Supplementary Methods

Library Preparation and Sequencing

DNA was extracted from blood lymphocytes by standard DNA extraction methods. The DNA concentration was determined using a Qubit dsDNA BR Assay (Invitrogen, Oregon, USA). Adapter-tagged libraries were prepared from 50 ng of genomic DNA according to the TruSight™ Rapid Capture workflow (Illumina) using a transposase-based method (Nextera) (Marine et al., 2011). The DNA was enzymatically tagmented and adaptor ligated. Prior to target sequence enrichment, 12 processed samples were pooled in equimolar amounts (500 ng of each). The pooled DNA was then hybridized to biotinylated oligonucleotides designed to enrich the ROIs, which were then captured by streptavidin magnetic beads. To increase specificity, two rounds of target sequence enrichment were performed. The final DNA libraries were quality checked on an Agilent 2100 Bioanalyzer using a high-sensitivity chip, and the DNA concentration was measured by the Qubit dsDNA BR assay.

The pool was denatured, diluted to a final concentration of 7-12 pM, and sequenced on a MiSeq instrument following the Illumina 300 sequencing cycle program with paired-end reads. Demultiplexing generated two FASTQ files per sample corresponding to the reverse and forward sequences, respectively.

Preprocessing

Sequence alignment (used for all methods except NextGENe and SeqNext)

Data was aligned to human reference genome build b37 using BWA vs. 0.7.12. Using SAMtools vs. 0.1.19 duplicate reads were removed and mate information was fixed. Local realignment around indels and base quality score recalibration was performed with GATK vs. 3.5.

```
## Alignment

bwa aln -t 10 $ref $file"_R1_001.fastq.gz" > $file"_R1.sai"

bwa aln -t 10 $ref $file"_R2_001.fastq.gz" > $file"_R2.sai"

bwa sampe -r "@RG\tID:group1\tSM:"$file"\tPL:ILLUMINA" $ref \
$file"_R1.sai" $file"_R2.sai" $file"_R1_001.fastq.gz"
$file"_R2_001.fastq.gz" | \
samtools view -bS - | \
samtools sort - $file.sorted

## Remove duplicates and fix mate information

samtools rmdup $file.sorted.bam - | \
samtools fixmate - $file".rmdup.fixmate.bam"

samtools index $file.rmdup.fixmate.bam

## Local realignment around indels

java -Xmx64g -Djava.io.tmpdir=$temp \
-jar $GATK/GenomeAnalysisTK.jar \
-log $file.RealignerTargetCreator.log \
```

```
-I $file.rmdup.fixmate.bam \
-T RealignerTargetCreator \
-nt 8 \
-DBQ 1 \
-known $datasets/Mills_and_1000G_gold_standard.indels.b37.vcf \
-known $datasets/1000G_phase1.indels.b37.vcf \
-R $ref \
-o $file.input.bam.list

java -Xmx64g -Djava.io.tmpdir=$temp \
-jar $GATK/GenomeAnalysisTK.jar \
-log $file.IndelRealigner.log \
-I $file.rmdup.fixmate.bam \
-R $ref \
-DBQ 1 \
-T IndelRealigner \
-compress 3 \
-known $datasets/Mills_and_1000G_gold_standard.indels.b37.vcf \
-known $datasets/1000G_phase1.indels.b37.vcf \
-targetIntervals $file.input.bam.list \
-o $file.realigned.bam

## Quality score recalibration

java -Xmx64g -Djava.io.tmpdir=$temp \
-jar $GATK/GenomeAnalysisTK.jar \
-log $file.BaseRecalibrator.log \
-T BaseRecalibrator \
-cov CycleCovariate \
-cov ContextCovariate \
-I $file.realigned.bam \
-R $ref \
-nct 8 \
-knownSites $datasets/dbsnp_138.b37.vcf \
-knownSites $datasets/Mills_and_1000G_gold_standard.indels.b37.vcf \
-knownSites $datasets/1000G_phase1.indels.b37.vcf \
-o $file.recal_data.table

java -Xmx64g -Djava.io.tmpdir=$temp \
-jar $GATK/GenomeAnalysisTK.jar \
-log $file.PrintReads.log \
-T PrintReads \
-I $file.realigned.bam \
-R $ref \
-dcov 500 \
-nct 8 \
-BQSR $file.recal_data.table \
-o $file.realigned.recal.bam
```

Read counts for panelcn.MOPS and ExomeDepth

RCs were calculated in R using panelcn.MOPS' function countBamListInGRanges with the read.width parameter set to 150.

DepthOfCoverage files for VisCap

DepthOfCoverage files were generated using GATK vs. 3.5 and human reference genome build b37.

```
java -Xmx64g -Djava.io.tmpdir=$temp \
-jar $GATK/GenomeAnalysisTK.jar \
-T DepthOfCoverage \
-R $ref \
-o $file.cov \
-I $file.realigned.recal.bam \
-L Genes.bed
```

Methods Compared

panelcn.MOPS

To increase sensitivity, we optimized the expected fold change of the copy number classes. After experiments on the validation set, we set the parameter to 0.57 and 1.46 for CN1 (deletion) and CN3 (duplication), respectively. Additionally, quantile normalization and re-scaling provided the best results.

ExomeDepth

ExomeDepth (Plagnol et al., 2012) fits a beta-binomial model to the coverage ratio which is calculated by dividing the RCs of each ROI of the test sample by the sum of the RCs from test sample and controls. First ExomeDepth ranks all control samples according to the correlation of their RCs to the RCs of the test sample. Samples are added sequentially to the aggregate reference set and the beta-binomial model is fitted. Fitting is stopped if the posterior probability in favor of an artificial single-exon heterozygous deletion call fails to increase, which might already happen after adding one sample to the reference set. In the second step, ExomeDepth calculates the likelihood for a deletion, duplication or normal copy number for each exon, and finally uses a hidden Markov model (HMM) for segmentation. The statistical support for each CNV is quantified by the Bayes factor. According to the authors, the correlation between test and reference should be larger than 0.97 to obtain meaningful results; test samples that do not reach this threshold are therefore considered to be of low quality. There is, however, no quality control measure for ROIs. ExomeDepth was originally developed as an R package for CNV detection in WES data, but it has already been applied successfully to targeted NGS panel sequencing data.

We replaced the default read counting procedure of ExomeDepth version 1.1.8 with the read counting procedure of panelcn.MOPS, which improved the performance of the method. Further, we optimized the parameters corresponding to the expected CNV length and the transition probability. The best performing choices were 100,000 for the expected CNV length and 0.01 for the transition probability (the default values are 50,000 and 0.0001, respectively).

CoNVaDING

CoNVaDING (Copy Number Variation Detection In Next-generation sequencing Gene panels) (Johansson et al., 2016) calculates the ratio of normalized average read depth of the test sample to that of the selected controls. Subsequently CoNVaDING performs a distribution analysis using the Z-

score for each ROI. Control samples that are most informative for a specific test sample are selected by looking at the overall coverage pattern. CoNVaDING implements three different quality control metrics. The first checks the fit of the selected control samples, the second shows the variability between all ROIs within the test sample, and the third computes the variability for each ROI within the control samples. CoNVaDING is implemented as a Perl command line script that depends on SAMtools (Li et al., 2009) for mean coverage calculations and duplicate marking.

Using version 1.1.6 of CoNVaDING, we optimized the number of samples to be used as controls by the method. The default value of 30 worked best on the validation set. Additionally, we increased the lower ratio cutoff value to 0.71 and decreased the higher ratio cutoff value to 1.35 in order to increase sensitivity without losing specificity. For the rest of the parameters the default values yielded the best performance. Samples with a sample ratio larger than 0.09 were considered to be of low quality.

VisCap

VisCap (Pugh et al., 2016) detects CNVs based on the log2 ratio of the sample-specific fractional coverage of each ROI divided by the median for the ROI across the entire batch of control samples. The quality control for samples is based on the whiskers of the boxplot of log2 ratios for each sample. No quality control criterion is provided for ROIs. VisCap is implemented as an R script on top of GATK's DepthOfCoverage tool (McKenna et al., 2010).

We downloaded VisCap from https://github.com/pughlab/VisCap and optimized the thresholds for the log2 ratios. A lower threshold of - 0.4 and an upper threshold of 0.3 increased the sensitivity compared to the default values of - 0.55 and 0.4 without loss of specificity.

NextGENe

Similar to ExomeDepth (Plagnol et al., 2012), the commercial tool NextGENe (Softgenetics, State College, PA, USA) fits a beta-binomial model to the normalized coverage ratios to model the amount of noise (dispersion). For segmentation, an HMM uses likelihood values that are based on the coverage ratios and the amount of dispersion to calculate CNs for each ROI. A high dispersion without large CNVs is an indicator of a low-quality sample, but there is no quality measure for ROIs.

NextGENe version 2.4.1.1 only accepts a maximum of 10 control samples. Analysis of the validation set showed that the program performed best when controls either from the same sequencing run or at least from recent runs were used. In accordance with our experiments on the validation set, we chose the default option of dispersion and HMM with normalized counts and auto fitting and used the average of multiple controls. If a test sample without a large CNV showed a dispersion value larger than 0.005, the analysis was rerun with a different set of 10 controls. If the dispersion value remained too high, the sample was considered to be of low quality.

SeqNext

The commercial tool SeqNext (*JSI* Medical Systems GmbH, Kippenheim, Germany) first computes the normalized coverage for each target ROI, that is, the ratio of the coverage of target ROIs to the sum of all other ROIs for the test sample and for one or more control sample(s). The relative coverage of each target ROI of the test sample is then calculated as the ratio between the normalized coverage of the test sample and the average normalized coverage of the control samples. The minimum standard deviation is determined by subtracting the average normalized coverage of the control samples from the normalized coverage of the test sample and dividing the result by the standard deviation of the

control samples. If the relative coverage threshold and the minimum standard deviation exceed user defined upper or lower limits, a CNV is called. There are several quality control metrics for both samples and ROIs. A CNV that does not pass all quality controls is marked as skipped.

Using version 4.3.0 we optimized the settings for the relative coverage threshold on the validation set. The default values of 75% and 125% achieved the best results. Since the default values achieved 100% specificity and changes in the parameters did not increase the sensitivity, the final analyses were performed with default parameters. ROIs with a CNV marked as skipped were considered to be of low quality.

Optimal parameter settings for each method

panelcn.MOPS

correlation threshold for control sample selection: 0.99
minimal number of controls: 8
maximal number of controls: `maxControls = 25`
normalization parameters: `normType = quant, sizeFactor = quant`
expected fold change for deletion: 0.57
expected fold change for amplification: 1.46
`I = c(0.025, 0.57, 1, 1.46, 2)`

ExomeDepth

expected CNV length: `expected.CNV.length = 100,000`
transition probability: `transition.probability = 0.01`

CoNVaDING

control samples: `controlSamples=30`
region threshold: `regionThreshold 20`
lower ratio cutoff value: `ratioCutOffLow 0.71`
higher ratio cutoff value: `ratioCutOffHigh 1.35`

VisCap

lower threshold for the log2 ratios: -0.4
upper threshold for the log2 ratios: 0.3
`threshold.cnv_log2_cutoffs = c(-0.4, 0.30)`
minimum number of exons: `threshold.min_exons = 1`

NextGENe

model: dispersion and HMM with normalized counts
control samples: average of multiple controls
dispersion HMM settings: auto fitting

SeqNext

lower relative coverage threshold: 75%
upper relative coverage threshold: 125%

References

Johansson LF, van Dijk F, Boer EN de, van Dijk-Bos KK, Jongbloed JDH, van der Hout AH, Westers H, Sinke RJ, Swertz MA, Sijmons RH, Sikkema-Raddatz B. 2016. CoNVaDING: Single exon variation detection in targeted NGS data. Hum Mutat 37:457–464.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

Marine R, Polson SW, Ravel J, Hatfull G, Russell D, Sullivan M, Syed F, Dumas M, Wommack KE. 2011. Evaluation of a transposase protocol for rapid generation of shotgun high-throughput sequencing libraries from nanogram quantities of DNA. Appl Environ Microbiol 77:8071–8079.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303.

Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, Wood NW, Hambleton S, Burns SO, Thrasher AJ, Kumararatne D, Doffinger R, et al. 2012. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 28:2747–2754.

Pugh TJ, Amr SS, Bowser MJ, Gowrisankar S, Hynes E, Mahanta LM, Rehm HL, Funke B, Lebo MS. 2016. VisCap: inference and visualization of germ-line copy-number variants from targeted clinical sequencing data. Genet Med 18:712–719.

Supp. Table S2: Number of Samples per CN Category for each Dataset

	Validation Set	Test Sets			
	TSC Panel (n=25)	**TSC Panel (n=141)**	**Small CNVs (n=4)**	**Custom Panel (n=10)**	**Total (n=180)**
CN2	13	110		5	**128**
Multi-Exon Deletions	5	12		2	**19**
Multi-Exon Duplications	2	1			**3**
Single-Exon Deletions		12		2	**15**
Single-Exon Duplications		2			**2**
Whole-Gene Deletions	4	1		1	**5**
Deletions < 1 ROI	1		2		**3**
Duplications < 1 ROI			2		**2**
	Cases excluded from further analysis				
***De novo Alu* Insertions**		3			**3**