SUPPLEMENTARY NOTES

DNA Fragmentation

Human solution hybrid selection

For the 189 human DNA samples for the hybrid selection (**application 1**), between 200 and 4,800 ng genomic DNA (final volume 200ul in 0.1xTE), were sheared in MicroAmp Fast 96-well reaction plates 0.1ml (Life Technology - Applied Biosystems, cat.#: 4346907) with the following parameters on the Covaris E210 instrument: off-set: 10mm, duty cycle: 10%, intensity: 5, cycles/burst: 200, duration: 780sec.

Human whole genome shotgun (WGS) sequencing

Ten genomic DNA samples (3µg DNA each, final volume 130ul in 0.1xTE) from human cell-lines for whole genome shotgun sequencing (**application 2**) were sheared in microTUBEs (Covaris, cat.#: 520045) with the following conditions on the Covaris E210 instrument: duty cycle: 5%, intensity: 10, cycles/burst: 200, duration: 120sec.

Bacterial sequencing

For the 12 *E. coli* bacterial libraries for whole genome sequencing (**application 3**), 3µg genomic DNA per sample (final volume 200ul in TE) were sheared in a MicroAmp Fast 96-well reaction plate with the following parameters on the Covaris E210 instrument: off-set: 10mm, duty cycle: 10%, intensity: 5, cycles/burst: 200, duration: 900sec. One third of this was then used for library preparation (diluted in TE).

Reaction Clean-up

MagNA composition

0.1% carboxyl-modified Sera-Mag Magnetic Speed-beads (FisherSci, cat.#: 09-981-123)
18% PEG-8000 (w/v) (e.g. Sigma Aldrich, cat.#: 89510)
1M NaCl
10mM Tris-HCl, ph 8.0
1 mM EDTA, pH 8.0
Optional: 0.05% Tween 20
note: beads contain sodium azide, wash beads 2x with water or TE before addition.
Store MagNA kit in the refrigerator in the dark and vigorously mix before use; kit is stable for at least 3 months.

General use of MagNA

Different volume ratios of MagNA to DNA result in different molecular size-cut-offs (Figure S5). We use different ratios throughout the protocol for the lower size-cut-off.

General use (as for AMPure XP) is as follows: After several mixing cycles of MagNA and DNA, a 4-5min incubation period for DNA precipitation on the beads is performed. Separation on a 96-well ring-magnet (e.g. Agencourt SPRIPlate, A32782) is done for different time intervals for each of the different clean-up steps, depending on the volume of the mix; visually inspect that solution has cleared. After supernatant is removed (or transferred in case of the first step in the dual size selection, see below), fresh 70% ethanol (150-200µl) is added while still on the magnet. After a short incubation time (~30 seconds), ethanol is removed and a second ethanol-wash cycle is performed. The beads are air-dried for about 5-15min at room temperature on the magnet. Elution is done in low salt buffers (usually 0.1xTE) in different volumes throughout the protocol while the plate is off the magnet. Eluted DNA (without beads) is transferred into the next enzymatic reaction. Carry-over of some beads does not negatively interfere with the enzymatic reactions (Fisher, Barry, Abreu, Minie, Nolan, Delorey, Young, Fennell, Allen, Ambrogio et al. 2011).

Comparison of MagNA with Agencourt AMPure XP

To compare the performance of MagNA with the commercial product (Agencourt AMPure XP, Beckman Coulter), the ratio of bead suspension to DNA was varied as indicated on the gel picture (Figure S5) between 0.9x and 3.0x. Recommended ratio is 1.8x (e.g. 36ul SPRI-beads and 20ul size marker 1ug). 1µg size marker (GeneRuler 50bp DNA ladder, Fermentas) was used for each reaction and, beside ratio variation, clean-up was performed according to the AMPure XP protocol.

Dual Fragment Size Selection

General remarks on dual size selection

Dual size selections with SPRI-beads are not as precise as size selections via gel-cuts or with commercial systems such as the Pippin Prep. In particular, dual size selections only lead to a reduction of unwanted fragment sizes, but do not eliminate them completely. In addition, some of the shorter, desired, fragments bind in the first step to the beads, which get removed, thus reducing the amount of material (however, sample loss also occurs during gel-based size selections). Nevertheless, dual size selection has advantages in that it is automatable, can be done in a 96-well set-up (Borgstrom, Lundin and Lundeberg 2011; Lennon, Lintner, Anderson, Alvarez, Barry, Brockman, Daza, Erlich, Giannoukos, Green et al. 2010), and is fast and cheap. Finally, the risk of cross-contamination is lower than when running several samples in parallel on the same gel.

Bacterial and human WGS (mean insert size ~300bp)

Fragments larger than 400bp bind to the carboxyl coated beads by adding 0.85x volume MagNA to the sample solution after fragmentation for **application 2 and 3**. After incubation, the beads are discarded and the supernatant is added to 0.15x fresh MagNA to bind the remaining shorter DNA, but exclude fragments shorter than ~200bp (example: 100ul DNA + 85ul MagNA; 185ul supernatant + 27.75ul MagNA; final MagNA: 1.13x). Washing and elution steps are carried out as described above.

Sample Barcoding

General remarks on the enzymatic reactions

The basic workflow, enzymatic steps and oligonucletide sequences are almost identical for the three different applications. Differences are highlighted in the following paragraphs and arose over time, mainly by implementing knowledge from recent publications, such as different DNA polymerases for enrichment PCR (Aird, Ross, Chen, Danielsson, Fennell, Russ, Jaffe, Nusbaum and Gnirke 2011), and because we tried to automate the entire process as far as our robot allows, which resulted mainly in different reaction and elution volumes. Workflow and specific reaction conditions for the 3 applications are given in Figure S1, S3, S4.

Blunting and 5' phosphorylation

Fragmented (for **application 2 and 3** also size selected) and concentrated DNA is blunted and 5'-phosphorylated using the NEB Quick Blunting Kit (E1201). For **application 1 and 2** 0.2-4.8µg input DNA were blunted using the recommended amounts of enzyme mix and buffers. As most samples we process do not have more than 1µg of input DNA, we are currently (as for **application 3**) only using ¼ of the recommended enzyme mix and substituting the kit buffers with a 'home-made' 10x Blunting Buffer and dNTP-Mix (10x Blunting Buffer: 100mM MgCl2, 500mM Tris-HCl pH 7.5, 75mM DTT; dNTP-Mix: 10mM ATP, 2mM dATP, 2mM dCTP, 2mM dGTP, 2mM dTTP). The blunting enzyme mix consists of two enzymes, with both enzymes having different temperature optima. We observed that by not incubating at room temperature as recommended in the manufacturer's instruction, but instead incubating consecutively at the two optima, the blunting reaction became more efficient (data not shown). Incubation temperatures and time intervals were therefore changed to 20min at 12°C followed by 15min at 37°C and implemented in **application 2 and 3**.

Clean-up

2x volume MagNA is added to the finished blunting reaction by the liquid handling robot, mixed, incubated for 5min and beads were collected on the walls of the 96-well plate using a magnet.

Supernatant is removed and 150µl fresh 70% Ethanol is added. After 30-40 sec incubation, the ethanol is removed and the ethanol-wash is repeated a second time. Beads are air-dried for about 5min at room temperature. The clean-up processes throughout the protocol are almost completely automated; only the plate-transfer to and from the magnet is done manually.

Adapter

Six-mer barcodes were designed such that they differ by at least 2 nucleotides and such that not more than 2 consecutive nucleotides are the same base; adapter sequences are given in Table S6a.

We are using a set of 159 truncated barcoded-P5-adapters; the truncated PE-P7-adapter is always identical and exhibits the Illumina paired end sequences. To achieve a higher degree of barcoding, a combination of barcoded-P5- and barcoded-P7-adapter (not provided) could be used or the number of bases exhibiting the barcode could be increased.

All adapters were ordered from IDT (Coralville, IA, USA) without modifications or special purification, with the goal of increasing yield and reducing costs.

Each barcoded-P5 is hybridized with its short barcoded reverse complement counterpart (barcoded-P5-comp) to obtain partially double stranded barcoded-P5-adapters by preparing the following reaction: 200µM barcoded-P5, 200µM barcoded-P5-comp, 50mM NaCl, 1mM Tris-HCl pH8.0, 0.1mM EDTA pH 8.0. Similarly, prepare such a mix for the universal PE-P7-adapter and incubate in a thermo cycler for 10sec at 95°C, followed by a slow ramp to room temperature at 0.1°C/sec. 2µl of barcoded-P5-adapter and 2µl of PE-P7-adapter are used per ligation reaction.

Ligation reaction

This step was slightly different for all 3 applications, because we further automated the process by no longer eluting in a small volume (11µl) manually (as was done for **application 1**), but rather in higher volume with the robot for the microbial sequencing (**application 3**). As potential ligation biases are a concern for blunt-end ligation, each of the 10 samples for the human WGS (**application 2**) were tagged with 4 different barcoded adapters in individual reactions, with each barcode ending with one of the four nucleotides. Therefore the elution volume was higher and split into 4 ligation reactions per sample (= 4 libraries). The amount of ligase and the buffer concentrations were identical for all three applications. Only the reaction volume and adapter concentration changed slightly for **application 3**, where we eluted in diluted adapter to automate the protocol further.

Clean-up

The ligation reactions were purified by adding 1.6x volume MagNA, and subsequently followed by incubation, washing and drying steps as described above.

Nick fill-in

A strand displacing DNA polymerase is used to close nicks and replace and extend the short complementary part of both truncated adapters. Concentrations and conditions are identical for **application 1 and 2**, but slightly different for **application 3** because of the increased elution volume of the cleaned fragments to allow further automation.

Clean-up

Identical to the previous clean-up step (1.6x volume of MagNA), but elution volume differs for the three applications.

Enrichment PCR or amplification of truncated libraries

At least a few cycles of amplification are necessary to amplify fragments with both adapters on either side; fragments with unwanted (truncated) adapter combinations (P5-P5 or P7-P7) will not be sequenced, but would compete in the hybrid selection step for **application 1**. For sequencing, the truncated adapter sites have to be completed via enrichment PCR to allow hybridization to the flow cell.

For **application 1**, the primer pair (PreHyb-PE) does not extend the adapter sites, as these short adapters seem to minimize 'daisy-chaining' during hybrid selection, which refers to off-target molecules hybridizing to the adapter site of an on-target molecule that gets selected (Figure 2b main manuscript). Enrichment PCR was performed after pooled hybrid selection was finished (see below).

For **application 2 and 3,** primer pair Sol-PE-PCR was used for enrichment PCR. The originally recommended DNA polymerase was used for **application 3** (Phusion HF DNA polymerase, Finnzymes) and, as the AccuPrimer HF *Taq* DNA polymerase (Invitrogen) in combination with extended melting periods was shown to eliminate GC-biases associated with many polymerases (Aird, Ross, Chen, Danielsson, Fennell, Russ, Jaffe, Nusbaum and Gnirke 2011), AccuPrime HF *Taq* DNA polymerase with buffer II was used for **application 2**. For the human WGS (**application 2**) we performed 4 PCR reactions per library per sample and the number of PCR cycles for this enrichment step was individually determined per library according to qPCR measurement prior to enrichment PCR. This QC and adjustment step was omitted for **application 1 and 3** for the sake of higher throughput. We performed 6 cycles of amplification of the truncated libraries for **application 1**, 12 cycles of enrichment PCR for **application 3** and between 12 and 14 cycles **for application 2** with the respective annealing temperature for the respective primer pair and cycling conditions for the respective DNA polymerase.

Instead of the universal Sol-PE-PCR primer to complete the adapter sites for sequencing, indexed PCR primer can be used to further increase the magnitude of pooling (see Figure 2a main manuscript). Primer sequences and annealing temperatures are given in Table S6b (only a subset of 16 indexing PCR primer are given, see (Meyer and Kircher 2010) for further indexing PCR primer).

Clean-up

For **application 1 and 3** this clean-up step was performed as described above (1.6x MagNA). For **application 2**, the Qiagen MinElute Kit was used.

Copy Number Determination for Equimolar Library Pooling

SYBR Green qPCR assay

The number of molecules with both completed adapter sites (P5 and P7) was determined for each of the 94 finished libraries individually for **application 1**, that were pooled into pool sizes of 14, 28 and 52 libraries before hybrid capture, using the DyNAmo HS SYBR Green qPCR kit (Fisher Scientific: F410-L) with primer pair PreHyb-PE (Table S6b). Normalization was done according to the mean of 2 measurements per library. The same SYBR Green qPCR kit was utilized with the Sol-PE-qPCR primer pair for **application 3**. Normalization of the 12 libraries was performed according to the mean of 2 measurements for these libraries.

Sequencing

One lane of 36 cycles single read Illumina sequencing was performed for a pool of the human WGS libraries (**application 2**). Normalization and subsequent pooling for deep sequencing was performed according to the number of reads per barcode, i.e. library.

Although no normalization was carried out for the pool of 95 samples in **application 1**, for data production for this project, we are performing copy number determination of the libraries and normalization before pooling. Copy numbers per truncated library are determined by sequencing. We are constantly reusing the set of barcoded adapters (set of 159), which means that only up to 159 samples can be pooled per sequencing lane to determine the copy number for each individual library. To estimate copy number for a large number of libraries simultaneously (e.g. the 2,152 libraries we made for our prostate cancer hybrid capture study), we thus took advantage of two dimensions of barcoding, our internal barcoding and the external barcoding provided by the Illumina indexing read. Specifically, to estimate the copy number, we introduced an index to the PE-P7 adapter site per pool of up to 144 samples during the enrichment PCR (indexing PCR; Figure 2a main text, Table S6b for 16 indexing PE primer (Meyer and Kircher 2010)).

Pooled Solution Hybrid Selection (14, 28, 52; 95 and 81 libraries per pool)

After normalization and pooling of 14, 28 and 52 samples (**application 1**), respectively, the three pools were purified with 1.8x MagNA. To achieve the required amount of 'pond' for hybrid selection of 500ng per pool, 10 PCR cycles with PreHyb-PE were performed and the products were cleaned with the Qiagen MinElute Kit.

A subsample of each finished truncated library for the pool of 95 samples was used (10µl, not normalized), pooled and purified with 1.8x MagNA. No amplification was necessary for this pool, as the required input amount was achieved by pooling the 95 samples.
81 libraries with similar complexity (as measured by counted molecules via sequencing) with at most 5x difference were normalized via cherry picking and pooled. Clean-up with 1.8x MagNA and 8 PCR cycles with PreHyb-PE was performed, followed by another MagNA clean-up.
For the hybrid selection, instructions were followed as given in the user guide (Hybridization in: G3360-90000_SureSelect_IlluminaPaired_1.1.1) but SureSelect PE Block #3 was substituted by Univ_Block (concentration: 315µM Univ_Block_P5, 315µM Univ_Block_P7; concentration in hybridization: 6.5µM; sequences in Table S7c). After 15 cycles of enrichment PCR with the Sol-PE-PCR primer pair to complete the adapter sites, a final clean-up and concentration step with 1.6x MagNA was performed. For the 81 library pool (cherry-picking) 16 cycles 'off-bead' PCR (with indexing PCR primer) was performed instead of NaOH treatment after capture (Fisher, Barry, Abreu, Minie, Nolan, Delorey, Young, Fennell, Allen, Ambrogio et al. 2011).

Sequencing and Analysis

Human solution hybrid selection – application 1

The three library pools of 14, 28 and 52 samples, enriched for the 2.2Mb target region in three independent reactions, were pooled in equimolar ratios and sequenced together on one lane single read 36 cycles on a GAII instrument. The pool of 95 samples was sequenced on one lane HiSeq2000 using 50 cycles paired end reads. The pool of 81 samples (cherry picking) was sequenced together with 3 other indexed libraries on one lane HiSeq2000 50 cycles paired end. Sequences were aligned to the human genome (hg18) using BWA (Li and Durbin 2009) by excluding the first 6 bases of the first read (= barcode as an identifier for the sample). Duplicates were removed in the resulting BAM files for each sample individually (according to the barcode) using Picard's MarkDuplicates(http://picard.sourceforge.net). Hybrid selection performance was calculated with Picard's CalculateHsMetrics.

Human WGS

The pool of 40 libraries (**application 2**) was sequenced on 58 lanes of 100 cycle paired end reads on HiSeq2000 instruments and analyzed. Sequences were mapped to the human genome (hg19) using BWA (Li and Durbin 2009) excluding the first 6 bases of the first read. Duplication rates were calculated with SAMtools (Li, Handsaker, Wysoker, Fennell, Ruan, Homer, Marth, Abecasis and Durbin 2009).

Microbial sequencing

The pool of 12 microbial libraries (**application 3**) was sequenced on a HiSeq2000 instrument on one lane with 50 cycles of paired end reads. Reads were aligned to the *E .coli* genome (K12, substrain MG1655) using BWA (Li and Durbin 2009) excluding the first 6 bases of the first read.

Influence of adapter length in pooled hybrid capture

An earlier version of the above presented protocol was used to prepare barcoded libraries of 50 human DNA samples derived from cell culture (HapMap- and CEPH-HGDP-samples). After the nick fill-in reaction, subsamples of each of the 50 samples were used for two different amplifications with different primer pairs resulting in different length of the adapters. First, enrichment PCR with primer pair Sol-PE-PCR resulted in 50 finished (sequence-able) libraries with full-length adapters on either side (= 'regular' [long] adapter including 6base barcode: 64bp and 61bp). Second, amplification with primer pair PreHyb-PE resulted in 50 libraries with short adapters (= 'truncated' adapter including 6 base barcode: 34bp and 33bp). The number of barcoded molecules per library (total 100) was measured with qPCR assays and the 50 libraries were pooled in equimolar ratio per library-type (regular and truncated). After an amplification step of these two pools with the respective primer pair to achieve a sufficient input amount of library, one solution hybrid selection was performed per pool with a custom Agilent SureSelect Target Enrichment System Kit (target size ~ 6Mb). Hybridization for the regular-length library-pool was performed according to the instructions using the provided components. The blocking oligos (Block #3 PE) were replaced with Univ_Block (see above – Pooled Solution Hybrid Selection) for the truncated library-pool. Enrichment PCR after finishing the hybrid capture was done with the Sol-PE-qPCR primer pair for the regular-length library-pool and with the Sol-PE-PCR primer pair for the truncated library-pool. One lane SR36 cycle Illumina sequencing was performed for each experiment. Analysis was done as explained above for the hybrid selection experiment (**application 1**). Results are shown in Table S1.

SUPPLEMENTARY FIGURES

Figure S1

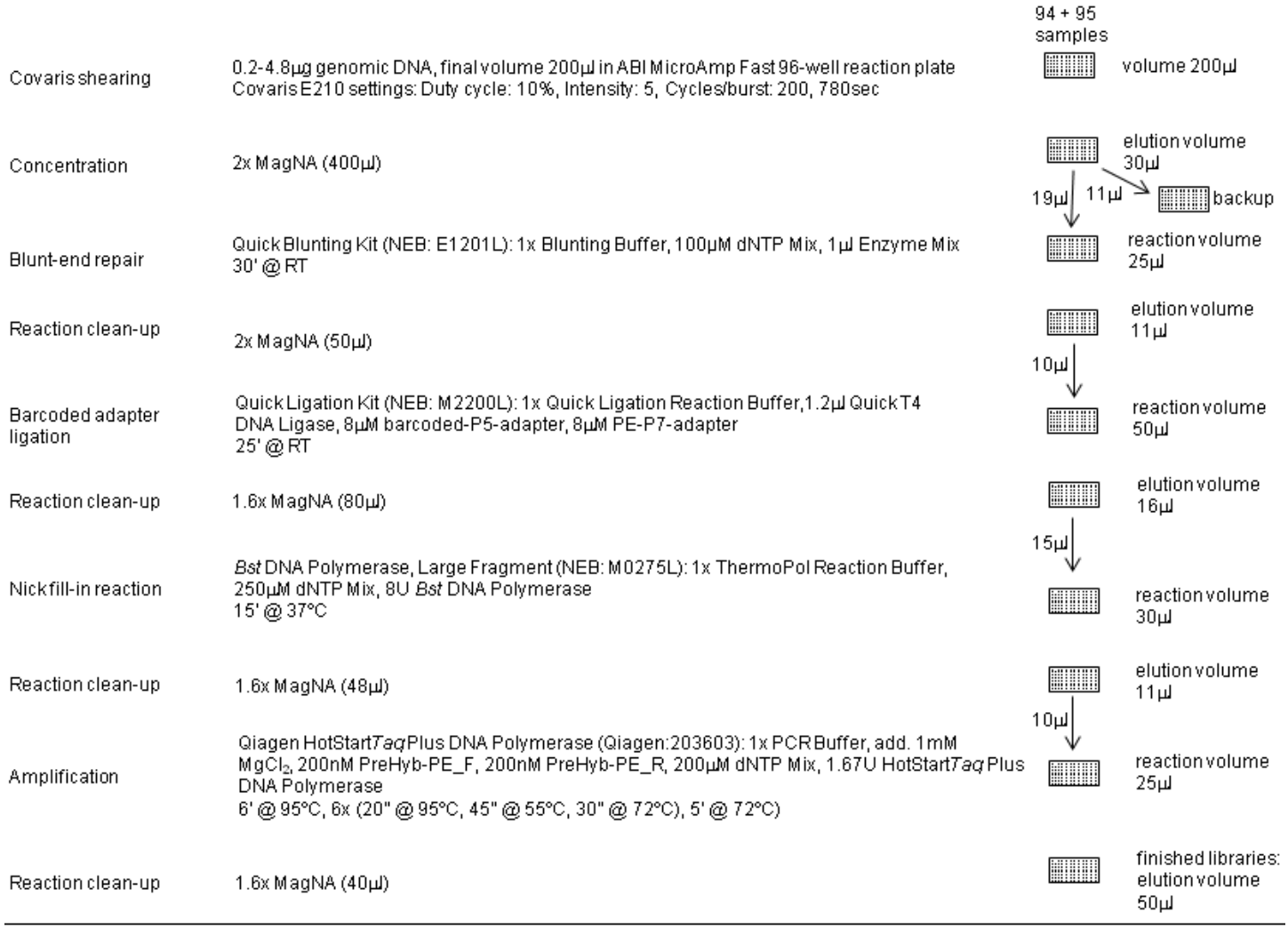

Figure S1: Experimental overview and specific reaction conditions for application 1 – pooled hybrid selection in 4 differently sized pools (14, 28, 52 and 95 libraries per hybrid selection reaction).

Figure S2

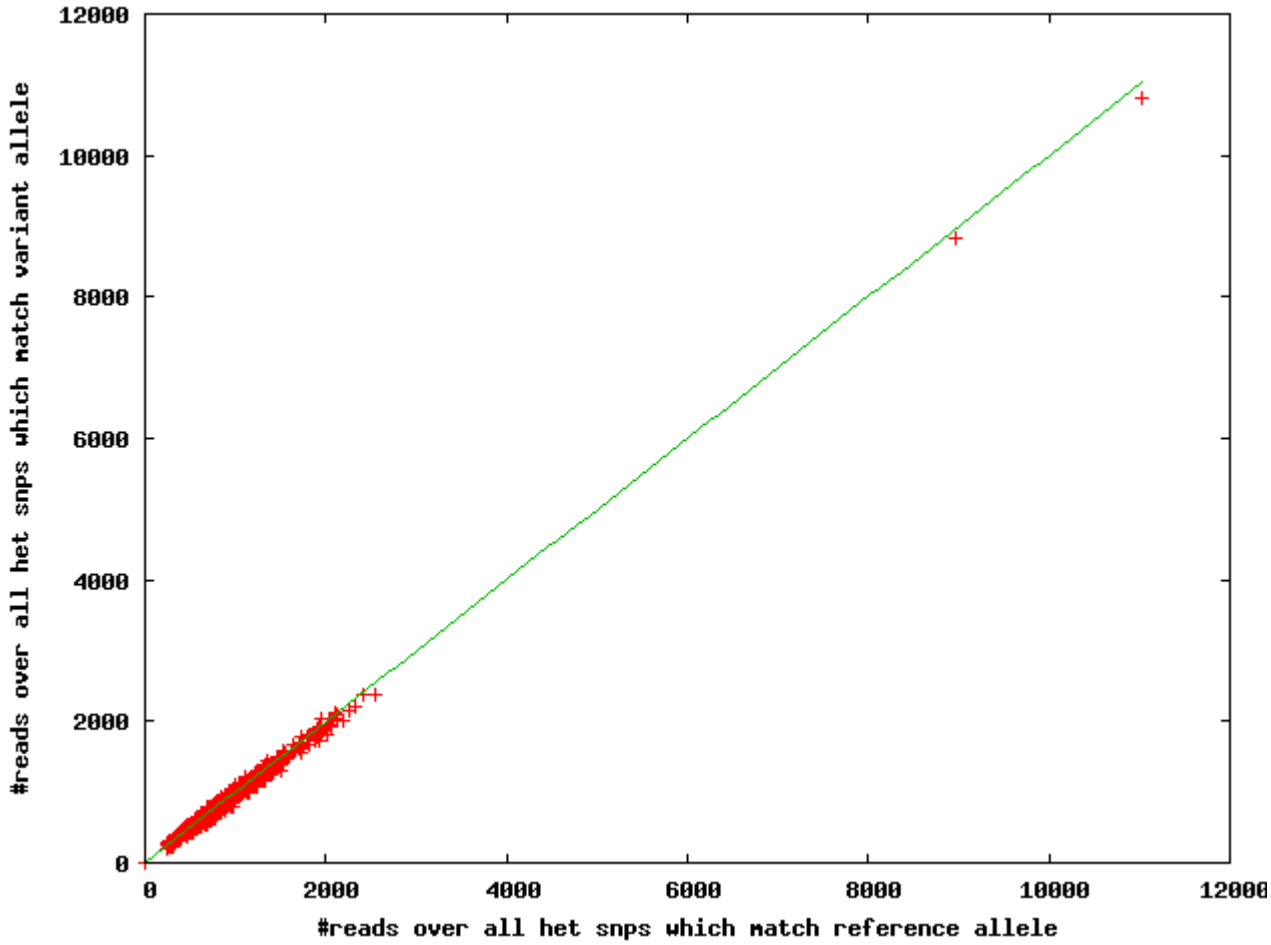

Figure S2: Number of reads matching the reference (x-axis) and variant allele (y-axis) from known heterozygous positions in overlapping data set of the gold standard and our targeted re-sequencing per sample.

Figure S3

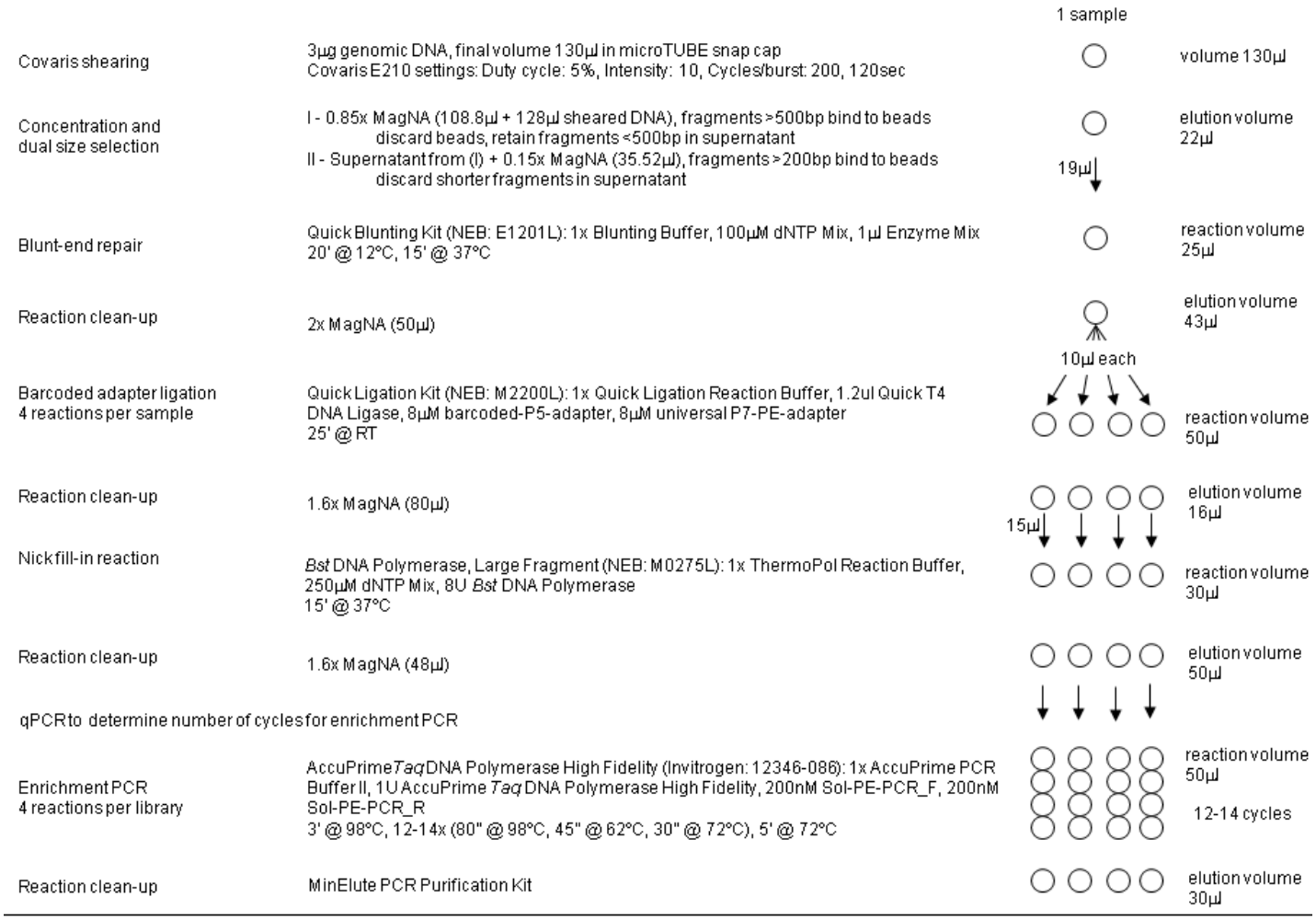

Figure S3: Experimental overview (shown for 1 sample) and specific reaction conditions for application 2 (10 human WGS sequencing)

Figure S4

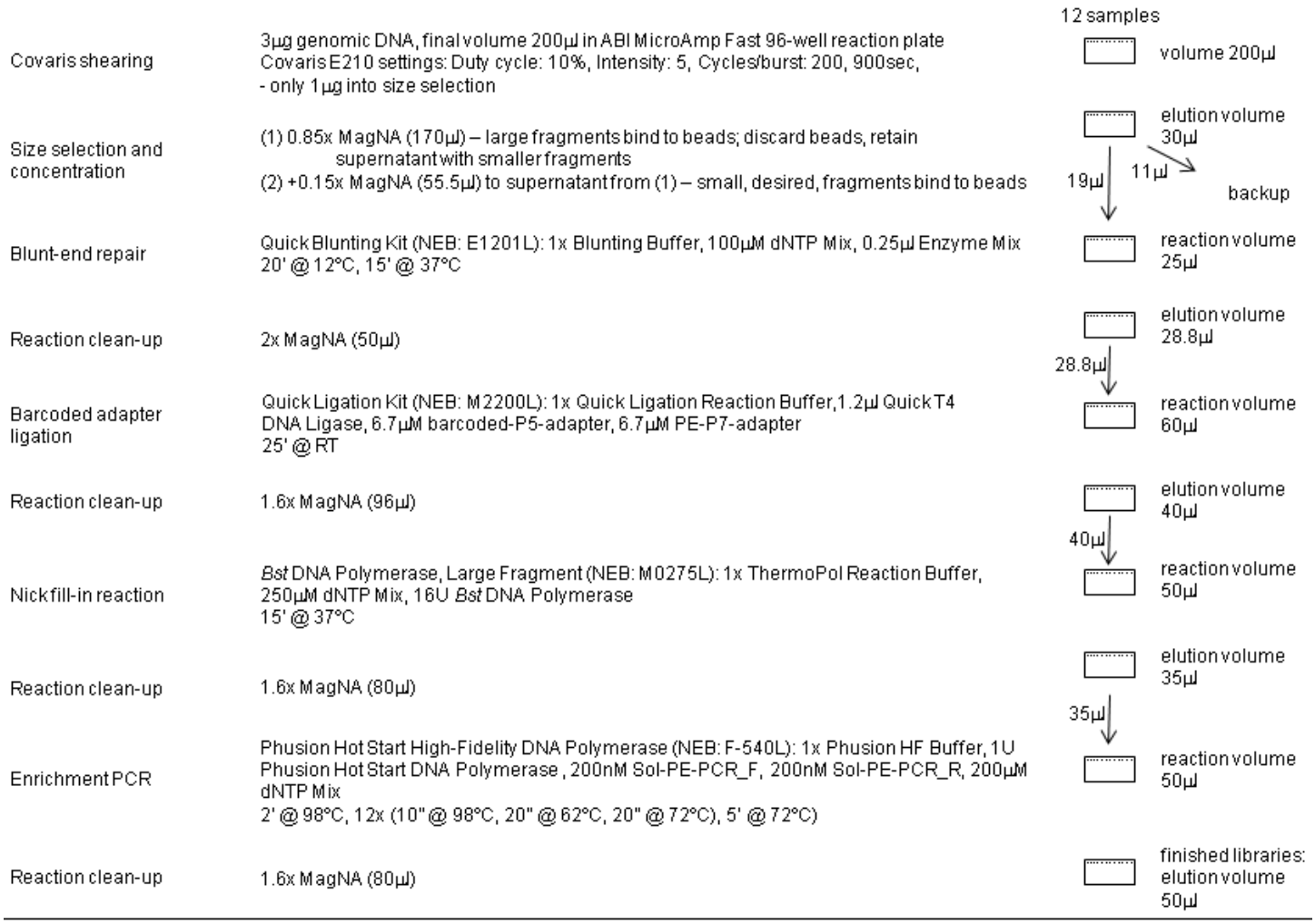

Figure S4: Experimental overview and specific reaction conditions for application 3 (WGS of 12 bacterial strains)

Figure S5

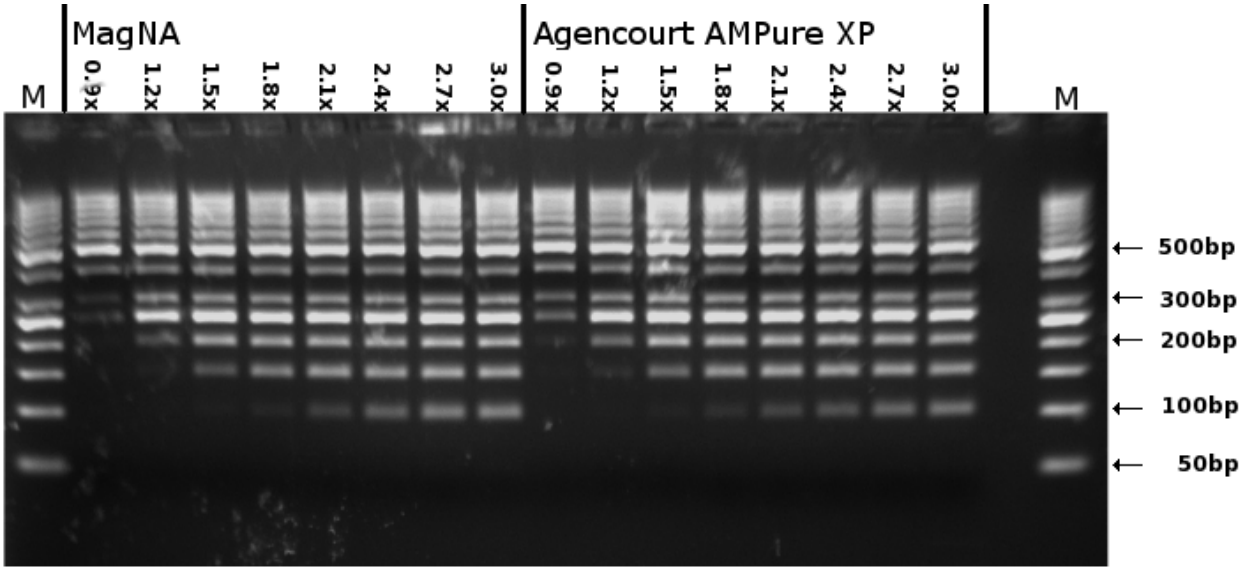

Figure S5: Gel picture after purification of 1µg DNA marker (GeneRuler 50bp ladder, Fermentas) using Agencourt AMPure XP (Beckman Coulter) and MagNA with varying ratios of bead-suspension to DNA marker from 0.9x to 3.0x. M: original size marker 1µg

SUPPLEMENTARY TABLES

Table S1: Influence of adapter length on capture efficiency

	'regular' (long) adapter					'truncated' (short) adapter				
barcode	PF unique bases aligned	on-bait bases	near-bait bases (250bp either side)	off-bait bases	% selected bases (on and near bait)	PF unique bases aligned	on-bait bases	near-bait bases (250bp either side)	off-bait bases	% selected bases (on and near bait)
AAGTTG	9,227,222	1,326,981	1,634,225	6,266,016	0.32	3,363,545	952,619	1,561,774	849,152	0.75
AATACA	8,797,458	1,598,157	1,763,191	5,436,110	0.38	3,786,937	1,165,622	1,703,385	917,930	0.76
AATTGA	6,917,620	1,207,812	1,387,396	4,322,412	0.38	2,846,952	825,619	1,235,278	786,055	0.72
ACAACC	10,517,399	1,256,333	1,521,478	7,739,588	0.26	3,005,346	821,591	1,347,725	836,030	0.72
ACAATA	8,124,432	1,440,466	1,655,470	5,028,496	0.38	4,442,280	1,255,907	1,941,018	1,245,355	0.72
ACCGCG	8,368,814	771,214	943,929	6,653,671	0.20	2,077,689	549,712	859,721	668,256	0.68
ACTCAC	8,520,728	1,104,393	1,346,438	6,069,897	0.29	3,195,273	909,328	1,412,783	873,162	0.73
ACTGAT	9,464,527	1,435,200	1,656,556	6,372,771	0.33	3,462,630	985,360	1,509,097	968,173	0.72
AGGCGA	15,140,020	1,658,930	1,908,080	11,573,010	0.24	4,300,345	1,230,863	1,896,650	1,172,832	0.73
AGGTTC	12,196,505	1,569,239	1,800,705	8,826,561	0.28	3,269,281	970,145	1,451,647	847,489	0.74
AGTGGC	12,171,931	1,361,410	1,571,651	9,238,870	0.24	2,729,344	805,692	1,250,206	673,446	0.75
ATACAC	8,157,394	1,311,759	1,542,701	5,302,934	0.35	3,732,233	1,085,228	1,727,401	919,604	0.75
ATGTAG	6,892,295	1,151,931	1,275,595	4,464,769	0.35	2,946,123	887,196	1,291,537	767,390	0.74
ATGTGT	7,709,627	1,162,830	1,379,003	5,167,794	0.33	3,472,836	999,963	1,514,308	958,565	0.72
CACCAC	13,710,606	1,532,015	1,712,979	10,465,612	0.24	3,034,310	882,758	1,377,655	773,897	0.74
CACTGC	9,726,089	1,064,314	1,203,859	7,457,916	0.23	3,287,709	945,772	1,485,567	856,370	0.74
CACTTA	7,066,957	1,048,692	1,206,713	4,811,552	0.32	3,772,493	1,116,034	1,680,635	975,824	0.74
CCATAG	9,436,188	1,262,930	1,528,685	6,644,573	0.30	2,751,456	777,047	1,219,075	755,334	0.73
CCGAAT	12,104,329	1,401,687	1,684,740	9,017,902	0.25	3,010,104	891,241	1,337,361	781,502	0.74
CCGCGT	9,930,491	909,232	1,263,585	7,757,674	0.22	2,974,992	698,712	1,412,592	863,688	0.71
CGTTAT	9,426,962	1,291,335	1,518,990	6,616,637	0.30	3,048,729	868,674	1,306,122	873,933	0.71
CTATTG	8,555,268	1,315,472	1,483,533	5,756,263	0.33	3,529,394	1,069,977	1,595,109	864,308	0.76
CTCAGC	7,501,034	845,052	1,059,657	5,596,325	0.25	4,789,966	1,288,950	2,158,000	1,343,016	0.72
CTCGGA	6,317,616	792,415	942,569	4,582,632	0.27	3,088,878	860,628	1,394,497	833,753	0.73
CTCTCG	9,524,400	1,110,287	1,241,288	7,172,825	0.25	3,318,780	1,007,506	1,494,791	816,483	0.75
CTGACA	11,156,332	1,333,381	1,601,858	8,221,093	0.26	3,008,422	851,128	1,338,185	819,109	0.73
CTTAGA	6,223,795	951,983	1,343,537	3,928,275	0.37	4,982,342	1,134,535	2,271,210	1,576,597	0.68
CTTATC	7,580,426	1,191,686	1,484,045	4,904,695	0.35	3,841,506	1,113,089	1,695,964	1,032,453	0.73
GAACTA	7,640,318	1,031,968	1,390,772	5,217,578	0.32	3,365,111	847,625	1,592,380	925,106	0.73
GAGGCA	15,507,338	1,644,618	2,055,229	11,807,491	0.24	2,497,469	708,018	1,150,141	639,310	0.74
GATCGG	10,447,741	1,169,484	1,428,993	7,849,264	0.25	3,982,205	1,115,726	1,800,896	1,065,583	0.73
GCAGCT	11,861,279	1,168,519	1,467,390	9,225,370	0.22	2,692,978	683,606	1,134,478	874,894	0.68
GCCTTG	14,046,507	1,478,832	1,677,928	10,889,747	0.22	3,128,055	903,485	1,354,962	869,608	0.72
GCTAAG	11,991,328	1,458,820	1,702,408	8,830,100	0.26	3,011,711	890,980	1,359,085	761,646	0.75
GGTAAC	11,460,700	1,322,776	1,832,368	8,305,556	0.28	3,958,253	993,542	1,792,737	1,171,974	0.70
GGTACG	13,998,576	1,552,355	1,777,336	10,668,885	0.24	3,023,112	897,733	1,363,529	761,850	0.75
GTAGCC	11,782,112	1,278,818	1,511,047	8,992,247	0.24	2,411,385	663,984	1,032,631	714,770	0.70
GTGGTC	10,253,868	1,113,770	1,343,058	7,797,040	0.24	2,605,654	691,920	1,098,065	815,669	0.69
GTGTCC	11,159,947	1,192,312	1,383,100	8,584,535	0.23	3,301,546	956,172	1,487,676	857,698	0.74
TAAGGT	10,487,663	1,643,700	1,828,407	7,015,556	0.33	2,907,242	871,951	1,317,803	717,488	0.75
TACGAT	11,775,620	1,659,130	1,961,821	8,154,669	0.31	3,416,526	993,065	1,585,118	838,343	0.75
TAGAGT	6,859,181	1,143,300	1,356,985	4,358,896	0.36	2,920,089	807,183	1,333,031	779,875	0.73
TAGCAG	10,332,013	1,343,058	1,525,141	7,463,814	0.28	2,311,778	696,758	1,034,345	580,675	0.75
TCCGGT	8,383,744	994,543	1,239,914	6,149,287	0.27	4,638,057	1,323,110	2,208,941	1,106,006	0.76
TCCTCC	9,437,635	1,039,356	1,358,008	7,040,271	0.25	3,262,089	839,463	1,403,452	1,019,174	0.69
TGACCT	14,292,082	1,624,789	2,185,314	10,481,979	0.27	5,263,076	1,356,622	2,473,917	1,432,537	0.73
TGCAGT	6,332,917	774,388	933,381	4,625,148	0.27	3,192,818	901,399	1,424,370	867,049	0.73
TGGACT	9,597,715	1,176,618	1,368,312	7,052,785	0.27	3,196,681	933,668	1,460,190	802,823	0.75
TGTGTG	9,735,206	1,125,087	1,439,172	7,170,947	0.26	3,145,916	835,412	1,420,180	890,324	0.72
TTCATA	3,726,781	735,041	824,769	2,166,971	0.42	3,683,343	1,094,144	1,722,788	866,411	0.76
average	9,831,535	1,241,568	1,485,066	7,104,900	**0.29**	3,339,700	939,129	1,500,400	900,170	**0.73**
stdev	2,534,506	250,614	287,116	2,152,852	**0.05**	667,578	179,131	321,612	194,727	**0.02**

Table S2a: Pooled solution hybrid selections statistics of application 1 (14 samples in pool, normalization via dilution)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	PF reads	PF unique reads	% duplication	PF unique reads aligned	On bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
621	197	ACGGTG	238,423	118,643	0.56	87,611	1,767,280	482,470	728,922	0.59	0.76	0.81	845	0.03	0.21	0.00
712	225	TCACAA	139,852	85,960	0.46	65,294	1,493,760	317,076	409,064	0.67	0.82	0.69	958	0.03	0.17	0.00
785	249	CTCAGC	175,196	100,177	0.50	74,939	1,686,424	361,498	499,898	0.66	0.80	0.78	942	0.03	0.20	0.00
790	250	CAGGTT	187,909	96,044	0.55	70,183	1,523,755	342,117	520,283	0.64	0.78	0.70	909	0.03	0.17	0.00
795	252	GAATCG	306,676	146,316	0.60	102,050	1,909,604	579,361	980,549	0.55	0.72	0.88	784	0.03	0.23	0.00
800	253	TAGATC	198,807	114,075	0.49	86,228	1,939,487	413,480	578,674	0.66	0.80	0.89	942	0.03	0.24	0.00
817	259	GGAATT	1,382,638	486,734	0.72	307,346	5,797,352	1,283,303	3,368,600	0.55	0.68	2.67	790	0.01	0.65	0.00
829	262	CATGAC	245,263	123,349	0.56	89,175	1,821,798	481,449	728,582	0.60	0.76	0.84	856	0.03	0.22	0.00
835	264	GCTATT	123,318	69,923	0.53	49,367	1,121,329	217,138	339,948	0.67	0.80	0.52	951	0.04	0.11	0.00
852	270	AGCGCA	180,976	95,096	0.54	69,132	1,446,609	357,137	546,665	0.62	0.77	0.67	876	0.03	0.16	0.00
882	279	TCTTCT	151,425	89,681	0.48	66,759	1,488,753	331,402	449,549	0.66	0.80	0.68	934	0.03	0.17	0.00
885	280	TCCGGT	167,565	104,778	0.44	81,037	1,780,726	438,369	536,050	0.65	0.81	0.82	920	0.03	0.22	0.00
891	282	CATCAA	279,215	142,948	0.55	104,160	2,262,920	499,385	778,996	0.64	0.78	1.04	910	0.02	0.29	0.00
897	284	ACATCA	193,266	105,886	0.52	79,051	1,738,588	383,590	565,455	0.65	0.79	0.80	921	0.03	0.21	0.00
814	258	average	283,609	134,258	0.54	95,167	1,984,170	463,413	787,945	**0.63**	**0.78**	0.91	896	0.03	0.23	0.00
75	24	stdev	320,579	103,596	0.07	62,812	1,130,717	253,138	761,189	**0.04**	**0.04**	0.52	57	0.01	0.13	0.00

Table S2b: Pooled solution hybrid selections statistics of application 1 (28 samples in pool, normalization via dilution)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	PF reads	PF unique reads	% duplication	PF unique reads aligned	On bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
166	53	GAGGCA	568,862	173,459	0.76	106,225	1,508,254	488,353	1,614,848	0.42	0.55	0.69	595	0.03	0.17	0.00
426	135	CTTATC	409,074	191,521	0.60	136,384	2,633,934	770,769	1,232,174	0.57	0.73	1.21	809	0.02	0.34	0.00
651	206	GTGGTC	304,578	138,020	0.60	101,008	2,070,078	507,184	856,886	0.60	0.75	0.95	858	0.03	0.26	0.00
690	219	ACCGCG	493,225	186,180	0.69	121,718	2,084,487	627,162	1,426,579	0.50	0.66	0.96	717	0.02	0.26	0.00
720	228	TGTGTG	272,008	157,079	0.48	122,319	2,866,122	545,732	746,839	0.69	0.82	1.32	981	0.02	0.37	0.00
755	239	ACAACC	278,359	154,378	0.52	112,430	2,515,050	529,990	777,389	0.66	0.80	1.16	937	0.02	0.33	0.00
760	241	CTGACA	306,018	148,221	0.58	107,215	2,194,962	570,820	879,386	0.60	0.76	1.01	857	0.03	0.28	0.00
773	245	GAACTA	255,763	139,259	0.52	103,424	2,313,547	493,534	709,194	0.66	0.80	1.06	937	0.02	0.30	0.00
809	256	CTCGGA	446,177	197,026	0.62	139,507	2,898,392	662,201	1,182,457	0.61	0.75	1.33	870	0.02	0.38	0.00
823	261	GCAGCT	350,470	175,825	0.57	123,809	2,470,467	679,148	1,059,714	0.59	0.75	1.14	836	0.02	0.32	0.00
825	261	CCGCGT	376,940	184,996	0.58	132,409	2,661,727	707,821	1,132,160	0.59	0.75	1.22	842	0.02	0.35	0.00
832	263	GGTAAC	216,458	121,622	0.50	91,594	2,048,560	447,069	618,456	0.66	0.80	0.94	937	0.03	0.26	0.00
849	269	CGTTAT	358,559	178,532	0.58	127,837	2,751,708	609,898	984,682	0.63	0.77	1.27	901	0.02	0.36	0.00
850	269	AATTGA	254,727	137,542	0.53	101,200	2,224,228	469,408	747,032	0.65	0.78	1.02	920	0.02	0.28	0.00
856	271	CCATAG	396,505	183,887	0.60	131,563	2,815,596	604,843	1,052,512	0.63	0.76	1.29	896	0.02	0.37	0.00
857	271	CTTAGA	223,263	118,408	0.55	87,174	1,931,878	402,622	629,294	0.65	0.79	0.89	928	0.03	0.24	0.00
865	274	TCACGG *	453,417	225,426	0.58	155,955	2,984,753	877,856	1,439,635	0.56	0.73	1.37	802	0.02	0.39	0.00
870	276	CCGAAT	271,363	140,953	0.55	103,784	2,271,639	499,460	757,403	0.64	0.79	1.04	917	0.02	0.29	0.00
872	276	ATACAC	556,937	265,355	0.59	187,798	3,492,518	1,074,832	1,817,506	0.55	0.72	1.61	779	0.02	0.46	0.00
877	278	AGCCGC	209,837	127,455	0.47	95,738	2,173,288	473,724	607,951	0.67	0.81	1.00	951	0.03	0.28	0.00
882	279	TTCATA	369,542	182,504	0.57	132,256	2,873,625	609,393	1,013,475	0.64	0.77	1.32	910	0.02	0.38	0.00
891	282	ATTACT	279,706	151,087	0.53	111,050	2,468,702	518,470	788,375	0.65	0.79	1.14	931	0.02	0.32	0.00
892	282	TGCAGT	220,524	116,576	0.54	87,690	1,984,929	395,704	600,714	0.67	0.80	0.91	948	0.03	0.25	0.00
897	284	TGACCT	253,253	137,833	0.52	103,169	2,302,056	484,514	721,021	0.66	0.79	1.06	934	0.03	0.30	0.00
900	285	TCCTCC	266,015	145,708	0.53	105,904	2,282,890	531,916	785,785	0.63	0.78	1.05	903	0.02	0.29	0.00
908	288	GTAGCC	293,362	160,515	0.52	119,269	2,621,107	591,563	842,287	0.65	0.79	1.21	920	0.02	0.34	0.00
915	290	GATCGG	333,172	163,332	0.59	115,253	2,436,069	550,717	931,637	0.62	0.76	1.12	885	0.02	0.31	0.00
916	290	TAGCGC	234,662	130,552	0.51	97,328	2,134,549	481,215	693,252	0.65	0.79	0.98	918	0.03	0.27	0.00
797	252	average	330,456	161,902	0.56	116,465	2,429,111	578,783	951,737	**0.62**	**0.76**	1.12	879	0.02	0.31	0.00
161	51	stdev	100,945	33,646	0.06	21,933	406,208	145,483	316,938	**0.06**	**0.05**	0.19	82	0.00	0.06	0.00

* Sequence data from samples marked with an asterisk were deleted from the fastq file before deposition in NCBI's SRA.

Table S2c: Pooled solution hybrid selections statistics of application 1 (52 samples in pool, normalization via dilution)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	PF reads	PF unique reads	% duplication	PF unique reads aligned	On bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
263	83	CCTCCA	530,978	183,656	0.72	121,408	2,058,462	608,956	1,460,293	0.50	0.65	0.95	710	0.02	0.26	0.00
534	169	GGCATC	175,284	100,048	0.50	73,804	1,675,521	348,611	485,077	0.67	0.81	0.77	951	0.03	0.20	0.00
571	181	TAAGGT	289,856	143,205	0.57	104,893	2,156,126	549,050	861,031	0.60	0.76	0.99	861	0.02	0.27	0.00
637	202	TACTCT	239,748	126,705	0.55	92,249	2,026,799	432,091	677,444	0.65	0.78	0.93	920	0.03	0.25	0.00
661	209	GCCTTG	296,719	158,400	0.54	114,521	2,326,323	640,694	926,550	0.60	0.76	1.07	851	0.02	0.30	0.00
669	212	GAAGTC	253,290	133,659	0.53	100,587	2,187,993	495,831	736,030	0.64	0.78	1.01	911	0.02	0.28	0.00
682	216	GCGCTA	326,981	179,100	0.53	130,729	2,691,972	730,997	1,021,644	0.61	0.77	1.24	862	0.02	0.35	0.00
691	219	CACTGC	289,619	156,944	0.52	117,357	2,357,274	682,713	949,965	0.59	0.76	1.08	841	0.03	0.30	0.00
721	228	CTGCTG	232,636	129,347	0.51	98,144	2,186,438	484,294	666,016	0.66	0.80	1.01	933	0.03	0.28	0.00
732	232	GCATTC	181,792	94,293	0.55	69,214	1,549,171	319,857	484,162	0.66	0.79	0.71	937	0.03	0.18	0.00
737	233	CTATGA	313,903	166,083	0.54	122,309	2,539,747	660,808	957,794	0.61	0.77	1.17	870	0.02	0.33	0.00
742	235	ATTCTT	231,746	141,173	0.46	108,395	2,508,768	504,600	671,941	0.68	0.82	1.15	969	0.02	0.32	0.00
756	239	CTCTCG	407,913	201,754	0.57	147,038	3,036,181	768,140	1,194,756	0.61	0.76	1.40	865	0.02	0.40	0.00
765	242	TGGACT	175,156	105,612	0.47	80,207	1,853,745	374,629	498,545	0.68	0.82	0.85	968	0.03	0.23	0.00
772	244	TCGGAC	241,909	155,074	0.42	120,992	2,714,695	644,796	754,071	0.66	0.82	1.25	940	0.02	0.35	0.00
787	249	TATCCT	256,335	147,043	0.50	109,755	2,363,238	588,792	779,503	0.63	0.79	1.09	902	0.02	0.30	0.00
793	251	ACTCAC	384,824	190,047	0.57	137,419	2,911,566	677,308	1,083,178	0.62	0.77	1.34	887	0.02	0.38	0.00
794	251	TTGGCG	255,354	138,027	0.53	99,875	1,986,205	571,908	837,512	0.58	0.75	0.91	833	0.03	0.25	0.00
803	254	CACCAC	306,875	163,901	0.53	120,251	2,487,823	663,935	936,614	0.61	0.77	1.14	866	0.02	0.32	0.00
805	255	GCTAAG	314,855	162,675	0.56	114,727	2,200,068	673,986	1,026,497	0.56	0.74	1.01	803	0.02	0.28	0.00
805	255	AAGTTG	346,332	167,102	0.57	124,331	2,527,180	662,892	1,037,014	0.60	0.75	1.16	851	0.02	0.33	0.00
807	256	GACAGA	319,343	162,942	0.56	117,528	2,298,217	671,456	1,026,127	0.58	0.74	1.06	819	0.02	0.29	0.00
809	256	ACCACA	311,852	173,013	0.51	130,354	2,871,332	643,478	917,067	0.65	0.79	1.32	922	0.02	0.37	0.00
812	257	GTGCCT	170,107	106,228	0.44	81,622	1,884,356	400,320	490,356	0.68	0.82	0.87	967	0.03	0.23	0.00
816	258	CGACTC	278,664	138,048	0.58	98,628	2,044,526	492,624	816,092	0.61	0.76	0.94	868	0.03	0.25	0.00
821	260	CACTTA	253,684	139,626	0.52	102,729	2,294,311	485,822	712,497	0.66	0.80	1.05	935	0.02	0.29	0.00
827	262	ACAATA	227,196	140,453	0.45	106,177	2,312,607	575,523	721,738	0.64	0.80	1.06	912	0.03	0.30	0.00
827	262	TTCCAT	305,038	168,378	0.51	125,134	2,726,022	645,887	882,481	0.64	0.79	1.25	912	0.02	0.35	0.00
836	265	ATGTAG *	235,614	131,984	0.50	100,402	2,308,383	438,263	666,897	0.68	0.80	1.06	963	0.02	0.29	0.00
839	266	ATGTGT	372,413	176,785	0.59	126,607	2,520,232	668,430	1,115,777	0.59	0.74	1.16	834	0.02	0.33	0.00
842	267	GTGTCC	281,338	173,159	0.45	133,801	2,936,520	725,084	887,448	0.65	0.80	1.35	919	0.02	0.38	0.00
847	268	AGAAGG	177,742	108,689	0.47	82,158	1,880,676	388,324	524,261	0.67	0.81	0.86	959	0.03	0.23	0.00
848	268	AACCTT	214,782	131,386	0.46	99,254	2,255,095	475,530	643,857	0.67	0.81	1.04	952	0.03	0.29	0.00
849	269	CTATTG	227,063	117,369	0.55	86,854	1,926,008	392,498	634,412	0.65	0.79	0.89	929	0.03	0.24	0.00
856	271	CTAAGT	267,015	136,034	0.55	101,044	2,205,490	482,904	746,973	0.64	0.78	1.01	914	0.02	0.28	0.00
857	271	TAGAGT	228,996	136,790	0.47	104,521	2,277,566	556,243	719,778	0.64	0.80	1.05	913	0.02	0.29	0.00
862	273	TGAGTA	253,594	133,757	0.53	101,431	2,221,157	500,130	727,228	0.64	0.79	1.02	917	0.02	0.28	0.00
866	274	AATACA	323,467	168,585	0.55	120,673	2,227,214	711,606	1,163,891	0.54	0.72	1.02	773	0.02	0.28	0.00
868	275	ACTGAT	188,915	119,272	0.44	90,428	2,042,577	461,423	570,461	0.66	0.81	0.94	946	0.03	0.26	0.00
874	277	GGTACG	323,868	177,721	0.52	131,231	2,600,111	776,987	1,084,590	0.58	0.76	1.20	830	0.02	0.34	0.00
875	277	TGTTCA	224,202	135,998	0.46	105,180	2,433,441	493,298	649,269	0.68	0.82	1.12	969	0.02	0.31	0.00
877	278	GGTCGT	151,663	102,888	0.39	81,216	1,919,984	385,006	456,250	0.70	0.83	0.88	990	0.03	0.24	0.00
885	280	TACGAT	286,714	162,832	0.50	121,951	2,497,331	697,964	950,845	0.60	0.77	1.15	858	0.02	0.32	0.00
886	281	AGGCGA	192,953	125,672	0.41	98,785	2,322,053	474,765	561,755	0.69	0.83	1.07	984	0.03	0.30	0.00
887	281	AGTGGC	352,705	250,206	0.50	147,392	3,011,071	847,735	1,152,252	0.60	0.77	1.38	856	0.02	0.39	0.00
894	283	CCTGTA	308,334	184,101	0.46	141,112	3,032,634	786,842	978,154	0.63	0.80	1.39	900	0.02	0.39	0.00
895	283	TAGCAG	261,406	138,675	0.54	102,943	2,257,587	490,213	752,152	0.65	0.79	1.04	918	0.02	0.29	0.00
901	285	GCTGCC	258,851	157,016	0.48	114,994	2,487,255	623,247	799,130	0.64	0.80	1.14	906	0.02	0.32	0.00
904	286	CCGACG	236,694	113,314	0.59	81,353	1,737,421	373,831	654,628	0.63	0.76	0.80	894	0.03	0.21	0.00
912	289	AGGTTC	343,573	174,962	0.55	131,425	2,588,299	752,691	1,127,267	0.58	0.75	1.19	825	0.02	0.34	0.00
916	290	GGCGAC	136,487	92,139	0.40	71,572	1,692,749	338,860	401,756	0.70	0.83	0.78	990	0.03	0.20	0.00
918	291	GGTTGG	154,335	99,972	0.43	75,124	1,700,464	363,500	490,143	0.67	0.81	0.78	948	0.03	0.20	0.00
797	252	average	267,706	146,574	0.51	108,112	2,304,961	561,642	809,061	**0.63**	**0.78**	1.06	899	0.03	0.29	0.00
115	36	stdev	73,098	30,944	0.06	20,021	369,956	137,423	231,525	**0.04**	**0.03**	0.17	59	0.00	0.05	0.00

* Sequence data from samples marked with an asterisk were deleted from the fastq file before deposition in NCBI's SRA.

Table S2d: Pooled solution hybrid selection statistics of application 1 (95 samples in pool, no normalization)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	insert size	PF reads	PF unique reads	% duplication	PF unique reads aligned	On bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
209	132	TTCCAT	322	263,340	160,621	0.40	142,688	3,851,127	1,075,911	1,728,645	0.58	0.74	1.77	824	0.02	0.48	0.00
496	314	GTGTCC	351	564,586	366,932	0.36	322,936	8,667,197	2,475,080	3,914,511	0.58	0.74	3.99	820	0.01	0.74	0.06
600	380	TGTGTG	343	793,236	503,126	0.38	445,345	11,958,711	3,328,016	5,479,577	0.58	0.74	5.50	820	0.01	0.80	0.17
622	394	CCGAAT	389	488,286	313,122	0.38	273,595	7,108,888	2,147,431	3,481,047	0.56	0.73	3.27	795	0.02	0.68	0.02
683	433	CTGCTG	357	1,225,194	772,913	0.38	684,283	18,159,556	5,332,066	8,420,543	0.57	0.74	8.35	810	0.01	0.84	0.42
688	436	GATCGG	402	1,325,460	831,377	0.39	723,128	18,206,053	5,805,728	9,685,663	0.54	0.71	8.37	769	0.01	0.85	0.42
699	443	TCCGGT	355	1,256,634	796,869	0.38	706,104	18,722,793	5,595,942	8,605,690	0.57	0.74	8.61	810	0.01	0.84	0.43
703	445	GGTAAC	329	665,298	419,989	0.38	372,688	10,131,895	2,784,567	4,465,453	0.58	0.74	4.66	830	0.01	0.76	0.10
707	448	ATTCTT	387	1,916,702	1,189,259	0.40	1,038,481	26,234,334	8,332,791	13,783,801	0.54	0.71	12.06	773	0.01	0.88	0.60
707	448	CTTATC	293	788,292	494,853	0.39	439,457	12,452,350	3,130,481	4,906,955	0.61	0.76	5.73	865	0.01	0.80	0.20
717	454	GACAGA	308	840,134	541,108	0.37	482,521	13,452,585	3,470,489	5,595,926	0.60	0.75	6.19	851	0.01	0.81	0.24
724	459	CATGAC	310	297,578	188,796	0.38	166,801	4,650,012	1,223,166	1,901,442	0.60	0.76	2.14	852	0.02	0.55	0.00
728	461	CACTGC	314	864,840	560,209	0.36	501,308	14,006,115	3,694,114	5,696,676	0.60	0.76	6.44	852	0.01	0.81	0.26
740	469	GGCGAC	322	637,390	415,915	0.36	368,951	10,309,630	2,757,928	4,144,223	0.60	0.76	4.74	853	0.01	0.76	0.12
742	470	ATGTGT	305	948,874	595,999	0.38	529,742	14,708,234	3,877,045	6,110,277	0.60	0.75	6.76	848	0.01	0.82	0.29
743	470	TACGAT	292	1,115,254	684,029	0.40	604,723	16,993,301	4,288,198	6,941,138	0.60	0.75	7.81	857	0.01	0.84	0.37
744	471	CCGCGT	312	813,690	536,469	0.35	478,217	13,528,188	3,495,756	5,283,089	0.61	0.76	6.22	864	0.01	0.80	0.25
750	475	CTATTG	310	924,548	564,526	0.40	503,267	13,814,616	3,718,486	5,959,523	0.59	0.75	6.35	837	0.01	0.81	0.25
759	481	AGGTTC	317	1,466,354	906,336	0.39	801,557	21,659,806	5,932,985	9,797,470	0.58	0.74	9.96	825	0.01	0.87	0.51
761	482	CGACTC	276	358,604	247,771	0.32	221,577	6,517,388	1,437,474	2,383,938	0.63	0.77	3.00	898	0.02	0.66	0.02
762	483	CGTTAT *	317	705,198	447,814	0.38	395,334	10,878,588	2,940,724	4,601,153	0.59	0.75	5.00	841	0.01	0.77	0.13
763	483	AGCCGC	334	1,483,074	941,393	0.38	836,312	22,677,477	6,344,020	10,013,727	0.58	0.74	10.43	827	0.01	0.86	0.53
764	484	AATTGA	317	2,175,534	1,321,185	0.41	1,169,753	31,619,490	8,675,797	14,257,762	0.58	0.74	14.54	825	0.01	0.90	0.67
765	484	GGTCGT	344	820,584	527,295	0.37	462,317	12,510,365	3,593,619	5,459,280	0.58	0.75	5.75	826	0.01	0.79	0.20
772	489	TTGATT	329	918,984	575,857	0.39	510,806	13,938,250	3,798,732	6,083,030	0.59	0.74	6.41	833	0.01	0.82	0.25
779	494	CCTGTA	326	775,634	497,905	0.37	438,933	12,079,925	3,232,989	5,151,771	0.59	0.75	5.55	840	0.01	0.79	0.18
781	494	GAGGCA	316	924,452	577,358	0.39	516,176	14,248,272	3,812,732	6,033,123	0.59	0.75	6.55	842	0.01	0.81	0.27
786	498	ACTCAC	371	468,896	304,830	0.37	267,006	7,010,668	2,096,501	3,327,623	0.56	0.73	3.22	803	0.02	0.68	0.02
787	498	GAAGTC	360	1,232,292	799,305	0.37	707,467	18,801,391	5,531,837	8,661,883	0.57	0.74	8.65	811	0.01	0.85	0.44
787	499	TCTTCT	309	1,264,278	807,615	0.37	718,874	20,114,874	5,289,758	8,119,323	0.60	0.76	9.25	854	0.01	0.85	0.47
796	504	CTATGA	360	893,658	572,238	0.37	505,768	13,279,813	3,956,184	6,325,857	0.56	0.73	6.11	803	0.01	0.81	0.23
797	505	GAATCG	378	1,026,106	645,798	0.38	571,840	14,795,008	4,556,891	7,331,204	0.55	0.73	6.80	789	0.01	0.82	0.29
802	508	GCTGCC	341	871,582	579,703	0.35	513,034	14,082,560	3,896,907	5,940,747	0.59	0.75	6.48	838	0.01	0.80	0.27
803	508	AGGCGA	379	1,420,718	889,382	0.39	780,475	20,104,221	6,203,745	10,077,018	0.55	0.72	9.24	787	0.01	0.85	0.47
804	509	CTCTCG	317	681,062	424,790	0.39	378,393	10,467,539	2,788,346	4,405,835	0.59	0.75	4.81	844	0.01	0.77	0.12
806	510	TAAGGT	341	983,948	612,266	0.39	538,370	14,455,561	4,012,157	6,625,678	0.58	0.74	6.65	820	0.01	0.82	0.28
812	515	CTTAGA	347	892,592	557,051	0.39	483,687	12,856,374	3,663,924	6,008,825	0.57	0.73	5.91	813	0.01	0.81	0.21
813	515	ACAATA	358	1,825,260	1,147,470	0.38	1,017,826	26,786,702	7,895,430	12,786,373	0.56	0.73	12.32	803	0.01	0.88	0.61
817	517	CATCAA	362	662,060	421,488	0.38	369,925	9,729,573	2,944,864	4,550,331	0.56	0.74	4.47	804	0.01	0.76	0.09
820	519	AGTGGC	311	1,071,370	686,470	0.37	612,906	17,062,119	4,501,213	7,045,760	0.60	0.75	7.85	849	0.01	0.83	0.38
821	520	GGAATT	328	57,582	42,574	0.27	38,167	1,096,434	279,465	403,820	0.62	0.77	0.50	877	0.05	0.11	0.00
825	522	CAGGTT	372	668,668	421,887	0.39	368,323	9,597,667	2,959,324	4,586,234	0.56	0.73	4.41	797	0.01	0.75	0.09
827	524	TTGGCG	343	1,397,766	875,947	0.38	781,265	20,920,399	6,051,840	9,521,028	0.57	0.74	9.62	816	0.01	0.86	0.49
828	525	CTCGGA	363	1,203,724	741,416	0.40	648,968	17,017,039	5,006,902	8,228,405	0.56	0.73	7.82	801	0.01	0.84	0.38
829	525	TCACGG	387	1,066,234	694,969	0.36	613,543	15,858,903	4,941,085	7,802,889	0.55	0.73	7.29	789	0.01	0.83	0.34

(continued)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	insert size	PF reads	PF unique reads	% duplication	PF unique reads aligned	On bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
832	527	GTGCCT	372	580,926	381,540	0.36	334,785	8,873,977	2,654,783	4,067,998	0.57	0.74	4.08	810	0.01	0.73	0.07
837	530	TGCAGT	325	1,440,546	912,157	0.38	806,791	22,119,359	6,117,216	9,391,628	0.59	0.75	10.17	837	0.01	0.86	0.52
840	532	ATTACT	327	1,744,858	1,086,083	0.39	969,488	26,128,751	7,372,893	11,700,436	0.58	0.74	12.01	823	0.01	0.88	0.60
841	532	GCGCTA	313	982,922	636,914	0.36	566,753	15,783,551	4,090,559	6,572,002	0.60	0.75	7.26	850	0.01	0.83	0.33
849	538	CACCAC	367	650,540	423,118	0.36	375,364	9,940,638	2,994,230	4,559,094	0.57	0.74	4.57	809	0.01	0.76	0.10
850	538	TAGATC	349	1,168,634	744,951	0.38	662,098	17,626,447	5,190,695	8,057,518	0.57	0.74	8.11	813	0.01	0.85	0.40
852	539	GCAGCT	358	1,113,036	710,656	0.38	630,834	16,718,759	4,947,782	7,765,234	0.57	0.74	7.69	809	0.01	0.83	0.37
852	540	CCGACG	324	676,768	428,669	0.38	380,617	10,520,277	2,869,233	4,387,753	0.59	0.75	4.84	843	0.01	0.76	0.12
854	541	ACTGAT	319	1,296,208	789,565	0.41	701,957	19,096,392	5,259,580	8,371,623	0.58	0.74	8.78	831	0.01	0.85	0.45
855	541	ACCACA	310	1,248,582	796,724	0.37	713,433	19,982,475	5,249,825	8,086,973	0.60	0.76	9.19	854	0.01	0.85	0.47
860	545	CACTTA	323	973,636	611,624	0.39	541,532	14,817,763	4,107,167	6,319,887	0.59	0.75	6.81	836	0.01	0.83	0.29
863	546	CTAAGT	328	1,533,694	955,548	0.39	849,395	23,268,432	6,297,875	10,039,696	0.59	0.75	10.70	837	0.01	0.86	0.54
863	546	TAGCAG	367	1,622,956	1,031,044	0.38	908,576	23,748,957	7,084,827	11,538,914	0.56	0.73	10.92	798	0.01	0.88	0.56
864	547	GGTTGG	403	886,970	575,940	0.37	499,671	12,686,167	4,039,455	6,539,068	0.55	0.72	5.83	776	0.01	0.80	0.20
868	549	CTGACA	343	1,265,822	807,016	0.38	714,952	19,304,320	5,396,707	8,646,938	0.58	0.74	8.88	824	0.01	0.85	0.45
868	549	GAACTA	332	1,067,360	670,323	0.39	591,258	16,091,659	4,421,048	7,051,989	0.58	0.74	7.40	831	0.01	0.84	0.34
868	549	GTAGCC	336	802,180	530,438	0.35	470,947	12,994,336	3,502,774	5,474,058	0.59	0.75	5.98	842	0.01	0.80	0.22
870	551	ACCGCG	348	503,026	315,120	0.39	280,833	7,534,773	2,154,736	3,423,804	0.57	0.74	3.46	818	0.01	0.69	0.03
874	554	TAGCGC	294	768,592	498,396	0.36	445,824	12,765,036	3,144,495	4,906,444	0.61	0.76	5.87	873	0.01	0.80	0.21
875	554	TCGGAC	353	878,890	573,154	0.36	506,698	13,634,587	3,890,852	6,101,065	0.58	0.74	6.27	822	0.01	0.81	0.25
876	555	TACTCT	345	1,004,016	641,178	0.37	571,608	15,210,664	4,421,084	7,025,487	0.57	0.74	6.99	812	0.01	0.83	0.31
876	555	ACAACC	262	89,028	66,224	0.27	59,143	1,827,819	345,284	587,938	0.66	0.79	0.84	943	0.04	0.22	0.00
880	557	ACGGTG	326	1,836,352	1,161,844	0.38	1,039,219	28,587,374	7,876,144	12,067,121	0.59	0.75	13.15	839	0.01	0.88	0.62
880	557	TGACCT	318	991,366	632,822	0.37	566,129	15,600,597	4,238,292	6,576,008	0.59	0.75	7.17	841	0.01	0.83	0.32
882	559	TTCATA	343	1,480,030	934,458	0.38	826,354	22,055,719	6,372,556	10,103,510	0.57	0.74	10.14	815	0.01	0.87	0.52
885	560	CTCAGC	355	946,068	613,988	0.36	543,476	14,623,081	4,282,041	6,430,207	0.58	0.75	6.72	822	0.01	0.81	0.29
886	561	GTGGTC	344	992,858	640,899	0.37	567,099	15,393,978	4,318,121	6,735,882	0.58	0.75	7.08	829	0.01	0.82	0.32
888	562	AAGTTG	329	930,982	569,148	0.40	503,907	13,651,212	3,759,429	6,118,314	0.58	0.74	6.28	826	0.01	0.81	0.25
889	563	TGAGTA	295	1,466,790	900,158	0.40	802,725	22,395,489	5,693,616	9,389,771	0.60	0.75	10.30	851	0.01	0.87	0.53
892	565	GGTACG	356	1,351,480	869,597	0.37	766,423	20,434,382	5,899,495	9,421,497	0.57	0.74	9.40	814	0.01	0.86	0.48
892	565	AACCTT	286	51,656	38,223	0.27	34,225	1,013,810	226,200	355,900	0.64	0.78	0.47	905	0.05	0.09	0.00
893	566	TGGACT	328	1,709,238	1,056,471	0.40	929,717	25,107,938	7,071,044	11,196,934	0.58	0.74	11.55	824	0.01	0.88	0.58
895	567	GGCATC	324	1,165,886	746,853	0.37	662,804	18,293,285	4,999,306	7,640,822	0.59	0.75	8.41	842	0.01	0.84	0.42
897	568	TCCTCC	336	1,079,258	704,285	0.36	620,624	16,952,120	4,622,384	7,345,693	0.59	0.75	7.80	835	0.01	0.84	0.37
897	568	AATACA	333	1,769,734	1,092,905	0.40	966,318	25,918,021	7,172,976	11,989,719	0.57	0.73	11.92	819	0.01	0.88	0.60
898	569	ATTCCG	377	1,850,654	1,180,039	0.38	1,043,215	27,043,059	8,312,040	13,312,224	0.56	0.73	12.44	791	0.01	0.88	0.61
899	569	TATCCT	340	1,298,666	815,276	0.39	722,667	19,577,155	5,504,376	8,594,525	0.58	0.74	9.00	828	0.01	0.85	0.46
903	572	ACATCA	385	1,080,720	696,441	0.37	612,740	15,762,149	4,901,691	7,885,104	0.55	0.72	7.25	786	0.01	0.84	0.33
906	574	GCATTC	407	464,290	310,717	0.35	270,166	6,961,150	2,189,201	3,430,442	0.55	0.73	3.20	788	0.02	0.68	0.02
906	574	GCCTTG	364	1,412,272	903,243	0.37	800,939	21,274,077	6,179,158	9,905,809	0.57	0.73	9.78	811	0.01	0.85	0.50
906	574	GCTAAG	334	975,634	605,017	0.40	533,283	14,342,210	4,025,379	6,499,699	0.58	0.74	6.59	821	0.01	0.82	0.27
908	575	TGTTCA	338	1,110,998	687,085	0.40	604,857	16,171,235	4,568,849	7,459,239	0.57	0.74	7.44	817	0.01	0.84	0.35
912	577	CCTCCA	294	1,913,334	1,176,054	0.40	1,053,767	29,924,710	7,363,486	11,925,795	0.61	0.76	13.76	866	0.01	0.88	0.63
913	578	TAGAGT	324	1,269,466	777,808	0.40	689,955	18,623,866	5,181,668	8,357,609	0.58	0.74	8.56	824	0.01	0.85	0.43
916	580	ATACAC	387	1,435,488	920,916	0.37	810,060	20,715,366	6,542,734	10,494,346	0.55	0.72	9.53	781	0.01	0.86	0.49
917	581	GCTATT *	323	1,520,468	974,305	0.37	867,736	23,945,896	6,508,232	10,015,624	0.59	0.75	11.01	842	0.01	0.86	0.55
918	581	AGAAGG	436	928,272	589,531	0.38	506,936	12,242,200	4,247,515	7,109,185	0.52	0.70	5.63	739	0.01	0.81	0.18
1689	1070	ATGTAG	303	1,214,582	772,066	0.38	685,881	19,064,479	4,940,630	7,973,045	0.60	0.75	8.77	849	0.01	0.85	0.44
2813	1781	CCATAG	326	1,882,744	1,243,204	0.35	1,101,288	30,308,078	8,291,879	12,754,729	0.59	0.75	13.94	840	0.01	0.89	0.65
4833	3061	AGCGCA	387	1,205,544	789,237	0.36	699,078	18,215,722	5,732,153	8,652,967	0.56	0.73	8.38	796	0.01	0.84	0.42
886	561	average	339	1,059,785	671,507	0.37	594,389	16,021,005	4,520,551	7,178,858	**0.58**	**0.74**	7.37	826	0.01	0.79	0.32
477	302	stdev	31	436,622	271,086	0.02	240,451	6,447,782	1,844,917	2,974,261	**0.02**	**0.01**	2.96	29	0.01	0.14	0.19

* Sequence data from samples marked with an asterisk were deleted from the fastq file before deposition in NCBI's SRA.

Table S2e: Pooled solution hybrid selection statistics of application 1 (81 samples in pool, normalization with cherry picking)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	insert size	PF reads	PF unique reads	% duplication	PF unique reads aligned	on bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
574	363	TGGACT	155	532,370	343,815	0.37	314,356	12,085,494	1,299,471	1,308,092	0.82	0.91	5.56	1,171	0.01	0.84	0.16
652	413	GGCATC	169	507,903	395,601	0.23	366,170	14,107,852	1,661,665	1,376,114	0.82	0.92	6.49	1,172	0.01	0.86	0.24
690	437	CTATGA	165	514,947	415,597	0.20	387,442	15,036,431	1,659,022	1,431,609	0.83	0.92	6.91	1,181	0.01	0.87	0.27
705	446	GCATTC	165	526,365	397,762	0.26	366,277	14,088,204	1,610,484	1,426,903	0.82	0.92	6.48	1,171	0.01	0.86	0.24
713	451	GCGCTA	163	503,953	389,572	0.24	357,950	13,874,963	1,535,833	1,341,201	0.83	0.92	6.38	1,179	0.01	0.86	0.23
723	458	TCACAA	140	703,939	530,070	0.26	491,879	19,656,887	1,611,712	1,743,511	0.85	0.92	9.04	1,216	0.00	0.89	0.42
724	458	AACCTT	159	549,662	443,590	0.20	412,649	16,162,695	1,665,979	1,493,981	0.84	0.92	7.43	1,191	0.01	0.87	0.31
752	477	TAGATC	173	645,036	470,878	0.28	435,770	16,553,150	2,083,132	1,749,791	0.81	0.91	7.61	1,156	0.00	0.89	0.33
758	480	TCACGG	171	541,771	411,738	0.25	382,790	14,674,057	1,782,216	1,464,585	0.82	0.92	6.75	1,166	0.01	0.87	0.26
790	500	TACTCT	166	472,522	367,499	0.24	339,954	13,085,375	1,487,870	1,318,101	0.82	0.92	6.02	1,172	0.01	0.85	0.20
790	500	GACAGA	160	492,174	378,316	0.24	351,994	13,678,025	1,479,441	1,321,760	0.83	0.92	6.29	1,182	0.01	0.86	0.23
829	525	TACGAT	156	541,996	422,584	0.24	388,323	15,169,820	1,548,770	1,437,637	0.84	0.92	6.98	1,190	0.01	0.87	0.28
866	549	AAGTTG	152	540,155	427,059	0.22	397,066	15,629,162	1,518,066	1,460,752	0.84	0.92	7.19	1,196	0.01	0.87	0.30
885	560	GCCTTG	154	713,484	554,781	0.23	515,284	20,263,350	1,987,741	1,887,345	0.84	0.92	9.32	1,195	0.00	0.90	0.44
888	563	GCTATT	167	643,923	486,475	0.26	447,264	17,122,751	2,019,432	1,748,239	0.82	0.92	7.87	1,167	0.00	0.89	0.35
916	580	GTGGTC	160	607,065	426,706	0.31	395,027	15,218,351	1,687,018	1,582,664	0.82	0.91	7.00	1,172	0.01	0.87	0.28
919	582	CCATAG	147	494,475	395,298	0.22	364,623	14,445,778	1,284,100	1,321,306	0.85	0.92	6.64	1,206	0.01	0.85	0.26
921	583	CCGAAT	156	557,947	423,163	0.26	391,703	15,299,597	1,553,707	1,468,978	0.84	0.92	7.04	1,189	0.01	0.87	0.28
960	608	TCCGGT	169	654,861	475,331	0.29	438,742	16,765,670	2,019,805	1,725,538	0.82	0.92	7.71	1,164	0.00	0.89	0.33
966	612	CTCGGA	163	611,595	463,105	0.25	429,427	16,616,998	1,844,587	1,626,719	0.83	0.92	7.64	1,178	0.01	0.89	0.33
969	614	TTGGCG	169	512,315	346,649	0.34	318,826	12,122,879	1,501,810	1,316,106	0.81	0.91	5.57	1,155	0.01	0.85	0.16
973	617	GGTCGT	171	539,006	426,246	0.22	395,315	15,215,610	1,805,309	1,485,456	0.82	0.92	7.00	1,171	0.01	0.88	0.28
996	631	ATTCCG	189	576,792	402,832	0.31	372,270	13,812,218	2,071,558	1,549,397	0.79	0.91	6.35	1,128	0.01	0.88	0.22
1,021	647	GAACTA	159	591,995	425,033	0.30	390,176	15,066,867	1,652,370	1,525,917	0.83	0.92	6.93	1,176	0.01	0.88	0.27
1,031	653	CGACTC	156	504,182	396,277	0.22	369,471	14,505,344	1,448,669	1,345,792	0.84	0.92	6.67	1,194	0.01	0.86	0.26
1,034	655	GAATCG	169	676,216	518,039	0.24	481,853	18,546,697	2,206,603	1,838,388	0.82	0.92	8.53	1,169	0.01	0.90	0.39
1,043	660	TGAGTA	155	496,828	393,558	0.22	364,933	14,313,462	1,415,031	1,349,491	0.84	0.92	6.58	1,193	0.01	0.85	0.25
1,054	667	ATTCTT	176	527,521	427,041	0.20	397,128	15,193,270	1,891,352	1,476,894	0.82	0.92	6.99	1,165	0.01	0.88	0.28
1,055	668	GTGTCC	154	526,491	394,143	0.27	364,346	14,216,109	1,445,953	1,382,221	0.83	0.92	6.54	1,188	0.01	0.86	0.25
1,092	691	AGCCGC	159	414,999	327,941	0.22	301,987	11,781,620	1,214,944	1,148,931	0.83	0.92	5.42	1,186	0.01	0.83	0.16
1,149	728	CTCAGC	157	668,675	527,793	0.22	488,334	19,058,452	1,930,790	1,836,658	0.83	0.92	8.76	1,189	0.01	0.89	0.41
1,152	729	TAAGGT	167	561,404	443,893	0.22	411,723	15,860,829	1,828,230	1,560,137	0.82	0.92	7.29	1,173	0.01	0.88	0.30
1,153	730	ACCGCG	174	571,901	410,500	0.29	379,829	14,403,148	1,845,190	1,562,164	0.81	0.91	6.62	1,151	0.01	0.87	0.25
1,175	744	CACCAC	158	527,198	389,797	0.28	354,576	13,715,623	1,475,182	1,386,675	0.83	0.92	6.31	1,178	0.01	0.85	0.23
1,178	746	TGCAGT	154	659,821	490,510	0.27	450,605	17,598,401	1,766,487	1,702,800	0.84	0.92	8.09	1,189	0.00	0.89	0.36
1,181	748	CATGAC	159	676,136	528,468	0.23	486,899	18,997,426	1,974,352	1,801,874	0.83	0.92	8.74	1,188	0.00	0.90	0.41
1,185	750	CTATTG	154	495,020	392,735	0.22	360,474	14,160,011	1,376,224	1,343,905	0.84	0.92	6.51	1,194	0.01	0.85	0.25
1,193	756	ACAATA	154	532,991	410,752	0.24	380,396	14,875,220	1,490,850	1,436,624	0.84	0.92	6.84	1,190	0.01	0.87	0.27
1,214	769	ACATCA	165	622,398	458,910	0.28	420,674	16,182,823	1,851,402	1,660,483	0.82	0.92	7.44	1,170	0.00	0.88	0.32
1,231	780	TTCCAT	169	502,933	384,657	0.25	355,542	13,663,103	1,629,674	1,331,816	0.82	0.92	6.28	1,170	0.01	0.86	0.22
1,233	781	CCGACG	158	505,955	398,963	0.22	370,459	14,542,794	1,477,435	1,351,735	0.84	0.92	6.69	1,192	0.01	0.86	0.26
1,241	786	ACGGTG	172	575,344	456,483	0.22	424,586	16,368,372	1,940,078	1,587,672	0.82	0.92	7.53	1,171	0.01	0.88	0.32
1,244	788	TATCCT	161	540,895	440,007	0.20	407,096	15,861,593	1,693,483	1,477,585	0.83	0.92	7.29	1,187	0.01	0.88	0.30
1,258	797	CTAAGT	150	523,270	416,878	0.22	384,760	15,136,341	1,423,552	1,415,323	0.84	0.92	6.96	1,199	0.01	0.86	0.28
1,263	800	GTAGCC	152	555,386	434,504	0.23	401,972	15,803,453	1,516,181	1,506,394	0.84	0.92	7.27	1,195	0.01	0.87	0.30
1,267	803	ATACAC	157	612,554	473,458	0.24	438,512	17,079,102	1,763,407	1,664,508	0.83	0.92	7.85	1,186	0.01	0.88	0.35
1,274	807	ATGTAG	151	519,922	409,142	0.23	378,101	14,868,546	1,414,241	1,391,922	0.84	0.92	6.84	1,198	0.01	0.86	0.27
1,312	831	CTTAGA	169	568,142	416,494	0.28	388,308	14,841,355	1,797,631	1,525,952	0.82	0.92	6.82	1,163	0.01	0.88	0.26
1,367	866	GGTTGG	186	551,678	395,741	0.30	364,922	13,652,801	1,941,701	1,476,560	0.80	0.91	6.28	1,139	0.01	0.88	0.21
1,375	871	AGAAGG	161	526,902	409,985	0.24	376,514	14,598,472	1,577,148	1,439,735	0.83	0.92	6.71	1,180	0.01	0.86	0.26
1,384	877	GGAATT	170	647,718	528,581	0.19	494,405	19,112,161	2,202,242	1,843,267	0.83	0.92	8.79	1,175	0.00	0.90	0.41
1,405	890	GGTACG	152	570,068	455,998	0.21	420,922	16,587,053	1,570,472	1,548,708	0.84	0.92	7.63	1,198	0.01	0.87	0.33
1,405	890	CTGACA	153	521,698	375,797	0.30	340,462	13,229,656	1,363,332	1,340,558	0.83	0.92	6.08	1,182	0.01	0.85	0.21
1,423	901	ACTGAT	156	545,362	425,432	0.24	390,846	15,240,985	1,555,521	1,468,040	0.83	0.92	7.01	1,188	0.01	0.87	0.28
1,423	901	TCGGAC	154	514,953	400,492	0.24	367,474	14,381,088	1,432,088	1,371,671	0.84	0.92	6.61	1,192	0.01	0.86	0.25
1,434	908	CGTTAT	159	583,475	430,710	0.27	398,047	15,416,016	1,663,262	1,517,932	0.83	0.92	7.09	1,180	0.01	0.88	0.29
1,505	953	CATCAA	160	621,411	493,653	0.22	456,162	17,775,671	1,881,578	1,686,750	0.83	0.92	8.17	1,186	0.01	0.89	0.37
1,505	953	ACTCAC	156	448,295	348,339	0.24	318,488	12,411,666	1,276,525	1,187,636	0.83	0.92	5.71	1,188	0.01	0.83	0.18
1,510	956	CTGCTG	153	518,490	421,578	0.20	391,823	15,508,708	1,460,888	1,391,743	0.84	0.92	7.13	1,203	0.01	0.86	0.29
1,515	959	GCAGCT	155	569,089	423,456	0.27	389,954	15,194,836	1,560,707	1,499,655	0.83	0.92	6.99	1,185	0.01	0.86	0.28
1,519	962	ATTACT	170	544,805	436,425	0.21	404,264	15,570,991	1,835,275	1,493,038	0.82	0.92	7.16	1,173	0.01	0.88	0.29

(continued)

DNA amount in ng for shearing	DNA amount in ng into library	barcode	insert size	PF reads	PF unique reads	% duplication	PF unique reads aligned	on bait bases	near bait bases (250bp either side)	off bait bases	% bases directly on bait	% selected bases (on and near bait)	mean bait coverage	fold enrichment	% target with no reads	% target with 2x coverage	% target with 10x coverage
1,610	1,020	AATTGA	168	533,575	434,168	0.20	403,248	15,687,223	1,729,401	1,463,624	0.83	0.92	7.21	1,183	0.01	0.85	0.29
1,625	1,029	TGTTCA	164	507,928	404,964	0.21	375,218	14,570,709	1,605,260	1,381,080	0.83	0.92	6.70	1,182	0.01	0.86	0.26
1,660	1,051	GAAGTC	165	485,547	375,165	0.24	345,055	13,346,183	1,501,144	1,303,489	0.83	0.92	6.14	1,177	0.01	0.86	0.21
1,671	1,058	TTCATA	147	592,363	452,860	0.25	415,525	16,407,740	1,502,154	1,514,002	0.84	0.92	7.54	1,203	0.01	0.87	0.32
1,685	1,067	AGCGCA	176	563,590	415,511	0.27	384,469	14,623,320	1,875,017	1,520,976	0.81	0.92	6.72	1,156	0.01	0.87	0.26
1,836	1,163	ATGTGT	144	526,899	406,798	0.24	378,362	15,000,699	1,306,631	1,360,137	0.85	0.92	6.90	1,209	0.01	0.86	0.28
1,896	1,201	TCTTCT	160	559,350	451,125	0.20	420,190	16,429,653	1,706,807	1,522,884	0.84	0.92	7.55	1,190	0.01	0.88	0.32
1,918	1,214	TGACCT	156	626,704	477,171	0.25	441,940	17,253,984	1,762,050	1,648,363	0.83	0.92	7.93	1,189	0.01	0.89	0.35
1,920	1,216	GATCGG	169	593,320	459,188	0.24	425,808	16,407,363	1,935,842	1,591,477	0.82	0.92	7.54	1,172	0.01	0.89	0.32
1,970	1,248	GTGCCT	180	496,825	362,454	0.29	332,696	12,560,415	1,657,187	1,337,485	0.81	0.91	5.78	1,150	0.01	0.85	0.18
1,980	1,254	CCGCGT	165	614,156	490,936	0.21	454,472	17,655,442	1,950,910	1,677,747	0.83	0.92	8.12	1,181	0.01	0.89	0.36
2,004	1,269	CCTCCA	162	578,509	471,397	0.20	435,937	17,025,992	1,794,882	1,584,065	0.83	0.92	7.83	1,188	0.01	0.88	0.34
2,012	1,274	ACCACA	161	461,958	367,258	0.22	339,545	13,237,396	1,392,368	1,263,522	0.83	0.92	6.09	1,186	0.01	0.85	0.21
2,099	1,329	GAGGCA	151	546,725	431,858	0.22	402,671	15,888,866	1,510,261	1,463,521	0.84	0.92	7.31	1,199	0.01	0.87	0.31
2,208	1,399	GCTGCC	168	497,232	401,914	0.21	371,419	14,382,822	1,619,678	1,382,728	0.83	0.92	6.61	1,178	0.01	0.86	0.25
2,262	1,433	TAGCAG	161	543,490	432,002	0.22	399,859	15,622,041	1,651,340	1,456,408	0.83	0.92	7.18	1,188	0.01	0.86	0.30
2,349	1,488	GGCGAC	172	464,376	377,252	0.20	350,373	13,484,853	1,620,129	1,305,708	0.82	0.92	6.20	1,170	0.01	0.86	0.22
2,460	1,558	AGGCGA	166	599,494	449,095	0.26	414,485	15,973,364	1,829,618	1,607,469	0.82	0.92	7.34	1,172	0.01	0.88	0.31
2,630	1,666	CAGGTT	163	583,197	420,366	0.29	388,574	15,002,473	1,692,875	1,475,663	0.83	0.92	6.90	1,176	0.01	0.87	0.27
2,644	1,674	TGTGTG	157	683,914	547,034	0.21	507,143	19,892,778	1,979,353	1,846,959	0.84	0.92	9.15	1,194	0.01	0.90	0.43
1,326	840	average	162	557,599	428,831	0.24	396,483	15,388,379	1,668,292	1,495,929	0.83	0.92	7.08	1,181	0.01	0.87	0.28
484	307	stdev	9	61,363	47,965	0.04	45,105	1,804,371	225,694	161,392	0.01	0.00	0.83	15	0.00	0.02	0.06

Table S3: Hybrid selection statistics comparison with literature, single sample and multiple sample selections, On-array (MGS) and In-solution (SHS)

study	capture type	barcoding technology	sequencing technology	number of samples per selection	% aligned bases / reads on target
Single sample selection					
Hodges (2007)	MGS		Illumina	1	55-85*
Teer (2010)	MGS		Illumina	1	58-61
Bainbridge (2010)	MGS		Illumina	1	78
Mokry (2010)	MGS		SOLiD	1	40-61
Bainbridge (2010)	MGS		SOLiD	1	48-51
Gnirke (2009)	SHS		Illumina	1	50-90*
Teer (2010)	SHS		Illumina	1	62-65
Blumenstiel (2010)	SHS		Illumina	1	66-92*
Cummings (2010)	SHS		Illumina	1	23
Fisher (2011)	SHS		Illumina	1	84-89*
Multiple sample selection					
Teer (2010)	MGS	Multiplex	Illumina	12	51-61
Meyer (2010)	MGS	Multiplex	Illumina	50	25
Nijman (2010)	MGS	SOLiD barcodes	SOLiD	20	58-62
Cummings (2010)	SHS	Multiplex	Illumina	2	20-34
Cummings (2010)	SHS	Multiplex	Illumina	5	35-56
Kenny (2010)	SHS	Internal	Illumina	1-3	20-22**
Kenny (2010)	SHS	Internal	Illumina	9	18**
this study	SHS	Internal	Illumina	14	72-82* (55-67)***
	SHS	Internal	Illumina	28	55-82* (42-69)***
	SHS	Internal	Illumina	52	65-83* (50-70)***
	SHS	Internal	Illumina	95	70-79* (52-66)***
	SHS	Internal	Illumina	81	91-92* (79-85)***

Notes: MGS - Microarray-based Genomic Selection, SHS - solution hybrid selection; definitions for % aligned reads (or bases) on target differs between the studies – specified when indicated, * within 250bp window on either side of target, ** within 50bp window on either side of target, *** directly on bait; in the cited papers, it is sometimes not clear if the duplicated reads were removed prior to calculation of on-target percentages.

Table S4: Results for application 2 - human WGS sequencing per library (barcode).

	barcode	PF reads	% duplication	genome coverage	mean insert size
	ACAATA	73,026,973	14.90	5.08	288
	CACCAC	73,817,069	11.29	5.14	286
	CGTTAT	78,343,835	16.19	5.45	287
sample 1	GATCGG	73,661,145	15.34	5.13	287
	ATGTGT	64,569,343	14.42	4.61	260
	CTCAGC	74,741,030	16.43	5.34	266
	GGTACG	72,578,526	12.66	5.18	270
sample 2	GAGGCA	69,472,479	14.17	4.96	260
	TGAGTA	69,091,931	13.93	5.09	273
	TAAGGT	55,670,585	14.83	4.10	278
	CCGACG	71,882,072	17.25	5.30	275
sample 3	GCATTC	66,490,070	12.93	4.90	270
	TAACAT	71,846,548	15.18	5.28	279
	AGGTCG	75,091,144	13.11	5.52	273
	GTGAAC	58,476,892	14.47	4.30	275
sample 4	CCTGTA	87,121,560	10.61	6.40	268
	TTGGCG	97,787,920	17.84	7.53	284
	TTCATA	84,435,130	14.14	6.50	284
	GCTGCC	75,106,399	11.71	5.79	283
sample 5	CCGCGT	73,728,855	11.66	5.68	283
	AGCGCA	76,412,147	16.25	5.31	274
	CCATAG	76,613,748	16.70	5.33	275
	GTGCCT	70,922,121	17.18	4.93	265
sample 6	TAGATC	69,361,327	15.02	4.82	270
	GCGCTA	74,253,094	12.14	7.46	280
	AGGTTC	75,336,265	17.03	7.57	277
	TACGAT	72,396,823	19.89	7.28	281
sample 7	AGAAGG	60,504,339	15.00	6.08	285
	ATACAC	55,074,060	19.62	4.03	260
	TGTGTG	83,491,000	15.40	6.12	274
	ACTGAT	50,946,362	16.14	3.73	268
sample 8	CACTTA	76,663,162	12.90	5.62	271
	CGACTC	75,605,403	8.85	5.93	254
	AATACA	63,646,367	12.26	4.99	263
	GCTAAG	67,421,275	9.89	5.29	261
sample 9	TTCCAT	75,142,367	13.38	5.89	265
	TCCTCC	105,507,570	17.00	7.78	272
	CTATGA	64,816,053	11.18	4.78	269
	ATGTAG	22,688,924	14.51	1.67	257
sample 10	GGTCGT	61,970,542	12.09	4.57	272

Table S5: Results for application 3 – microbial sequencing.

sample	barcode	PF reads	% duplication	genome coverage	mean insert size in bp
strain01	AGCGCA	17,736,070	1.53	169	287
strain02	TAGATC	17,685,204	0.88	171	320
strain03	CCATAG	20,828,422	0.96	201	331
strain04	GTGCCT	11,837,054	0.71	113	312
strain05	ATACAC	12,966,184	0.89	124	281
strain06	TGTGTG	13,489,352	0.80	131	278
strain07	ACTGAT	15,439,906	0.91	149	314
strain08	CACTTA	14,021,406	0.88	136	294
strain09	CTCAGC	14,337,584	0.86	138	283
strain10	GAGGCA	13,929,826	0.70	135	308
strain11	GGTACG	15,839,758	1.11	152	294
strain12	ATGTGT	15,056,200	1.89	143	305

Table S6: Oligonucleotide sequences: Adapter (a), primer (b), and blocker oligonucleotides for hybrid selection with 'truncated' adapters sequences (c). All oligonucleotides are given in 5'->3' direction, 'molecular barcodes' of the barcoded-P5 oligonucleotides are given in capital letters and are read in the first 6 cycles of the first sequencing read. Indexing PCR primers (b) are designed for the PE adapter, a custom sequencing primer to read out the index is needed. See Meyer & Kircher 2010(Meyer and Kircher 2010) for multiplexing indexing primer sequences. Oligonucleotide sequences (in lowercase letters)

(a)

barcoded-P5	Barcoded-P5-comp	Barcode
ctttccctacacgacgctcttccgatctCAGGTT	AACCTGagatcggaa	CAGGTT
ctttccctacacgacgctcttccgatctTCACAA	TTGTGAagatcggaa	TCACAA
ctttccctacacgacgctcttccgatctACATCA	TGATGTagatcggaa	ACATCA
ctttccctacacgacgctcttccgatctAGCGCA	TGCGCTagatcggaa	AGCGCA
ctttccctacacgacgctcttccgatctCATCAA	TTGATGagatcggaa	CATCAA
ctttccctacacgacgctcttccgatctGCTATT	AATAGCagatcggaa	GCTATT
ctttccctacacgacgctcttccgatctTAGATC	GATCTAagatcggaa	TAGATC
ctttccctacacgacgctcttccgatctCATGAC	GTCATGagatcggaa	CATGAC
ctttccctacacgacgctcttccgatctGAATCG	CGATTCagatcggaa	GAATCG
ctttccctacacgacgctcttccgatctTCTTCT	AGAAGAagatcggaa	TCTTCT
ctttccctacacgacgctcttccgatctATTCCG	CGGAATagatcggaa	ATTCCG
ctttccctacacgacgctcttccgatctGGAATT	AATTCCagatcggaa	GGAATT
ctttccctacacgacgctcttccgatctACGGTG	CACCGTagatcggaa	ACGGTG
ctttccctacacgacgctcttccgatctCTCAGC	GCTGAGagatcggaa	CTCAGC
ctttccctacacgacgctcttccgatctTCCGGT	ACCGGAagatcggaa	TCCGGT
ctttccctacacgacgctcttccgatctTGCAGT	ACTGCAagatcggaa	TGCAGT
ctttccctacacgacgctcttccgatctTTCATA	TATGAAagatcggaa	TTCATA
ctttccctacacgacgctcttccgatctATACAC	GTGTATagatcggaa	ATACAC
ctttccctacacgacgctcttccgatctCGTTAT	ATAACGagatcggaa	CGTTAT
ctttccctacacgacgctcttccgatctCTCGGA	TCCGAGagatcggaa	CTCGGA
ctttccctacacgacgctcttccgatctTGTGTG	CACACAagatcggaa	TGTGTG
ctttccctacacgacgctcttccgatctACCGCG	CGCGGTagatcggaa	ACCGCG
ctttccctacacgacgctcttccgatctGATCGG	CCGATCagatcggaa	GATCGG
ctttccctacacgacgctcttccgatctTCACGG	CCGTGAagatcggaa	TCACGG
ctttccctacacgacgctcttccgatctATTACT	AGTAATagatcggaa	ATTACT
ctttccctacacgacgctcttccgatctCTTAGA	TCTAAGagatcggaa	CTTAGA
ctttccctacacgacgctcttccgatctGCAGCT	AGCTGCagatcggaa	GCAGCT
ctttccctacacgacgctcttccgatctTCCTCC	GGAGGAagatcggaa	TCCTCC
ctttccctacacgacgctcttccgatctGAACTA	TAGTTCagatcggaa	GAACTA
ctttccctacacgacgctcttccgatctACAACC	GGTTGTagatcggaa	ACAACC
ctttccctacacgacgctcttccgatctGGTAAC	GTTACCagatcggaa	GGTAAC
ctttccctacacgacgctcttccgatctGTGGTC	GACCACagatcggaa	GTGGTC
ctttccctacacgacgctcttccgatctCCGCGT	ACGCGGagatcggaa	CCGCGT
ctttccctacacgacgctcttccgatctCTGACA	TGTCAGagatcggaa	CTGACA
ctttccctacacgacgctcttccgatctCCGAAT	ATTCGGagatcggaa	CCGAAT
ctttccctacacgacgctcttccgatctAGCCGC	GCGGCTagatcggaa	AGCCGC
ctttccctacacgacgctcttccgatctTAGCGC	GCGCTAagatcggaa	TAGCGC
ctttccctacacgacgctcttccgatctTGACCT	AGGTCAagatcggaa	TGACCT
ctttccctacacgacgctcttccgatctCTTATC	GATAAGagatcggaa	CTTATC
ctttccctacacgacgctcttccgatctGTAGCC	GGCTACagatcggaa	GTAGCC
ctttccctacacgacgctcttccgatctCCATAG	CTATGGagatcggaa	CCATAG
ctttccctacacgacgctcttccgatctGAGGCA	TGCCTCagatcggaa	GAGGCA
ctttccctacacgacgctcttccgatctAATTGA	TCAATTagatcggaa	AATTGA
ctttccctacacgacgctcttccgatctACTCAC	GTGAGTagatcggaa	ACTCAC

ctttccctacacgacgctcttccgatctAAGTTG	CAACTTagatcggaa	AAGTTG
ctttccctacacgacgctcttccgatctTACGAT	ATCGTAagatcggaa	TACGAT
ctttccctacacgacgctcttccgatctCACCAC	GTGGTGagatcggaa	CACCAC
ctttccctacacgacgctcttccgatctGCATTC	GAATGCagatcggaa	GCATTC
ctttccctacacgacgctcttccgatctTTCCAT	ATGGAAagatcggaa	TTCCAT
ctttccctacacgacgctcttccgatctAGGCGA	TCGCCTagatcggaa	AGGCGA
ctttccctacacgacgctcttccgatctCTCTCG	CGAGAGagatcggaa	CTCTCG
ctttccctacacgacgctcttccgatctGCCTTG	CAAGGCagatcggaa	GCCTTG
ctttccctacacgacgctcttccgatctAGGTTC	GAACCTagatcggaa	AGGTTC
ctttccctacacgacgctcttccgatctTAGCAG	CTGCTAagatcggaa	TAGCAG
ctttccctacacgacgctcttccgatctCACTGC	GCAGTGagatcggaa	CACTGC
ctttccctacacgacgctcttccgatctAGTGGC	GCCACTagatcggaa	AGTGGC
ctttccctacacgacgctcttccgatctATGTAG	CTACATagatcggaa	ATGTAG
ctttccctacacgacgctcttccgatctGGTACG	CGTACCagatcggaa	GGTACG
ctttccctacacgacgctcttccgatctTAAGGT	ACCTTAagatcggaa	TAAGGT
ctttccctacacgacgctcttccgatctCCTGTA	TACAGGagatcggaa	CCTGTA
ctttccctacacgacgctcttccgatctCCTCCA	TGGAGGagatcggaa	CCTCCA
ctttccctacacgacgctcttccgatctACAATA	TATTGTagatcggaa	ACAATA
ctttccctacacgacgctcttccgatctACTGAT	ATCAGTagatcggaa	ACTGAT
ctttccctacacgacgctcttccgatctATGTGT	ACACATagatcggaa	ATGTGT
ctttccctacacgacgctcttccgatctGTGTCC	GGACACagatcggaa	GTGTCC
ctttccctacacgacgctcttccgatctTAGAGT	ACTCTAagatcggaa	TAGAGT
ctttccctacacgacgctcttccgatctGCTAAG	CTTAGCagatcggaa	GCTAAG
ctttccctacacgacgctcttccgatctTGGACT	AGTCCAagatcggaa	TGGACT
ctttccctacacgacgctcttccgatctCACTTA	TAAGTGagatcggaa	CACTTA
ctttccctacacgacgctcttccgatctCTATTG	CAATAGagatcggaa	CTATTG
ctttccctacacgacgctcttccgatctAATACA	TGTATTagatcggaa	AATACA
ctttccctacacgacgctcttccgatctATTCTT	AAGAATagatcggaa	ATTCTT
ctttccctacacgacgctcttccgatctGGCATC	GATGCCagatcggaa	GGCATC
ctttccctacacgacgctcttccgatctGGTTGG	CCAACCagatcggaa	GGTTGG
ctttccctacacgacgctcttccgatctTTGGCG	CGCCAAagatcggaa	TTGGCG
ctttccctacacgacgctcttccgatctCTGCTG	CAGCAGagatcggaa	CTGCTG
ctttccctacacgacgctcttccgatctTGAGTA	TACTCAagatcggaa	TGAGTA
ctttccctacacgacgctcttccgatctCTAAGT	ACTTAGagatcggaa	CTAAGT
ctttccctacacgacgctcttccgatctAGAAGG	CCTTCTagatcggaa	AGAAGG
ctttccctacacgacgctcttccgatctTCGGAC	GTCCGAagatcggaa	TCGGAC
ctttccctacacgacgctcttccgatctTGTTCA	TGAACAagatcggaa	TGTTCA
ctttccctacacgacgctcttccgatctAACCTT	AAGGTTagatcggaa	AACCTT
ctttccctacacgacgctcttccgatctGGTCGT	ACGACCagatcggaa	GGTCGT
ctttccctacacgacgctcttccgatctTATCCT	AGGATAagatcggaa	TATCCT
ctttccctacacgacgctcttccgatctGGCGAC	GTCGCCagatcggaa	GGCGAC
ctttccctacacgacgctcttccgatctGTGCCT	AGGCACagatcggaa	GTGCCT
ctttccctacacgacgctcttccgatctGCTGCC	GGCAGCagatcggaa	GCTGCC
ctttccctacacgacgctcttccgatctCCGACG	CGTCGGagatcggaa	CCGACG
ctttccctacacgacgctcttccgatctACCACA	TGTGGTagatcggaa	ACCACA
ctttccctacacgacgctcttccgatctCTATGA	TCATAGagatcggaa	CTATGA
ctttccctacacgacgctcttccgatctTACTCT	AGAGTAagatcggaa	TACTCT
ctttccctacacgacgctcttccgatctGACAGA	TCTGTCagatcggaa	GACAGA
ctttccctacacgacgctcttccgatctGCGCTA	TAGCGCagatcggaa	GCGCTA
ctttccctacacgacgctcttccgatctGAAGTC	GACTTCagatcggaa	GAAGTC
ctttccctacacgacgctcttccgatctTTGATT	AATCAAagatcggaa	TTGATT
ctttccctacacgacgctcttccgatctCGACTC	GAGTCGagatcggaa	CGACTC
ctttccctacacgacgctcttccgatctAGGATG	CATCCTagatcggaa	AGGATG
ctttccctacacgacgctcttccgatctATAGAA	TTCTATagatcggaa	ATAGAA
ctttccctacacgacgctcttccgatctACATGT	ACATGTagatcggaa	ACATGT
ctttccctacacgacgctcttccgatctACCAGG	CCTGGTagatcggaa	ACCAGG
ctttccctacacgacgctcttccgatctAGGTAA	TTACCTagatcggaa	AGGTAA
ctttccctacacgacgctcttccgatctATCAAG	CTTGATagatcggaa	ATCAAG
ctttccctacacgacgctcttccgatctATACCA	TGGTATagatcggaa	ATACCA

ctttccctacacgacgctcttccgatctCCGAGC	GCTCGGagatcggaa	CCGAGC
ctttccctacacgacgctcttccgatctACTTGC	GCAAGTagatcggaa	ACTTGC
ctttccctacacgacgctcttccgatctACTTAA	TTAAGTagatcggaa	ACTTAA
ctttccctacacgacgctcttccgatctCTTGGT	ACCAAGagatcggaa	CTTGGT
ctttccctacacgacgctcttccgatctAGATAC	GTATCTagatcggaa	AGATAC
ctttccctacacgacgctcttccgatctCTCCGG	CCGGAGagatcggaa	CTCCGG
ctttccctacacgacgctcttccgatctAGGCAT	ATGCCTagatcggaa	AGGCAT
ctttccctacacgacgctcttccgatctAGAGAT	ATCTCTagatcggaa	AGAGAT
ctttccctacacgacgctcttccgatctACGTGG	CCACGTagatcggaa	ACGTGG
ctttccctacacgacgctcttccgatctCAGTAC	GTACTGagatcggaa	CAGTAC
ctttccctacacgacgctcttccgatctATCGCT	AGCGATagatcggaa	ATCGCT
ctttccctacacgacgctcttccgatctACCTAT	ATAGGTagatcggaa	ACCTAT
ctttccctacacgacgctcttccgatctGATGCG	CGCATCagatcggaa	GATGCG
ctttccctacacgacgctcttccgatctTCTCTG	CAGAGAagatcggaa	TCTCTG
ctttccctacacgacgctcttccgatctAATAAT	ATTATTagatcggaa	AATAAT
ctttccctacacgacgctcttccgatctTGAGCG	CGCTCAagatcggaa	TGAGCG
ctttccctacacgacgctcttccgatctTATCTC	GAGATAagatcggaa	TATCTC
ctttccctacacgacgctcttccgatctTGACAG	CTGTCAagatcggaa	TGACAG
ctttccctacacgacgctcttccgatctCTATCT	AGATAGagatcggaa	CTATCT
ctttccctacacgacgctcttccgatctGTCTTC	GAAGACagatcggaa	GTCTTC
ctttccctacacgacgctcttccgatctGTGAAC	GTTCACagatcggaa	GTGAAC
ctttccctacacgacgctcttccgatctGGCCTG	CAGGCCagatcggaa	GGCCTG
ctttccctacacgacgctcttccgatctAGTGAA	TTCACTagatcggaa	AGTGAA
ctttccctacacgacgctcttccgatctGTAGTT	AACTACagatcggaa	GTAGTT
ctttccctacacgacgctcttccgatctTGTGCT	AGCACAagatcggaa	TGTGCT
ctttccctacacgacgctcttccgatctTATAGC	GCTATAagatcggaa	TATAGC
ctttccctacacgacgctcttccgatctACCAAC	GTTGGTagatcggaa	ACCAAC
ctttccctacacgacgctcttccgatctCGCGTT	AACGCGagatcggaa	CGCGTT
ctttccctacacgacgctcttccgatctATGCAA	TTGCATagatcggaa	ATGCAA
ctttccctacacgacgctcttccgatctCTCTAA	TTAGAGagatcggaa	CTCTAA
ctttccctacacgacgctcttccgatctTGAACA	TGTTCAagatcggaa	TGAACA
ctttccctacacgacgctcttccgatctTGTCTA	TAGACAagatcggaa	TGTCTA
ctttccctacacgacgctcttccgatctAGGTCG	CGACCTagatcggaa	AGGTCG
ctttccctacacgacgctcttccgatctGCAGGA	TCCTGCagatcggaa	GCAGGA
ctttccctacacgacgctcttccgatctGACACG	CGTGTCagatcggaa	GACACG
ctttccctacacgacgctcttccgatctACCTGA	TCAGGTagatcggaa	ACCTGA
ctttccctacacgacgctcttccgatctTACGTG	CACGTAagatcggaa	TACGTG
ctttccctacacgacgctcttccgatctGATTAG	CTAATCagatcggaa	GATTAG
ctttccctacacgacgctcttccgatctGTTGAG	CTCAACagatcggaa	GTTGAG
ctttccctacacgacgctcttccgatctACACTT	AAGTGTagatcggaa	ACACTT
ctttccctacacgacgctcttccgatctATTCGC	GCGAATagatcggaa	ATTCGC
ctttccctacacgacgctcttccgatctAATTCC	GGAATTagatcggaa	AATTCC
ctttccctacacgacgctcttccgatctGTGAGG	CCTCACagatcggaa	GTGAGG
ctttccctacacgacgctcttccgatctGATATG	CATATCagatcggaa	GATATG
ctttccctacacgacgctcttccgatctAGCTAG	CTAGCTagatcggaa	AGCTAG
ctttccctacacgacgctcttccgatctTTGGAT	ATCCAAagatcggaa	TTGGAT
ctttccctacacgacgctcttccgatctCTAACG	CGTTAGagatcggaa	CTAACG
ctttccctacacgacgctcttccgatctTACGCC	GGCGTAagatcggaa	TACGCC
ctttccctacacgacgctcttccgatctCAAGAG	CTCTTGagatcggaa	CAAGAG
ctttccctacacgacgctcttccgatctTAACAT	ATGTTAagatcggaa	TAACAT
ctttccctacacgacgctcttccgatctCTGTTA	TAACAGagatcggaa	CTGTTA
ctttccctacacgacgctcttccgatctTCCGTC	GACGGAagatcggaa	TCCGTC
ctttccctacacgacgctcttccgatctGGCCAA	TTGGCCagatcggaa	GGCCAA
ctttccctacacgacgctcttccgatctTAATAA	TTATTAagatcggaa	TAATAA
ctttccctacacgacgctcttccgatctGTTGCA	TGCAACagatcggaa	GTTGCA
ctttccctacacgacgctcttccgatctGCCGAA	TTCGGCagatcggaa	GCCGAA
PE-P7	**PE-P7-comp**	
ctcggcattcctgctgaaccgctcttccgatct	agatcggaagagc	

(b)

PreHyb-PE_F	ctttccctacacgacgctcttc	Annealing temp
PreHyb-PE_R	ctcggcattcctgctgaacc	55°C
Sol-PE-PCR_F	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttc	
Sol-PE-PCR_R	caagcagaagacggcatacgagatcggtctcggcattcctgctgaacc	62°C
Sol-PE-index_R: (with Sol-PE-PCR_F)		index read
Index01-PE-Primer	caagcagaagacggcatacgagatCCTGCGAcggtctcggcattcctgctgaacc	TCGCAGG
Index02-PE-Primer	caagcagaagacggcatacgagatTGCAGAGcggtctcggcattcctgctgaacc	CTCTGCA
Index03-PE-Primer	caagcagaagacggcatacgagatCGCATTAcggtctcggcattcctgctgaacc	TAATGCG
Index04-PE-Primer	caagcagaagacggcatacgagatTTGATCCcggtctcggcattcctgctgaacc	GGATCAA
Index05-PE-Primer	caagcagaagacggcatacgagatATCTTGCcggtctcggcattcctgctgaacc	GCAAGAT
Index06-PE-Primer	caagcagaagacggcatacgagatTCTCCATcggtctcggcattcctgctgaacc	ATGGAGA
Index07-PE-Primer	caagcagaagacggcatacgagatCATCGAGcggtctcggcattcctgctgaacc	CTCGATG
Index08-PE-Primer	caagcagaagacggcatacgagatTTCGAGCcggtctcggcattcctgctgaacc	GCTCGAA
Index09-PE-Primer	caagcagaagacggcatacgagatAGTTGGTcggtctcggcattcctgctgaacc	ACCAACT
Index10-PE-Primer	caagcagaagacggcatacgagatGGACGCAcggtctcggcattcctgctgaacc	TGCGTCC
Index11-PE-Primer	caagcagaagacggcatacgagatCGGAGTTcggtctcggcattcctgctgaacc	AACTCCG
Index12-PE-Primer	caagcagaagacggcatacgagatAGGTATTcggtctcggcattcctgctgaacc	AATACCT
Index13-PE-Primer	caagcagaagacggcatacgagatTGATAGTcggtctcggcattcctgctgaacc	ACTATCA
Index14-PE-Primer	caagcagaagacggcatacgagatGATCCAAcggtctcggcattcctgctgaacc	TTGGATC
Index15-PE-Primer	caagcagaagacggcatacgagatCAGGTCGcggtctcggcattcctgctgaacc	CGACCTG
Index16-PE-Primer	caagcagaagacggcatacgagatCCGGATGcggtctcggcattcctgctgaacc	CATCCGG
index-PE-sequencing-Primer	gatcggaagagcggttcagcaggaatgccgagaccg	
Sol-PE-qPCR_F	aatgatacggcgaccacc	
Sol-PE-qPCR_R	caagcagaagacggcatacg	55°C

(c)

Univ_Block_P5	agatcggaagagcgtcgtgtagggaaag
Univ_Block_P7	agatcggaagagcggttcagcaggaatgccgag

References

Aird, D., M.G. Ross, W.S. Chen, M. Danielsson, T. Fennell, C. Russ, D.B. Jaffe, C. Nusbaum, and A. Gnirke. 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* **12:** R18.

Bainbridge, M.N., M. Wang, D.L. Burgess, C. Kovar, M.J. Rodesch, M. D'Ascenzo, J. Kitzman, Y.Q. Wu, I. Newsham, T.A. Richmond et al. 2010. Whole exome capture in solution with 3 Gbp of data. *Genome Biol* **11:** R62.

Blumenstiel, B., K. Cibulskis, S. Fisher, M. DeFelice, A. Barry, T. Fennell, J. Abreu, B. Minie, M. Costello, G. Young et al. 2010. Targeted exon sequencing by in-solution hybrid selection. *Curr Protoc Hum Genet* **Chapter 18:** Unit 18 14.

Borgstrom, E., S. Lundin, and J. Lundeberg. 2011. Large scale library generation for high throughput sequencing. *PLoS One* **6:** e19119.

Cummings, N., R. King, A. Rickers, A. Kaspi, S. Lunke, I. Haviv, and J.B. Jowett. 2010. Combining target enrichment with barcode multiplexing for high throughput SNP discovery. *BMC Genomics* **11:** 641.

Fisher, S., A. Barry, J. Abreu, B. Minie, J. Nolan, T.M. Delorey, G. Young, T.J. Fennell, A. Allen, L. Ambrogio et al. 2011. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* **12:** R1.

Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E.M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher, C. Russ et al. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* **27:** 182-189.

Hodges, E., Z. Xuan, V. Balija, M. Kramer, M.N. Molla, S.W. Smith, C.M. Middle, M.J. Rodesch, T.J. Albert, G.J. Hannon et al. 2007. Genome-wide in situ exon capture for selective resequencing. *Nat Genet* **39:** 1522-1527.

http://picard.sourceforge.net.

Kenny, E.M., P. Cormican, W.P. Gilks, A.S. Gates, C.T. O'Dushlaine, C. Pinto, A.P. Corvin, M. Gill, and D.W. Morris. 2010. Multiplex target enrichment using DNA indexing for ultra-high throughput SNP detection. *DNA Res* **18:** 31-38.

Lennon, N.J., R.E. Lintner, S. Anderson, P. Alvarez, A. Barry, W. Brockman, R. Daza, R.L. Erlich, G. Giannoukos, L. Green et al. 2010. A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. *Genome Biol* **11:** R15.

Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:** 1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078-2079.

Meyer, M. and M. Kircher. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010:** pdb prot5448.

Mokry, M., H. Feitsma, I.J. Nijman, E. de Bruijn, P.J. van der Zaag, V. Guryev, and E. Cuppen. 2010. Accurate SNP and mutation detection by targeted custom microarray-based genomic enrichment of short-fragment sequencing libraries. *Nucleic Acids Res* **38:** e116.

Nijman, I.J., M. Mokry, R. van Boxtel, P. Toonen, E. de Bruijn, and E. Cuppen. 2010. Mutation discovery by targeted genomic enrichment of multiplexed barcoded samples. *Nat Methods* **7:** 913-915.

Teer, J.K., L.L. Bonnycastle, P.S. Chines, N.F. Hansen, N. Aoyama, A.J. Swift, H.O. Abaan, T.J. Albert, E.H. Margulies, E.D. Green et al. 2010. Systematic comparison of three genomic enrichment methods for massively parallel DNA sequencing. *Genome Res* **20:** 1420-1431.