], Timothy Defreitas[3], John A Demchok[39], Noreen Dhalla[629], Rajiv Dhir[649], Li Ding[130,131,138], HarshaVardhan Doddapaneni[155], Adel El-Naggar[352,603], Ina Felau[39], Martin L Ferguson[650], Gaetano

Finocchiaro[651], Kwun M Fong[652], Scott Frazer[3], William Friedman[653], Catrina C Fronick[131,609], Lucinda A Fulton[131], Robert S Fulton[130,131,138], Stacey B Gabriel[3], Jianjiong Gao[630], Nils Gehlenborg[3,654], Jeffrey E Gershenwald[655,656], Gad Getz[3,4,5,6], Ronald Ghossein[481], Nasra H Giama[657], Richard A Gibbs[155], Carmen Gomez[658], James Gossage[598], Ramaswamy Govindan[130], Nicola Grehan[310], George B Hanna[599], D Neil Hayes[234,659,660], Stephen J Hayes[600,601], Apurva M Hegde[375,376], David I Heiman[3], Zachary Heins[187], Austin J Hepperla[234], Katherine A Hoadley[233,234], Andrea Holbrook[638], Robert A Holt[629], Alan P Hoyle[234], Ralph H Hruban[320], Jianhong Hu[155], Mei Huang[608], David Huntsman[661], Ted R Hupp[602], Jason Huse[187], **Carolyn M Hutter#**[23], Christine A Iacobuzio-Donahue[481], Michael Ittmann[662,663,664], Joy C Jayaseelan[155], Stuart R Jefferys[234], Corbin D Jones[665], Steven JM Jones[629], Hartmut Juhl[666], Koo Jeong Kang[667], Beth Karlan[668], Katayoon Kasaian[629], Electron Kebebew[669,670], David Khoo[603], Hark Kyun Kim[671], Jaegil Kim[3], Tari A King[487,488,489], Viktoriya Korchina[155], Ritika Kundra[595,630], Jesper Lagergren[593,604], Phillip H Lai[638], Peter W Laird[265], Eric Lander[3], Michael S Lawrence[3,46,100], Alexander J Lazar[158,352], Xuan Le[672], Darlene Lee[673], Douglas A Levine[187,674], Lora Lewis[155], Tim Ley[675], Haiyan Irene Li[673], Pei Lin[3], W M Linehan[676], Eric Minwei Liu[110,111,187], Fei Fei Liu[368], Laurence B Lovat[237], Yiling Lu[376], Lisa Lype[677], Yussanne Ma[673], Shona MacRae[375], Dennis T Maglinte[638,678], Elaine R Mardis[131,641,679], Jeffrey Marks[396,680], Marco A Marra[673], Thomas J Matthew[11], Michael Mayo[673], Karen McCune[681], Michael D McLellan[130,131,138], Samuel R Meier[3], Shaowu Meng[234], Matthew Meyerson[3,6,48], Piotr A Mieczkowski[233], Tom Mikkelsen[682], Christopher A Miller[131], Gordon B Mills[683], Richard A Moore[673], Carl Morrison[394,684], Lisle E Mose[234], Catherine D Moser[657], Andrew J Mungall[673], Karen Mungall[673], David Mutch[685], Donna M Muzny[155], Jerome Myers[686], Yulia Newton[11], Michael S Noble[3], Peter O'Donnell[687], Brian Patrick O'Neill[688], Angelica Ochoa[187], Akinyemi I Ojesina[240,241,242], Joong-Won Park[689], Joel S Parker[690], Simon L Parsons[606], Harvey Pass[691], Alessandro Pastore[85], Chandra Sekhar Pedamallu[3,6,163], Nathan A Pennell[692], Charles M Perou[234,690,693], Gloria M Petersen[438], Nicholas Petrelli[694], Olga Potapova[695], Shaun R Preston[607], Sonia Puig[608], Janet S Rader[696], Suresh Ramalingam[697], W Kimryn Rathmell[698], Victor Reuter[481], Sheila M Reynolds[677], Matthew Ringel[699], Jeffrey Roach[700], Lewis R Roberts[657], A Gordon Robertson[673], Tom Roques[609], Mark A Rubin[183,197,198,199,200], Sara Sadeghi[673], Gordon Saksena[3], Charles Saller[701], Francisco Sanchez-Vega[595,630], Chris Sander[85,202,203], Grant Sanders[234], Dirk Schadendorf[77,702], Jacqueline E Schein[673], Heather K Schmidt[131], Nikolaus Schultz[630], Steven E Schumacher[3,204], Richard A Scolyer[409,413,419], Raja Seethala[703], Yasin Senbabaoglu[85], Troy Shelton[647], Yan Shi[234], Juliann Shih[3,163], Ilya Shmulevich[677], Craig Shriver[704], Sabina Signoretti[163,168,705], Janae V Simons[234], Samuel Singer[396,706], Payal Sipahimalani[673], Tara J Skelly[233], Karen Smith-McCune[681], Nicholas D Socci[85], Heidi J Sofia[23], Matthew G Soloway[690], Anil K Sood[707,708,709], Sharmila Sothi[610], Angela Tam[673], Donghui Tan[233], Roy Tarnuzzer[39], Nina Thiessen[673], R Houston Thompson[710], Leigh B Thorne[608], Ming Tsao[437,606], Olga Tucker[611], Richard Turkington[612], Christopher Umbricht[274,597,711], Timothy J Underwood[613], David J Van Den Berg[638], Erwin G Van Meir[712], Umadevi Veluvolu[233], Douglas Voet[3], Jiayin Wang[131,145,152], Linghua Wang[155], Zhining Wang[39], Paul Weinberger[713], John N Weinstein[375,376], Daniel J Weisenberger[638], Ian Welch[614], David A Wheeler[154,155], Dennis Wigle[714], Matthew D Wilkerson[233], Richard K Wilson[131,715], Boris Winterhoff[716], Maciej Wiznerowicz[717,718], Tina Wong[131,673], Winghing Wong[719], Liu Xi[155], Liming Yang[39], Christina Yau[636], Venkata D Yellapantula[138,157], **Jean C Zenklusen#**[39], Hailei Zhang[3], Hongxin Zhang[630] and Jiashan Zhang[39]

# Denotes **working group or project co-leader**

Author Affiliations

1. Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK.
2. Department of Haematology, University of Cambridge, Cambridge CB2 2XY, UK.
3. Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.
4. Center for Cancer Research, Massachusetts General Hospital, Boston, MA 02129, USA.
5. Department of Pathology, Massachusetts General Hospital, Boston, MA 02115, USA.
6. Harvard Medical School, Boston, MA 02115, USA.
7. European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge CB10 1SD, UK.
8. Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg 69117, Germany.
9. Computational Biology Program, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
10. Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada.
11. Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA.
12. UC Santa Cruz Genomics Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA.
13. International Cancer Genome Consortium (ICGC)/ICGC Accelerating Research in Genomic Oncology (ARGO) Secretariat, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
14. King Faisal Specialist Hospital and Research Centre, Al Maather, Riyadh 12713, Saudi Arabia.
15. DLR Project Management Agency, Bonn 53227, Germany.
16. Genome Canada, Ottawa, ON K2P 1P1, Canada.
17. Instituto Carlos Slim de la Salud, Mexico City, Mexico.
18. Federal Ministry of Education and Research, Berlin 10117, Germany.
19. Institut Gustave Roussy, Villejuif 94805, France.
20. Institut National du Cancer (INCA), Boulogne-Billancourt 92100, France.
21. The Wellcome Trust, London NW1 2BE, UK.
22. Prostate Cancer Canada, Toronto, ON M5C 1M1, Canada.
23. National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA.
24. Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi, Delhi 110003, India.
25. Science Writer, Garrett Park, MD 20896, USA.
26. Cancer Research UK, London EC1V 4AD, UK.
27. Department of Biochemistry, College of Medicine, Ewha Womans University, Seoul 07895, South Korea.
28. Chinese Cancer Genome Consortium, Shenzhen 518083, China.
29. Department of Medical Oncology, Beijing Hospital, Beijing 100730, China.
30. Laboratory of Molecular Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital and Institute, Beijing 100142, China.
31. National Cancer Center, Tokyo 104-0045, Japan.

32. German Cancer Aid, Bonn 53113, Germany.
33. Division of Cancer Genomics, National Cancer Center Research Institute, National Cancer Center, Tokyo 104-0045, Japan.
34. Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan.
35. Japan Agency for Medical Research and Development, Tokyo 100-0004, Japan.
36. Medical Oncology, University and Hospital Trust of Verona, Verona 37134, Italy.
37. University of Verona, Verona 37129, Italy.
38. BGI-Shenzhen, Shenzhen 518083, China.
39. National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
40. Centre for Law and Genetics, University of Tasmania, Sandy Bay Campus, Hobart, TAS 7001, Australia.
41. Centre of Genomics and Policy, McGill University and Génome Québec Innovation Centre, Montreal, QC H3A 1A4, Canada.
42. Heidelberg Academy of Sciences and Humanities, Heidelberg 69120, Germany.
43. CAPHRI Research School, Maastricht University, Maastricht, ER 6200MD, The Netherlands.
44. Genome Informatics Program, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
45. Barcelona Supercomputing Center (BSC), Barcelona 08034, Spain.
46. Laboratory for Medical Science Mathematics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
47. RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
48. Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA.
49. Biomedical Engineering, Oregon Health and Science University, Portland, OR 97239, USA.
50. Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
51. Heidelberg Center for Personalized Oncology (DKFZ-HIPO), German Cancer Research Center, Heidelberg 69120, Germany.
52. Institute of Pharmacy and Molecular Biotechnology and BioQuant, Heidelberg University, Heidelberg 69120, Germany.
53. University of California San Diego, San Diego, CA 92093, USA.
54. PDXen Biosystems Inc, Seoul 4900, South Korea.
55. Electronics and Telecommunications Research Institute, Daejeon 34129, South Korea.
56. Seven Bridges Genomics, Charlestown, MA 02129, USA.
57. Annai Systems, Inc, Carlsbad, CA 92013, USA.
58. Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, CA 94305, USA.
59. Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA.
60. The Francis Crick Institute, London NW1 1AT, UK.
61. University of Leuven, Leuven B-3000, Belgium.
62. The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada.
63. Heidelberg University, Heidelberg 69120, Germany.
64. New BIH Digital Health Center, Berlin Institute of Health (BIH) and Charité - Universitätsmedizin Berlin, Berlin 10117, Germany.
65. Department of Biochemistry and Molecular Medicine, University of Montreal, Montreal, QC

H3C 3J7, Canada.
66. CIBIO/InBIO - Research Center in Biodiversity and Genetic Resources, Universidade do Porto, Vairão 4485-601, Portugal.
67. Department Biochemistry and Molecular Biomedicine, University of Barcelona, Barcelona 08028, Spain.
68. Department of Medicine, Section of Hematology/Oncology, University of Chicago, Chicago, IL 60637, USA.
69. Department of Medicine, Division of Biomedical Informatics, UC San Diego School of Medicine, San Diego, CA 92093, USA.
70. UC San Diego Moores Cancer Center, San Diego, CA 92093, USA.
71. Children's Hospital of Philadelphia, Philadelphia, PA 19146, USA.
72. Massachusetts General Hospital Center for Cancer Research, Charlestown, MA 02129, USA.
73. University of Melbourne Centre for Cancer Research, Melbourne, VIC 3010, Australia.
74. Syntekabio Inc, Daejon 34025, South Korea.
75. AbbVie, North Chicago, IL 60064, USA.
76. Genomics Research Program, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
77. German Cancer Consortium (DKTK), Heidelberg 69120, Germany.
78. Heidelberg Center for Personalized Oncology (DKFZ-HIPO), German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
79. National Center for Tumor Diseases (NCT) Heidelberg, Heidelberg 69120, Germany.
80. Department of Pediatric Immunology, Hematology and Oncology, University Hospital, Heidelberg 69120, Germany.
81. German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
82. Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM), Heidelberg 69120, Germany.
83. Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan.
84. Genome Integration Data Center, Syntekabio, Inc, Daejon, 34025, South Korea.
85. Computational Biology Center, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
86. Department of Biology, ETH Zurich, Zürich 8093, Switzerland.
87. Department of Computer Science, ETH Zurich, Zurich 8092, Switzerland.
88. SIB Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland.
89. University Hospital Zurich, Zurich 8091, Switzerland.
90. Health Sciences Department of Biomedical Informatics, University of California San Diego, La Jolla, CA 92093, USA.
91. Department of Health Sciences and Technology, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea.
92. Samsung Genome Institute, Seoul 06351, South Korea.
93. Functional and Structural Genomics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
94. Leidos Biomedical Research, Inc, McLean, VA 22102, USA.
95. Center for Biomolecular Science and Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA.
96. Sage Bionetworks, Seattle, WA 98109, USA.

97. Department of Cell and Systems Biology, University of Toronto, Toronto, ON M5S 3G5, Canada.
98. Department of Radiation Oncology, University of California San Francisco, San Francisco, CA 94518, USA.
99. CSRA Incorporated, Fairfax, VA 22042, USA.
100. Massachusetts General Hospital, Boston, MA 02114, USA.
101. Weill Cornell Medical College, New York, NY 10065, USA.
102. Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
103. Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL 60637, USA.
104. Computational Biology Program, School of Medicine, Oregon Health and Science University, Portland, OR 97239, USA.
105. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona 08010, Spain.
106. Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim 7030, Norway.
107. Department of Urology, Charité Universitätsmedizin Berlin, Berlin 10117, Germany.
108. Finsen Laboratory and Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen 2200, Denmark.
109. Department of Biological Oceanography, Leibniz Institute of Baltic Sea Research, Rostock 18119, Germany.
110. Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA.
111. Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY 10021, USA.
112. Division of Applied Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
113. Department of Computer Science, Yale University, New Haven, CT 06520, USA.
114. Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA.
115. Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06520, USA.
116. Department of Internal Medicine, Stanford University, Stanford, CA 94305, USA.
117. Department of Molecular Medicine (MOMA), Aarhus University Hospital, Aarhus N 8200, Denmark.
118. Clinical Bioinformatics, Swiss Institute of Bioinformatics, Geneva 1202, Switzerland.
119. Institute for Pathology and Molecular Pathology, University Hospital Zurich, Zurich 8091, Switzerland.
120. Institute of Molecular Life Sciences, University of Zurich, Zurich 8057, Switzerland.
121. MIT Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
122. Controlled Department and Institution, New York, NY 10065, USA.
123. Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY 10065, USA.
124. Bioinformatics Research Centre (BiRC), Aarhus University, Aarhus 8000, Denmark.
125. Department of Medical Biophysics, University of Toronto, Toronto, ON M5S 1A8, Canada.

126. Institute of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, Zurich 8057, Switzerland.
127. CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona 08028, Spain.
128. Universitat Pompeu Fabra (UPF), Barcelona 08003, Spain.
129. Office of Cancer Genomics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
130. Alvin J. Siteman Cancer Center, Washington University School of Medicine, St Louis, MO 63110, USA.
131. The McDonnell Genome Institute at Washington University, St Louis, MO 63108, USA.
132. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China.
133. Center for Digital Health, Berlin Institute of Health and Charitè - Universitätsmedizin Berlin, Berlin 10117, Germany.
134. Department of Human Genetics, University of California Los Angeles, Los Angeles, CA 90095, USA.
135. Department of Pharmacology, University of Toronto, Toronto, ON M5S 1A8, Canada.
136. University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
137. Department of Genetics, Informatics Institute, University of Alabama at Birmingham, Birmingham, AL 35294, USA.
138. Department of Medicine and Department of Genetics, Washington University School of Medicine, St. Louis, St Louis, MO 63110, USA.
139. Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona 08003, Spain.
140. Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
141. Department of Urologic Sciences, University of British Columbia, Vancouver, BC V5Z 1M9, Canada.
142. Vancouver Prostate Centre, Vancouver, BC V6H 3Z6, Canada.
143. Division of Life Science and Applied Genomics Center, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China.
144. Geneplus-Shenzhen, Shenzhen 518122, China.
145. School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710048, China.
146. Biobyte solutions GmbH, Heidelberg 69126, Germany.
147. Division of Oncology, Washington University School of Medicine, St Louis, MO 63110, USA.
148. Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen 72076, Germany.
149. Indiana University, Bloomington, IN 47405, USA.
150. Simon Fraser University, Burnaby, BC V5A 1S6, Canada.
151. Department of Computer Science, University of Toronto, Toronto, ON M5S 1A8, Canada.
152. School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710048, China.
153. Department of Genetics, Washington University School of Medicine, St Louis, MO 63110, USA.

154. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA.
155. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA.
156. The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an 710048, China.
157. Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
158. Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
159. Quantitative and Computational Biosciences Graduate Program, Baylor College of Medicine, Houston, TX 77030, USA.
160. The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA.
161. Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Boston, MA 02215, USA.
162. Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA.
163. Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA.
164. Department of Mathematics, Aarhus University, Aarhus 8000, Denmark.
165. Center for Biological Sequence Analysis, Department of Bio and Health Informatics, Technical University of Denmark, Lyngby 2800, Denmark.
166. Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen 2200, Denmark.
167. Department for BioMedical Research, University of Bern, Bern 3008, Switzerland.
168. Department of Medical Oncology, Inselspital, University Hospital and University of Bern, Bern 3010, Switzerland.
169. Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern 3012, Switzerland.
170. Department of Urology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
171. Faculty of Biosciences, Heidelberg University, Heidelberg 69120, Germany.
172. Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea.
173. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona 8003, Spain.
174. Research Program on Biomedical Informatics, Universitat Pompeu Fabra, Barcelona 08002, Spain.
175. Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala SE-75124, Sweden.
176. Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, University of Queensland, St Lucia, Brisbane, QLD 4072, Australia.
177. University of Milano Bicocca, Monza 20052, Italy.
178. Sir Peter MacCallum Department of Oncology, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, VIC 3000, Australia.
179. Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center, Houston, TX 77030, USA.
180. The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada.
181. Health Data Science Unit, University Clinics, Heidelberg 69120, Germany.
182. Technical University of Denmark, Lyngby 2800, Denmark.

183. Department for Biomedical Research, University of Bern, Bern 3008, Switzerland.
184. Research Core Center, National Cancer Centre Korea, Goyang-si 410-769, South Korea.
185. Institute of Computer Science, Polish Academy of Sciences, Warsawa 01-248, Poland.
186. Harvard University, Cambridge, MA 02138, USA.
187. Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
188. Department of Information Technology, Ghent University, Ghent B-9000, Belgium.
189. Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent B-9000, Belgium.
190. Yale School of Medicine, Yale University, New Haven, CT 06520, USA.
191. Institute for Research in Biomedicine (IRB Barcelona), Barcelona 08028, Spain.
192. Division of Hematology-Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea.
193. Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea.
194. Cheonan Industry-Academic Collaboration Foundation, Sangmyung University, Cheonan 31066, South Korea.
195. Spanish National Cancer Research Centre, Madrid 28029, Spain.
196. Department of Computer Science, Princeton University, Princeton, NJ 08540, USA.
197. Bern Center for Precision Medicine, University Hospital of Bern, University of Bern, Bern 3008, Switzerland.
198. Englander Institute for Precision Medicine, Weill Cornell Medicine and New York Presbyterian Hospital, New York, NY 10021, USA.
199. Meyer Cancer Center, Weill Cornell Medicine, New York, NY 10065, USA.
200. Pathology and Laboratory, Weill Cornell Medical College, New York, NY 10021, USA.
201. Vall d'Hebron Institute of Oncology: VHIO, Barcelona 08035, Spain.
202. cBio Center, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA.
203. Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA.
204. Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA.
205. Dana-Farber Cancer Institute, Boston, MA 02215, USA.
206. cBio Center, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA.
207. Core Research for Evolutional Science and Technology (CREST), JST, Tokyo 102-8666, Japan.
208. Department of Biological Sciences, Laboratory for Medical Science Mathematics, Graduate School of Science, University of Tokyo, Yokohama 230-0045, Japan.
209. Department of Medical Science Mathematics, Medical Research Institute, Tokyo Medical and Dental University (TMDU), Tokyo 113-8510, Japan.
210. Department of Oncology-Pathology, Science for Life Laboratory, Karolinska Institutet, Stockholm 17121, Sweden.
211. Department of Gene Technology, Tallinn University of Technology, Tallinn 12616, Estonia.
212. Genetics and Genome Biology Program, SickKids Research Institute, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada.
213. Department of Information Technology, Ghent University, Interuniversitair Micro-Electronica Centrum (IMEC), Ghent B-9000, Belgium.

214. Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala SE-75108, Sweden.
215. Department of Medical Informatics and Clinical Epidemiology, Division of Bioinformatics and Computational Biology, OHSU Knight Cancer Institute, Oregon Health and Science University, Portland, OR 97239, USA.
216. Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China.
217. The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
218. Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH 43210, USA.
219. The Ohio State University Comprehensive Cancer Center (OSUCCC – James), Columbus, OH 43210, USA.
220. The University of Texas School of Biomedical Informatics (SBMI) at Houston, Houston, TX 77030, USA.
221. Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL 60637, USA.
222. Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Glasgow G61 1BD, UK.
223. Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA.
224. Department of Chemistry, Centre for Molecular Science Informatics, University of Cambridge, Cambridge CB2 1EW, UK.
225. Ludwig Center at Harvard Medical School, Boston, MA 02115, USA.
226. Physics Division, Optimization and Systems Biology Lab, Massachusetts General Hospital, Boston, MA 02114, USA.
227. Department of Medicine, Baylor College of Medicine, Houston, TX 77030, USA.
228. Computational and Systems Biology, Genome Institute of Singapore, Singapore 138672, Singapore.
229. School of Computing, National University of Singapore, Singapore 117417, Singapore.
230. The Azrieli Faculty of Medicine, Bar-Ilan University, Safed 13195, Israel.
231. National Cancer Centre Singapore, Singapore 169610, Singapore.
232. Peking University, Beijing 100871, China.
233. Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
234. Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
235. China National GeneBank-Shenzhen, Shenzhen 518083, China.
236. Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin 13125, Germany.
237. University College London, London WC1E 6BT, UK.
238. School of Life Sciences, Peking University, Beijing 100180, China.
239. Genome Institute of Singapore, Singapore 138672, Singapore.
240. Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA.
241. HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA.
242. O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham,

AL 35294, USA.
243. Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm 14183, Sweden.
244. German Cancer Consortium (DKTK), Partner site Berlin.
245. Human Genetics, University of Kiel, Kiel 24118, Germany.
246. Institute of Human Genetics, Ulm University and Ulm University Medical Center, Ulm 89081, Germany.
247. Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
248. Korea University, Seoul 02481, South Korea.
249. Division of Computational Genomics and Systems Genetics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
250. Cancer Science Institute of Singapore, National University of Singapore, Singapore 169609, Singapore.
251. Programme in Cancer and Stem Cell Biology, Duke-NUS Medical School, Singapore 169857, Singapore.
252. SingHealth, Duke-NUS Institute of Precision Medicine, National Heart Centre Singapore, Singapore 169609, Singapore.
253. Institute of Molecular and Cell Biology, Singapore 169609, Singapore.
254. Laboratory of Cancer Epigenome, Division of Medical Science, National Cancer Centre Singapore, Singapore 169610, Singapore.
255. BIOPIC, ICG and College of Life Sciences, Peking University, Beijing 100871, China.
256. Genome Science Division, Research Center for Advanced Science and Technology, University of Tokyo, Tokyo 153-8904, Japan.
257. Center for Bioinformatics and Functional Genomics, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA.
258. Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA.
259. The Hebrew University Faculty of Medicine, Jerusalem 91120, Israel.
260. Department of Computer Science, Bioinformatics Group, University of Leipzig, Leipzig 04109, Germany.
261. Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig 04109, Germany.
262. German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
263. Computational Biology, Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Jena 07745, Germany.
264. Transcriptome Bioinformatics, LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig 04109, Germany.
265. Center for Epigenetics, Van Andel Research Institute, Grand Rapids, MI 49503, USA.
266. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona 08036, Spain.
267. Research Center for Advanced Science and Technology, University of Tokyo, Tokyo 108-8639, Japan.
268. Cancer Epigenomics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
269. Van Andel Research Institute, Grand Rapids, MI 49503, USA.
270. Centre for Research in Molecular Medicine and Chronic Diseases (CiMUS), Universidade de Santiago de Compostela, Santiago de Compostela 15706, Spain.

271. Department of Zoology, Genetics and Physical Anthropology, (CiMUS), Universidade de Santiago de Compostela, Santiago de Compostela 15706, Spain.
272. The Biomedical Research Centre (CINBIO), Universidade de Vigo, Vigo 36310, Spain.
273. Department of Veterinary Medicine, Transmissible Cancer Group, University of Cambridge, Cambridge CB3 0ES, UK.
274. Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
275. McKusick-Nathans Institute of Genetic Medicine, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.
276. Foundation Medicine, Inc, Cambridge, MA 02141, USA.
277. Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada.
278. Li Ka Shing Centre, Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, UK.
279. University of Cambridge, Cambridge CB2 1TN, UK.
280. Quantitative Genomics Laboratories (qGenomics), Barcelona 08950, Spain.
281. Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences (NIEHS), Durham, NC 27709, USA.
282. Brandeis University, Waltham, MA 02254, USA.
283. Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY 10065, USA.
284. New York Genome Center, New York, NY 10013, USA.
285. Hopp Children's Cancer Center (KiTZ), Heidelberg 69120, Germany.
286. Pediatric Glioma Research Group, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
287. A.A. Kharkevich Institute of Information Transmission Problems, Moscow 127051, Russia.
288. Oncology and Immunology, Dmitry Rogachev National Research Center of Pediatric Hematology, Moscow 117997, Russia.
289. Skolkovo Institute of Science and Technology, Moscow 121205, Russia.
290. Integrative Bioinformatics Support Group, National Institute of Environmental Health Sciences (NIEHS), Durham, NC 27709, USA.
291. Center For Medical Innovation, Seoul National University Hospital, Seoul 03080, South Korea.
292. Department of Internal Medicine, Seoul National University Hospital, Seoul 03080, South Korea.
293. Division of Genetics and Genomics, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA.
294. School of Medicine/School of Mathematics and Statistics, University of St Andrews, St Andrews, Fife KY16 9SS, UK.
295. Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia.
296. Institute for Molecular Bioscience, University of Queensland, St Lucia, Brisbane, QLD 4072, Australia.
297. School of Molecular Biosciences and Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA.

298. Cancer Research Institute, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA.
299. Ben May Department for Cancer Research and Department of Human Genetics, University of Chicago, Chicago, IL 60637, USA.
300. Tri-Institutional PhD Program in Computational Biology and Medicine, Weill Cornell Medicine, New York, NY 10065, USA.
301. Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02215, USA.
302. Department of Cellular and Molecular Medicine and Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA.
303. Department of Cellular and Molecular Medicine and Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA.
304. Centre for Computational Biology, Duke-NUS Medical School, Singapore 169857, Singapore.
305. Department of Computer Science, University of Helsinki, Helsinki 00014, Finland.
306. Institute of Biotechnology, University of Helsinki, Helsinki 00014, Finland.
307. Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki 00014, Finland.
308. Programme in Cancer and Stem Cell Biology, Centre for Computational Biology, Duke-NUS Medical School, Singapore 169857, Singapore.
309. Academic Department of Medical Genetics, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK.
310. MRC Cancer Unit, University of Cambridge, Cambridge CB2 0XZ, UK.
311. The University of Cambridge School of Clinical Medicine, Cambridge CB2 0SP, UK.
312. National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India.
313. Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, UK.
314. Department of Statistics, Columbia University, New York, NY 10027, USA.
315. Duke-NUS Medical School, Singapore 169857, Singapore.
316. Faculty of Medicine and Health Technology, Tampere University and Tays Cancer Center, Tampere University Hospital, Tampere Fl-33014, Finland.
317. Bakar Computational Health Sciences Institute and Department of Pediatrics, University of California, San Francisco, CA 94158-2549, USA.
318. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
319. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA.
320. Department of Oncology, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.
321. Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, CT 06520, USA.
322. Department of Computational Biology, University of Lausanne, Lausanne 1015, Switzerland.
323. Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva CH 1211, Switzerland.

324. Swiss Institute of Bioinformatics, University of Geneva, Geneva CH 1211, Switzerland.
325. Independent Consultant, Wellesley 02481, USA.
326. Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge CB1 8RN, UK.
327. Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge CB1 8RN, UK.
328. CIBER Epidemiología y Salud Pública (CIBERESP), Madrid 28029, Spain.
329. Research Group on Statistics, Econometrics and Health (GRECS), UdG, Barcelona 8041, Spain.
330. Oxford Nanopore Technologies, New York, NY 10013, USA.
331. Department of Medical Genetics, College of Medicine, Hallym University, Chuncheon 24252, South Korea.
332. Department of Experimental and Health Sciences, Institute of Evolutionary Biology (UPF-CSIC), Universitat Pompeu Fabra, Barcelona 08003, Spain.
333. Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
334. Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
335. Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Barcelona 08193, Spain.
336. Applications Department, Oxford Nanopore Technologies, Oxford OX4 4DQ, UK.
337. Department of Genetics, Microbiology and Statistics, University of Barcelona, IRSJD, IBUB, Barcelona 08028, Spain.
338. Department of Ophthalmology and Ocular Genomics Institute, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA 02114, USA.
339. Department of Medical and Clinical Genetics, Genome-Scale Biology Research Program, University of Helsinki, Helsinki 00100, Finland.
340. Big Data Institute, Li Ka Shing Centre, University of Oxford, Oxford OX3 7LF, UK.
341. Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford OX4 2PG, UK.
342. School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China.
343. Bioinformatics Unit, Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain.
344. Vector Institute, Toronto, ON M5G 0A3, Canada.
345. Cancer Division, Garvan Institute of Medical Research, Kinghorn Cancer Centre, University of New South Wales (UNSW Sydney), Sydney, NSW 2010, Australia.
346. South Western Sydney Clinical School, Faculty of Medicine, University of New South Wales (UNSW Sydney), Liverpool, NSW 2170, Australia.
347. West of Scotland Pancreatic Unit, Glasgow Royal Infirmary, Glasgow G31 2ER, UK.
348. Centre for Cancer Research, Victorian Comprehensive Cancer Centre, University of Melbourne, Melbourne, VIC 3010, Australia.
349. MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh EH4 2XU, UK.
350. Department of Biology, Bioinformatics Group, Division of Molecular Biology, Faculty of Science, University of Zagreb, Zagreb 10000, Croatia.
351. Department of Bioinformatics, Division of Cancer Genomics, National Cancer Center Research Institute, Tokyo 104-0045, Japan.

352. Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
353. University of Glasgow, Glasgow G61 1BD, UK.
354. Oregon Health and Science University, Portland, OR 97239, USA.
355. MRC-University of Glasgow Centre for Virus Research, Glasgow G61 1QH, UK.
356. School of Computing Science, University of Glasgow, Glasgow G12 8RZ, UK.
357. Molecular and Medical Genetics, OHSU Knight Cancer Institute, Oregon Health and Science University, Portland, OR 97239, USA.
358. Department of Surgery, University of Melbourne, Parkville, VIC 3010, Australia.
359. The Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, VIC 3052, Australia.
360. Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia.
361. University of Cologne, Cologne 50931, Germany.
362. The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada.
363. University of Ljubljana, Ljubljana 1000, Slovenia.
364. Department of Public Health Sciences, University of Chicago, Chicago, IL 60637, USA.
365. Research Institute, NorthShore University HealthSystem, Evanston, IL 60201, USA.
366. Department of Statistics, University of California Santa Cruz, Santa Cruz, CA 95064, USA.
367. Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK.
368. University of Toronto, Toronto, ON M5G 2M9, Canada.
369. Department of Computer Science, Carleton College, Northfield, MN 55057, USA.
370. Center for Psychiatric Genetics, NorthShore University HealthSystem, Evanston, IL 60201, USA.
371. Argmix Consulting, North Vancouver, BC V7M 2J5, Canada.
372. Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
373. Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
374. Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA.
375. Cancer Unit, MRC University of Cambridge, Cambridge CB2 0XZ, UK.
376. Department of Bioinformatics and Computational Biology and Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
377. Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, Fukuoka 812-8582, Japan.
378. Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD 21218, USA.
379. Bioinformatics Core Facility, University Medical Center Hamburg, Hamburg 20246, Germany.
380. Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Hamburg 20251, Germany.
381. Earlham Institute, Norwich NR4 7UZ, UK.
382. Norwich Medical School, University of East Anglia, Norwich NR4 7TJ, UK.
383. The Institute of Cancer Research, London SW7 3RP, UK.

384. University of East Anglia, Norwich NR4 7TJ, UK.
385. German Center for Infection Research (DZIF), Partner Site Hamburg-Borstel-Lübeck-Riems, Hamburg, Germany.
386. Division of Molecular Genetics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
387. Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, VIC 3000, Australia.
388. QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia.
389. Victorian Institute of Forensic Medicine, Southbank, VIC 3006, Australia.
390. University of Pennsylvania, Philadelphia, PA 19104, USA.
391. Department of Gynaecological Oncology, Westmead Hospital, Sydney, NSW 2145, Australia.
392. Genetics and Molecular Pathology, SA Pathology, Adelaide, SA 5000, Australia.
393. Centre for Cancer Research, Westmead Institute for Medical Research, University of Sydney, Sydney, NSW 2145, Australia.
394. Department of Clinical Pathology, University of Melbourne, Melbourne, VIC 3052, Australia.
395. Faculty of Medicine and Health, University of Sydney, Sydney, NSW 2145, Australia.
396. Department of Surgery, Pancreas Institute, University and Hospital Trust of Verona, Verona 37134, Italy.
397. Department of Surgery, Princess Alexandra Hospital, Brisbane, QLD 4102, Australia.
398. Surgical Oncology Group, Diamantina Institute, University of Queensland, Brisbane, QLD 4102, Australia.
399. Department of Diagnostics and Public Health, University and Hospital Trust of Verona, Verona 37134, Italy.
400. ARC-Net Centre for Applied Research on Cancer, University and Hospital Trust of Verona, Verona 37134, Italy.
401. Illawarra Shoalhaven Local Health District L3 Illawarra Cancer Care Centre, Wollongong Hospital, Wollongong, NSW 2500, Australia.
402. School of Biological Sciences, University of Auckland, Auckland 1010, New Zealand.
403. Department of Pathology and Diagnostics, University and Hospital Trust of Verona, Verona 37134, Italy.
404. Department of Medicine, Section of Endocrinology, University and Hospital Trust of Verona, Verona 37134, Italy.
405. Department of Pathology, Queen Elizabeth University Hospital, Glasgow G51 4TF, UK.
406. University of Sydney, Sydney, NSW 2006, Australia.
407. Department of Medical Oncology, Beatson West of Scotland Cancer Centre, Glasgow G12 0YN, UK.
408. Academic Unit of Surgery, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow Royal Infirmary, Glasgow G4 OSF, UK.
409. Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Sydney, NSW 2050, Australia.
410. Discipline of Surgery, Western Sydney University, Penrith, NSW 2751, Australia.
411. Institute of Cancer Sciences, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK.
412. Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW 2109,

Australia.
413. Sydney Medical School, University of Sydney, Sydney, NSW 2050, Australia.
414. School of Environmental and Life Sciences, Faculty of Science, The University of Newcastle, Ourimbah, NSW 2258, Australia.
415. Eastern Clinical School, Monash University, Melbourne, VIC 3128, Australia.
416. Epworth HealthCare, Richmond, VIC 3121, Australia.
417. Applied Tumor Genomics Research Program, Research Programs Unit, University of Helsinki, Helsinki 00290, Finland.
418. Olivia Newton-John Cancer Research Institute, La Trobe University, Heidelberg, VIC 3084, Australia.
419. Melanoma Institute Australia, University of Sydney, Sydney, NSW 2065, Australia.
420. Children's Hospital at Westmead, University of Sydney, Sydney, NSW 2145, Australia.
421. Australian Institute of Tropical Health and Medicine, James Cook University, Douglas, QLD 4814, Australia.
422. Bioplatforms Australia, North Ryde, NSW 2109, Australia.
423. Melanoma Institute Australia, Macquarie University, Sydney, NSW 2109, Australia.
424. Children's Medical Research Institute, Sydney, NSW 2145, Australia.
425. Discipline of Pathology, Sydney Medical School, University of Sydney, Sydney, NSW 2065, Australia.
426. School of Mathematics and Statistics, University of Sydney, Sydney, NSW 2006, Australia.
427. Diagnostic Development, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
428. Ontario Tumour Bank, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
429. PanCuRx Translational Research Initiative, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
430. UHN Program in BioSpecimen Sciences, Toronto General Hospital, Toronto, ON M5G 2C4, Canada.
431. Hepatobiliary/Pancreatic Surgical Oncology Program, University Health Network, Toronto, ON M5G 2C4, Canada.
432. Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada.
433. Division of Medical Oncology, Princess Margaret Cancer Centre, Toronto, ON M5G 2M9, Canada.
434. University of Nebraska Medical Center, Omaha, NE 68198-6880, USA.
435. Transformative Pathology, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.
436. Department of Biochemistry and Molecular Medicine, University California at Davis, Sacramento, CA 95817, USA.
437. University Health Network, Princess Margaret Cancer Centre, Toronto, ON M5G 1L7, Canada.
438. Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA.
439. Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON M5S 1A8, Canada.
440. Department of Pathology, Human Oncology and Pathogenesis Program, Memorial Sloan

Kettering Cancer Center, New York, NY 10053, USA.
441. Department of Pathology, University Health Network, Toronto General Hospital, Toronto, ON M5G 2C4, Canada.
442. Human Longevity Inc, San Diego, CA 92121, USA.
443. CRUKManchester Institute and Centre, Manchester M20 4GJ, UK.
444. Department of Radiation Oncology, University of Toronto, Toronto, ON M5S 1A8, Canada.
445. Division of Cancer Sciences, Manchester Cancer Research Centre, University of Manchester, Manchester M20 4GJ, UK.
446. Radiation Medicine Program, Princess Margaret Cancer Centre, Toronto, ON M5G 2M9, Canada.
447. Department of Surgical Oncology, Princess Margaret Cancer Centre, Toronto, ON M5G 2M9, Canada.
448. STTARR Innovation Facility, Princess Margaret Cancer Centre, Toronto, ON M5G 1L7, Canada.
449. Hefei University of Technology, Anhui 230009, China.
450. State Key Laboratory of Cancer Biology, and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Shaanxi 710032, China.
451. Fourth Military Medical University, Shaanxi 710032, China.
452. Department of Surgery, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China.
453. Leeds Institute of Medical Research, University of Leeds, St James's University Hospital, Leeds LS9 7TF, UK.
454. Canadian Center for Computational Genomics, McGill University, Montreal, QC H3A 0G1, Canada.
455. Department of Human Genetics, McGill University, Montreal, QC H3A 1B1, Canada.
456. International Agency for Research on Cancer, Lyon 69008, France.
457. McGill University and Genome Quebec Innovation Centre, Montreal, QC H3A 0G1, Canada.
458. Centre National de Génotypage, CEA - Institute de Génomique, Evry 91000, France.
459. Leeds Institute of Medical Research @ St James's, University of Leeds, St James's University Hospital, Leeds LS9 7TF, UK.
460. Institute of Mathematics and Computer Science, University of Latvia, Riga LV 1459, Latvia.
461. Department of Oncology, Gil Medical Center, Gachon University, Incheon 405-760, South Korea.
462. Department of Molecular Oncology, BC Cancer Research Centre, Vancouver, BC V5Z 1L3, Canada.
463. Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
464. Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo O310, Norway.
465. Department of Clinical Sciences, Lund, Division of Oncology and Pathology, Skåne University Hospital, Lund University, Lund 223 62, Sweden.
466. Translational Research Lab, Centre Léon Bérard, Lyon 69373, France.
467. Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Radboud University, Nijmegen 6500 HB, The Netherlands.
468. Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

469. Department of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam 1066 CX, The Netherlands.
470. Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo O310, Norway.
471. Department of Oncology, University of Cambridge, Cambridge CB2 1TN, UK.
472. Breast Cancer Translational Research Laboratory JC Heuson, Institut Jules Bordet, Brussels 1000, Belgium.
473. Department of Oncology, Laboratory for Translational Breast Cancer Research, KU Leuven, Leuven 3000, Belgium.
474. Translational Cancer Research Unit, GZA Hospitals St.-Augustinus, Center for Oncological Research, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp 2000, Belgium.
475. Department of Laboratory Medicine, Translational Cancer Research, Lund University Cancer Center at Medicon Village, Lund University, Lund SE-221 85, Sweden.
476. Icelandic Cancer Registry, Icelandic Cancer Society, Reykjavik 125, Iceland.
477. Department of Medical Oncology, Josephine Nefkens Institute and Cancer Genomics Centre, Erasmus Medical Center, Rotterdam 3015 CN, The Netherlands.
478. National Genotyping Center, Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan.
479. Department of Pathology, Oslo University Hospital Ulleval, Oslo 0450, Norway.
480. Faculty of Medicine and Institute of Clinical Medicine, University of Oslo, Oslo NO-0316, Norway.
481. Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
482. Department of Pathology, Skåne University Hospital, Lund University, Lund SE-221 85, Sweden.
483. Department of Pathology, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands.
484. Department of Pathology, College of Medicine, Hanyang University, Seoul 133-791, South Korea.
485. Department of Pathology, Asan Medical Center, College of Medicine, Ulsan University, Songpa-gu, Seoul 05505, South Korea.
486. The Netherlands Cancer Institute, Amsterdam 1066 CX, The Netherlands.
487. Breast Oncology Program, Dana-Farber/Brigham and Women's Cancer Center, Boston, MA 02115, USA.
488. Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.
489. Division of Breast Surgery, Brigham and Women's Hospital, Boston, MA 02115, USA.
490. Department of Clinical Science, University of Bergen, Bergen 5020, Norway.
491. Morgan Welch Inflammatory Breast Cancer Research Program and Clinic, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
492. The University of Queensland Centre for Clinical Research, Royal Brisbane and Women's Hospital, Herston, QLD 4029, Australia.
493. Department of Pathology, Institut Jules Bordet, Brussels 1000, Belgium.
494. Institute for Bioengineering and Biopharmaceutical Research (IBBR), Hanyang University, Seoul 133-791, South Korea.

495. University of Oslo, Oslo 0316, Norway.
496. Institut Bergonié, Bordeaux 33076, France.
497. Department of Research Oncology, Guy's Hospital, King's Health Partners AHSC, King's College London School of Medicine, London SE1 9RT, UK.
498. University Hospital of Minjoz, INSERM UMR 1098, Besançon 25000, France.
499. Cambridge Breast Unit, Addenbrooke's Hospital, Cambridge University Hospital NHS Foundation Trust and NIHR Cambridge Biomedical Research Centre, Cambridge CB2 2QQ, UK.
500. East of Scotland Breast Service, Ninewells Hospital, Aberdeen AB25 2XF, UK.
501. Department of Clinical Sciences, Lund, Division of Oncology and Pathology, Lund University, Lund 223 62, Sweden.
502. University of Copenhagen, Copenhagen 2200, Denmark.
503. Oncologie Sénologie, ICM Institut Régional du Cancer, Montpellier 34298, France.
504. Department of Radiation Oncology, Radboud University Nijmegen Medical Centre, Nijmegen 6525 GA, The Netherlands.
505. University of Iceland, Reykjavik 101, Iceland.
506. Dundee Cancer Centre, Ninewells Hospital, Dundee DD2 1SY, UK.
507. Institut Curie, INSERM Unit 830, Paris 75248, France.
508. Department of Laboratory Medicine, Radboud University Nijmegen Medical Centre, Nijmegen 6525 GA, The Netherlands.
509. Department of General Surgery, Singapore General Hospital, Singapore 169608, Singapore.
510. Universite Lyon, INCa-Synergie, Centre Léon Bérard, Lyon 69008, France.
511. Giovanni Paolo II / I.R.C.C.S. Cancer Institute, Bari BA 70124, Italy.
512. Department of Biopathology, Centre Léon Bérard, Lyon 69008, France.
513. Université Claude Bernard Lyon 1, Villeurbanne 69100, France.
514. Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
515. NCCS-VARI Translational Research Laboratory, National Cancer Centre Singapore, Singapore 169610, Singapore.
516. Department of Pathology, Erasmus Medical Center Rotterdam, Rotterdam 3015 GD, The Netherlands.
517. Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam 1066 CX, The Netherlands.
518. Institute of Human Genetics, Christian-Albrechts-University, Kiel 24118, Germany.
519. Institute of Human Genetics, Ulm University and Ulm University Medical Center of Ulm, Ulm 89081, Germany.
520. Hematopathology Section, Institute of Pathology, Christian-Albrechts-University, Kiel 24118, Germany.
521. Institute of Human Genetics, University of Ulm and University Hospital of Ulm, Ulm 89081, Germany.
522. Department of Human Genetics, Hannover Medical School, Hannover 30625, Germany.
523. Department of Pediatric Oncology, Hematology and Clinical Immunology, Heinrich-Heine-University, Düsseldorf 40225, Germany.
524. Department of Internal Medicine/Hematology, Friedrich-Ebert-Hospital, Neumünster 24534, Germany.
525. Pediatric Hematology and Oncology, University Hospital Muenster, Muenster 24534,

Germany.
526. Department of Pediatrics, University Hospital Schleswig-Holstein, Kiel 24105, Germany.
527. Department of Medicine II, University of Würzburg, Wuerzburg 97070, Germany.
528. Senckenberg Institute of Pathology, University of Frankfurt Medical School, Frankfurt 60596, Germany.
529. Institute of Pathology, Charité – University Medicine Berlin, Berlin 10117, Germany.
530. Department for Internal Medicine II, University Hospital Schleswig-Holstein, Kiel 24105, Germany.
531. Institute for Medical Informatics Statistics and Epidemiology, University of Leipzig, Leipzig 04109, Germany.
532. Department of Hematology and Oncology, Georg-Augusts-University of Göttingen, Göttingen 37073, Germany.
533. Institute of Cell Biology (Cancer Research), University of Duisburg-Essen, Essen D-45147, Germany.
534. MVZ Department of Oncology, PraxisClinic am Johannisplatz, Leipzig 04109, Germany.
535. Institute of Pathology, Ulm University and University Hospital of Ulm, Ulm 89081, Germany.
536. Department of Pathology, Robert-Bosch-Hospital, Stuttgart, Germany, Stuttgart 70376, Germany.
537. University Hospital Giessen, Pediatric Hematology and Oncology, Giessen 35392, Germany.
538. Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel 24118, Germany.
539. Institute of Pathology, University of Wuerzburg, Wuerzburg 97070, Germany.
540. Department of General Internal Medicine, University Kiel, Kiel 24118, Germany.
541. Clinic for Hematology and Oncology, St.-Antonius-Hospital, Eschweiler D-52249, Germany.
542. Department for Internal Medicine III, University of Ulm and University Hospital of Ulm, Ulm 89081, Germany.
543. Neuroblastoma Genomics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
544. Department of Pediatric Oncology and Hematology, University of Cologne, Cologne 50937, Germany.
545. University of Düsseldorf, Düsseldorf 40225, Germany.
546. Department of Vertebrate Genomics/Otto Warburg Laboratory Gene Regulation and Systems Biology of Cancer, Max Planck Institute for Molecular Genetics, Berlin 14195, Germany.
547. St. Jude Children's Research Hospital, Memphis, TN 38105-3678, USA.
548. Heidelberg University Hospital, Heidelberg 69120, Germany.
549. Genomics and Proteomics Core Facility High Throughput Sequencing Unit, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
550. Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
551. Institute of Pathology, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany.
552. Institute of Pathology, University Medical Center Hamburg-Eppendorf, Hamburg 20251, Germany.
553. Martini-Clinic, Prostate Cancer Center, University Medical Center Hamburg-Eppendorf, Hamburg 20095, Germany.

554. Division of Cancer Genome Research, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
555. National Institute of Biomedical Genomics, Kalyani 741235, West Bengal, India.
556. Advanced Centre for Treatment Research and Education in Cancer, Tata Memorial Centre, Navi Mumbai, Maharashtra 410210, India.
557. Department of Pathology, General Hospital of Treviso, Department of Medicine, University of Padua, Treviso 31100, Italy.
558. Department of Medicine (DIMED), Surgical Pathology Unit, University of Padua, Padua 35121, Italy.
559. Department of Pathology and Clinical Laboratory, National Cancer Center Hospital, Tokyo 104-0045, Japan.
560. Department of Pathology, Keio University School of Medicine, Tokyo 160-8582, Japan.
561. Department of Hepatobiliary and Pancreatic Oncology, National Cancer Center Hospital, Tokyo 104-0045, Japan.
562. Department of Pathology, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan.
563. Preventive Medicine, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan.
564. Department of Gastric Surgery, National Cancer Center Hospital, Tokyo 104-0045, Japan.
565. Department of Gastroenterology and Hepatology, Yokohama City University Graduate School of Medicine, Kanagawa 236-0004, Japan.
566. Department of Cancer Genome Informatics, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan.
567. Hiroshima University, Hiroshima 734-8553, Japan.
568. Tokyo Women's Medical University, Tokyo 162-8666, Japan.
569. Osaka International Cancer Center, Osaka 541-8567, Japan.
570. Wakayama Medical University, Wakayama 641-8509, Japan.
571. Hokkaido University, Sapporo 060-8648, Japan.
572. Department of Surgery, Yokohama City University Graduate School of Medicine, Kanagawa 236-0004, Japan.
573. Division of Medical Oncology, National Cancer Centre, Singapore 169610, Singapore.
574. Cholangiocarcinoma Screening and Care Program and Liver Fluke and Cholangiocarcinoma Research Centre, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand.
575. Lymphoma Genomic Translational Research Laboratory, National Cancer Centre, Singapore 169610, Singapore.
576. Center of Digestive Diseases and Liver Transplantation, Fundeni Clinical Institute, Bucharest 022328, Romania.
577. Department of Surgery, Division of Hepatobiliary and Pancreatic Surgery, School of Medicine, Keimyung University Dongsan Medical Center, Daegu 41931, South Korea.
578. Pathology, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona 8034, Spain.
579. Anatomia Patológica, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona 8036, Spain.
580. Spanish Ministry of Science and Innovation, Madrid 28046, Spain.
581. Hematology, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer

(IDIBAPS), University of Barcelona, Barcelona 8034, Spain.
582. Department of Biochemistry and Molecular Biology, Faculty of Medicine, University Institute of Oncology-IUOPA, Oviedo 33006, Spain.
583. Royal National Orthopaedic Hospital - Bolsover, London W1W 5AQ, UK.
584. Department of Pathology, Oslo University Hospital, The Norwegian Radium Hospital, Oslo O310, Norway.
585. Institute of Clinical Medicine and Institute of Oral Biology, University of Oslo, Oslo O310, Norway.
586. Department of Pathology (Research), University College London Cancer Institute, London WC1E 6BT, UK.
587. Research Department of Pathology, University College London Cancer Institute, London WC1E 6BT, UK.
588. East Anglian Medical Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK.
589. Royal National Orthopaedic Hospital - Stanmore, Stanmore, Middlesex HA7 4LP, UK.
590. Division of Orthopaedic Surgery, Oslo University Hospital, Oslo 0379, Norway.
591. Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK.
592. University of Pavia, Pavia 27100, Italy.
593. Karolinska Institute, Stockholm SE-171 76, Sweden.
594. University of Oxford, Oxford OX3 9DU, UK.
595. Salford Royal NHS Foundation Trust, Salford M6 8HD, UK.
596. Gloucester Royal Hospital, Gloucester GL1 3NL, UK.
597. Royal Stoke University Hospital, Stoke-on-Trent ST4 6QG, UK.
598. St Thomas's Hospital, London SE1 7EH, UK.
599. Imperial College NHS Trust, Imperial College, London W2 INY, UK.
600. Department of Histopathology, Salford Royal NHS Foundation Trust, Salford M6 8HD, UK.
601. Faculty of Biology, Medicine and Health, University of Manchester, Manchester M13 9PL, UK.
602. Edinburgh Royal Infirmary, Edinburgh EH16 4SA, UK.
603. Barking Havering and Redbridge University Hospitals NHS Trust, Romford RM7 0AG, UK.
604. King's College London and Guy's and St Thomas' NHS Foundation Trust, London SE1 7EH, UK.
605. Cambridge Oesophagogastric Centre, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK.
606. Nottingham University Hospitals NHS Trust, Nottingham NG7 2UH, UK.
607. St Luke's Cancer Centre, Royal Surrey County Hospital NHS Foundation Trust, Guildford GU2 7XX, UK.
608. University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
609. Norfolk and Norwich University Hospital NHS Trust, Norwich NR4 7UY, UK.
610. University Hospitals Coventry and Warwickshire NHS Trust, Coventry CV2 2DX, UK.
611. University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK.
612. Centre for Cancer Research and Cell Biology, Queen's University, Belfast BT9 7AB, UK.
613. School of Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK.
614. Wythenshawe Hospital, Manchester M23 9LT, UK.

615. Barts Cancer Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK.
616. Royal Marsden NHS Foundation Trust, London and Sutton SW3 6JJ, UK.
617. University Hospital Southampton NHS Foundation Trust, Southampton SO16 6YD, UK.
618. HCA Laboratories, London W1G 8AQ, UK.
619. University of Liverpool, Liverpool L69 3BX, UK.
620. Department of Surgery, Academic Urology Group, University of Cambridge, Cambridge CB2 0QQ, UK.
621. Department of Urology, James Buchanan Brady Urological Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.
622. Second Military Medical University, Shanghai 200433, China.
623. Department of Surgery and Cancer, Imperial College, London W2 INY, UK.
624. The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China.
625. Nuffield Department of Surgical Sciences, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK.
626. Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK.
627. Department of Surgery, University of Chicago, Chicago, IL 60637, USA.
628. Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA.
629. Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 4S6, Canada.
630. Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
631. Department of Pathology and Laboratory Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
632. Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH 44016, USA.
633. Research Health Analytics and Informatics, University Hospitals Cleveland Medical Center, Cleveland, OH 44106, USA.
634. Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, AB T2N 4N2, Canada.
635. Departments of Surgery and Oncology, University of Calgary, Calgary, AB T2N 4N2, Canada.
636. Buck Institute for Research on Aging, Novato, CA 94945, USA.
637. Duke University Medical Center, Durham, NC 27710, USA.
638. USC Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033, USA.
639. The Preston Robert Tisch Brain Tumor Center, Duke University Medical Center, Durham, NC 27710, USA.
640. Departments of Dermatology and Pathology, Yale University, New Haven, CT 06510, USA.
641. Fox Chase Cancer Center, Philadelphia, PA 19111, USA.
642. Department of Surgery, Division of Thoracic Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.
643. University of Michigan Comprehensive Cancer Center, Ann Arbor, MI 48109, USA.
644. University of Alabama at Birmingham, Birmingham, AL 35294, USA.
645. Division of Anatomic Pathology, Mayo Clinic, Rochester, MN 55905, USA.

646. Division of Experimental Pathology, Mayo Clinic, Rochester, MN 55905, USA.
647. International Genomics Consortium, Phoenix, AZ 85004, USA.
648. Departments of Pediatrics and Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
649. Department of Pathology, UPMC Shadyside, Pittsburgh, PA 15232, USA.
650. Center for Cancer Genomics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
651. Department of Neuro-Oncology, Istituto Neurologico Besta, Milano 20133, Italy.
652. The University of Queensland Thoracic Research Centre, The Prince Charles Hospital, Brisbane, QLD 4032, Australia.
653. Department of Neurosurgery, University of Florida, Gainesville, FL 32610, USA.
654. Center for Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA.
655. Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
656. Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
657. Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN 55905, USA.
658. Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL 33136, USA.
659. Department of Internal Medicine, Division of Medical Oncology, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
660. University of Tennessee Health Science Center for Cancer Research, Memphis, TN 38163, USA.
661. Centre for Translational and Applied Genomics, British Columbia Cancer Agency, Vancouver, BC V5Z 1L3, Canada.
662. Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA.
663. Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030, USA.
664. Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX 77030, USA.
665. Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
666. Indivumed GmbH, Hamburg 20251, Germany.
667. Department of Surgery, Division of Hepatobiliary and Pancreatic Surgery, School of Medicine, Keimyung University Dong-san Medical Center, Daegu 41931, South Korea.
668. Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA.
669. Department of Surgery, The George Washington University, School of Medicine and Health Science, Washington, DC 20052, USA.
670. Endocrine Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
671. National Cancer Center, Gyeonggi 10408, South Korea.
672. ILSbio, LLC Biobank, Chestertown, MD 21620, USA.
673. Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 4S6, Canada.

674. Gynecologic Oncology, NYU Laura and Isaac Perlmutter Cancer Center, New York University, New York, NY 10016, USA.
675. Division of Oncology, Stem Cell Biology Section, Washington University School of Medicine, St. Louis, MO 63110, USA.
676. Urologic Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
677. Institute for Systems Biology, Seattle, WA 98109, USA.
678. Department of Pathology and Laboratory Medicine, Center for Personalized Medicine, Children's Hospital Los Angeles, Los Angeles, CA 90027, USA.
679. Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH 43215, USA.
680. Department of Surgery, Duke University, Durham, NC 27710, USA.
681. Department of Obstetrics, Gynecology and Reproductive Services, University of California San Francisco, San Francisco, CA 94143, USA.
682. Departments of Neurology and Neurosurgery, Henry Ford Hospital, Detroit, MI 48202, USA.
683. Precision Oncology, OHSU Knight Cancer Institute, Oregon Health and Science University, Portland, OR 97239, USA.
684. Department of Pathology, Roswell Park Cancer Institute, Buffalo, NY 14263, USA.
685. Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Washington University School of Medicine, St. Louis, MO 63110, USA.
686. Penrose St. Francis Health Services, Colorado Springs, CO 80907, USA.
687. Department of Medicine, University of Chicago, Chicago, IL 60637, USA.
688. Department of Neurology, Mayo Clinic, Rochester, MN 55905, USA.
689. Center for Liver Cancer, Research Institute and Hospital, National Cancer Center, Gyeonggi 410-769, South Korea.
690. Department of Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
691. NYU Langone Medical Center, New York, NY 10016, USA.
692. Department of Hematology and Medical Oncology, Cleveland Clinic, Cleveland, OH 44195, USA.
693. Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, NC 27599, USA.
694. Helen F. Graham Cancer Center at Christiana Care Health Systems, Newark, DE 19713, USA.
695. Cureline, Inc, South San Francisco, CA 94080, USA.
696. Department of Obstetrics and Gynecology, Medical College of Wisconsin, Milwaukee, WI 53226, USA.
697. Hematology and Medical Oncology, Winship Cancer Institute of Emory University, Atlanta, GA 30322, USA.
698. Vanderbilt Ingram Cancer Center, Vanderbilt University, Nashville, TN 37232, USA.
699. Ohio State University College of Medicine and Arthur G. James Comprehensive Cancer Center, Columbus, OH 43210, USA.
700. Research Computing Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

701. Analytical Biological Services, Inc, Wilmington, DE 19801, USA.
702. Department of Dermatology, University Hospital of Essen, Essen 45122, Germany.
703. University of Pittsburgh, Pittsburgh, PA 15213, USA.
704. Murtha Cancer Center, Walter Reed National Military Medical Center, Bethesda, MD 20889, USA.
705. Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.
706. Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
707. Center for RNA Interference and Noncoding RNA, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
708. Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
709. Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
710. Department of Urology, Mayo Clinic, Rochester, MN 55905, USA.
711. Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
712. Departments of Neurosurgery and Hematology and Medical Oncology, Winship Cancer Institute and School of Medicine, Emory University, Atlanta, GA 30322, USA.
713. Georgia Regents University Cancer Center, Augusta, GA 30912, USA.
714. Thoracic Oncology Laboratory, Mayo Clinic, Rochester, MN 55905, USA.
715. Institute for Genomic Medicine, Nationwide Children’s Hospital, Columbus, OH 43205, USA.
716. Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Mayo Clinic, Rochester, MN 55905, USA.
717. International Institute for Molecular Oncology, Poznań 60-203, Poland.
718. Poznan University of Medical Sciences, Poznań 61-701, Poland.
719. Edison Family Center for Genome Sciences and Systems Biology, Washington University, St. Louis, MO 63110, USA.