Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform

Martin Kircher, Susanna Sawyer, and Matthias Meyer

CONTENTS

SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1. Primer sequences
Supplementary Table 2. Index combination used for each library
Supplementary Table 3: Fraction of false pairs observed when using different image
analysis and base calling software
Supplementary Table 4: Quantifying false sample assignments in human sequencing
data from single-indexed libraries
Supplementary Table 5: Quantifying false sample assignments in human and mouse
transcriptome sequencing data from single-indexed libraries
Supplementary Figure 1: Raw intensity values of false index pairs
Supplementary Figure 2 : Changes in the fraction of false index pairs

S	UPPLEMENTARY METHODS	7
	1. Quantifying false index pairs expected based on sequencing error alone	7
	2. Analyzing intensity values from a random set of clusters with false index pairs	8
	3. Reading each index twice from short-insert molecules	11
	4. Quantifying jumping PCR and other effects	11
	5. Estimating false-assignment rates based on the occurrence of unused indexes	16
	5. Identification of false sample assignments in single indexed data	17

SUPPLEMENTARY REFERENCES	, 1	17	1
--------------------------	-----	----	---

Supplementary Table 1. Primer sequences All primers were synthesized by Sigma-Aldrich (Steinheim, Germany). Indexing primers were purified using reverse phase cartridges (RPC). All other primers were purified by HPLC.

Primer ID / index	Sequence (5' -> 3')
P7 indexing primers (first	index)
Sol_iPCR-MPI-97 AATCTTC	CAAGCAGAAGACGGCATACGAGATgaagattGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-98 ACCAACG	CAAGCAGAAGACGGCATACGAGATcgttggtGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-99 AGATGGC	CAAGCAGAAGACGGCATACGAGATgccatctGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-100 CCAGGTT	CAAGCAGAAGACGGCATACGAGATaacctggGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-101 CCGTTAG	CAAGCAGAAGACGGCATACGAGATctaacggGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-102 CGCCTCT	CAAGCAGAAGACGGCATACGAGATagaggcgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-103 CTTGCGG	CAAGCAGAAGACGGCATACGAGATccgcaagGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-104 GGCGGAG	CAAGCAGAAGACGGCATACGAGATctccgccGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-105 TGGACGT	CAAGCAGAAGACGGCATACGAGATacgtccaGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-106 AACCATG	CAAGCAGAAGACGGCATACGAGATcatggttGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-107 CAGGAAG	CAAGCAGAAGACGGCATACGAGATcttcctgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-108 CATACCT	CAAGCAGAAGACGGCATACGAGATaggtatgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-109 CCAATCC	CAAGCAGAAGACGGCATACGAGATggattggGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-110 CCGGCGT	CAAGCAGAAGACGGCATACGAGATacgccggGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-111 CGCATAG	CAAGCAGAAGACGGCATACGAGATctatgcgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-112 CGTAATC	CAAGCAGAAGACGGCATACGAGATgattacgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-113 CGTTGGT	CAAGCAGAAGACGGCATACGAGATaccaacgGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-114 CTATACG	CAAGCAGAAGACGGCATACGAGATcgtatagGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-115 GACCTAC	CAAGCAGAAGACGGCATACGAGATgtaggtcGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-116 GATATTG	CAAGCAGAAGACGGCATACGAGATcaatatcGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-117 AAGACGC	CAAGCAGAAGACGGCATACGAGATgcgtcttGTGACTGGAGTTCAGACGTGT
Sol_iPCR-MPI-118 GCAGTAT	CAAGCAGAAGACGGCATACGAGATatactgcGTGACTGGAGTTCAGACGTGT
Sol_iPCR- ϕX TTGCCGC	CAAGCAGAAGACGGCATACGAGATgcggcaaGTGACTGGAGTTCAGACGTGT
P5 indexing primers (second	l index)
P5_iPCR-LP-1 TCGCAGG	AATGATACGGCGACCACCGAGATCTACACcctgcgaACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-2 CTCTGCA	AATGATACGGCGACCACCGAGATCTACACtgcagagACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-3 CCTAGGT	AATGATACGGCGACCACCGAGATCTACACacctaggACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-4 GGATCAA	AATGATACGGCGACCACCGAGATCTACACttgatccACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-5 GCAAGAT	AATGATACGGCGACCACCGAGATCTACACatcttgcACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-6 ATGGAGA	${\tt AATGATACGGCGACCACCGAGATCTACACtctccatACACTCTTTCCCTACACGACGCTCTT}$
P5_iPCR-LP-7 CTCGATG	${\tt AATGATACGGCGACCACCGAGATCTACACcatcgagACACTCTTTCCCTACACGACGCTCTT}$
P5_iPCR-LP-8 GCTCGAA	${\tt AATGATACGGCGACCACCGAGATCTACACttcgagcACACTCTTTCCCTACACGACGCTCTT}$
P5_iPCR-LP-9 ACCAACT	AATGATACGGCGACCACCGAGATCTACACagttggtACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-10 CCGGTAC	${\tt AATGATACGGCGACCACCGAGATCTACACgtaccggACACTCTTTCCCTACACGACGCTCTT}$
P5_iPCR-LP-11 AACTCCG	AATGATACGGCGACCACCGAGATCTACACcggagttACACTCTTTCCCTACACGACGCTCTT
P5_iPCR-LP-12 TTGAAGT	${\tt AATGATACGGCGACCACCGAGATCTACACacttcaaACACTCTTTCCCTACACGACGCTCTT}$
P5_iPCR-LP-13 ACTATCA	${\tt AATGATACGGCGACCACCGAGATCTACACtgatagtACACTCTTTCCCTACACGACGCTCTT}$
IS4 AGATCTC	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT
amplification primers for i	ndexed libraries
IS5	AATGATACGGCGACCACCGA
IS6	CAAGCAGAAGACGGCATACGA
sequencing primer	
P5 index sequencing	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

1	Experiment MP	-CAP	Experiments no-CAP and SP-CAP						
Sample	Index 1	Index 2	Sample	Index 1	Index 2	Sample	Index 1	Index 2	
L1	AATCTTC (97)	TCGCAGG (1)	L1	AATCTTC (97)	TCGCAGG (1)	L10	AACCATG (106)	CCGGTAC (10)	
L2	ACCAACG (98)	CTCTGCA (2)	L2	ACCAACG (98)	CTCTGCA (2)	L11	CCAATCC (109)	ACTATCA (13)	
L3	AGATGGC (99)	CCTAGGT (3)	L3	AGATGGC (99)	CCTAGGT (3)	L12	CCGGCGT (110)	TCGCAGG (1)	
L4	CCAGGTT (100)	GGATCAA (4)	L4	CCAGGTT (100)	GGATCAA (4)	L13	CGCATAG (111)	CTCTGCA (2)	
L5	CCGTTAG (101)	GCAAGAT (5)	L5	CCGTTAG (101)	GCAAGAT (5)	L14	CGTAATC (112)	CCTAGGT (3)	
L6	CGCCTCT (102)	ATGGAGA (6)	L6	CGCCTCT (102)	ATGGAGA (6)	L15	CGTTGGT (113)	GGATCAA (4)	
L7	CTTGCGG (103)	CTCGATG (7)	L7	CTTGCGG (103)	CTCGATG (7)	L16	CAGGAAG (107)	AACTCCG (11)	
L8	GGCGGAG (104)	GCTCGAA (8)	L8	GGCGGAG (104)	GCTCGAA (8)	L17	CTATACG (114)	GCAAGAT (5)	
L9	TGGACGT (105)	ACCAACT (9)	L9	TGGACGT (105)	ACCAACT (9)	φX	TTGCCGC (control)	AGATCTC (IS4)	
L10	AACCATG (106)	CCGGTAC (10)							
L16	CAGGAAG (107)	AACTCCG (11)							
37	TTGCCGC	AGATCTC							

Supplementary Table 2. Index combination used for each library

φX

(control)

(IS4)

Supplementary Table 3: Fraction of false pairs observed when using different image analysis and base calling software

The sequencing run was performed close to the release date of a new image analysis version and images were transferred off the instrument. Thus, image analysis of this sequencing run was done once with the Illumina RTA software version 1.6 (on the instrument) and once with Illumina OLB version 1.8. Later, the analysis was repeated using OLB 1.9. Results reported in the manuscript are all based on OLB 1.8, which identified between 18%-25% more clusters for the different lanes than the original instrument software run with RTA1.6/SCS2.6. Both new image analysis software versions increased the fraction of perfect index pairings in these lanes by 1-5% (using the same base caller). This indicates a lower sequencing error due to the improved identification and tracking of cluster positions in the images of the flow cell. However, when comparing the fraction of false index pairs, increased values are observed for the two new versions.

RTA /	Raw		Bust	tard			Ibi	is				
OLB	clusters	Correct p	airs	False	pairs	Correct p	airs	False	pairs			
				no-Ca	р							
1.6	29133246	24116308	83%	97315	0.40%	25091043	86%	126518	0.50%			
1.8	34241955	26765784	78%	120236	0.45%	29950674	87%	175266	0.58%			
1.9	34241993	29962477	88%	170564	0.57%	29873025	87%	177594	0.59%			
	SP-Cap											
1.6	38937256	27155457	70%	125147	0.46%	29042714	75%	143824	0.49%			
1.8	48546372	34564560	71%	159982	0.46%	37690136	78%	192995	0.51%			
1.9	48546297	37889964	78%	193461	0.51%	37426552	77%	202081	0.54%			
	МР-Сар											
1.6	29245166	24246039	83%	154789	0.63%	25187487	86%	179188	0.71%			
1.8	34684183	28316453	82%	192938	0.68%	30261682	87%	230005	0.75%			
1.9	34684220	30525097	88%	233016	0.76%	30159122	87%	230668	0.76%			

Supplementary Table 4: Quantifying false sample assignments in human sequencing data from single-indexed libraries

All forward reads were aligned to the $\phi X174$ genome using BWA. False sample assignments (FSA) are defined as reads showing a sample index but an alignment to the phage genome. QF columns show the changes after applying a minimum quality score filter of 15 to the index read.

HGDP ID	Raw sequences	Algn φX index	Algn Sample index	FSA rate	Algn φX index (QF)	Algn Sample index (QF)	FSA rate (QF)	Kept øX index (QF)	Kept Sample index (QF)
HGDP00456	30562322	755869	1471	0.194%	743437	231	0.031%	98.36%	15.70%
HGDP00998	31188058	483305	409	0.085%	473923	56	0.012%	98.06%	13.69%
HGDP00665	34994524	781560	803	0.103%	768559	58	0.008%	98.34%	7.22%
HGDP00491	37133303	866342	943	0.109%	852285	49	0.006%	98.38%	5.20%
HGDP00711	39665263	758411	876	0.115%	738413	77	0.010%	97.36%	8.79%
HGDP01224	35608002	968626	1238	0.128%	956959	85	0.009%	98.80%	6.87%
HGDP00551	36576315	720162	1583	0.219%	691189	144	0.021%	95.98%	9.10%

Supplementary Table 5: Quantifying false sample assignments in human and mouse transcriptome sequencing data from single-indexed libraries, which were prepared using Illumina's TruSeq RNA Sample Prep Kit

All single reads were aligned to the $\varphi X174$ genome using BWA. False sample assignments are defined as reads showing one of the 12 sample indexes but an alignment to the phage genome. QF columns show the changes after applying a minimum quality score filter of 15 to the index read.

Lane	Raw reads	Algn PhiX index	Ave. algn sample index	False sample assignments	Algn PhiX index QF	Ave. algn sample index QF	False sample assignments QF
1	45990409	563645	906	0.160%	551229	304	0.055%
2	42400537	611873	872	0.142%	600333	307	0.051%
3	44098398	482999	736	0.152%	474908	270	0.057%
4	43447369	597251	882	0.147%	586811	315	0.054%
5	50775674	412854	691	0.167%	404888	242	0.060%

Supplementary Figure 1: Raw intensity values of false index pairs

We extracted the raw intensities from single clusters that were identified with a false index pair in each of the tiles 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, and 120 in the no-CAP experiment. The intensity values for both index reads of these twelve clusters are available in the Supplementary Text. All of them show intensity distributions indicative of non-pure clusters. Here, we show a visualization of the intensity values for tile 1 and 30. Illumina uses four fluorescent dyes to distinguish the four nucleotides A, C, G and T. Of these, two pairs (A/C and G/T) are excited using the same laser and are similar in their emission spectra. They are only partially separated using optical filters. A fluorophores are also measured in the C channel and G fluorophores are also measured in the C channel. Thus, an A can be identified by strong signals in both the A and the C channel, while a C will only show a strong signal in the C channel. The same applies for G/T. Hence, a strong signal in A/C excludes another signal in G/T channels for pure clusters.



Supplementary Figure 2: Changes in the fraction of false index pairs when accepting template reads with an increasing number of bases with quality scores smaller or equal to 15 (no-CAP blue, SP-CAP red, MP-CAP blue).

When considering a cutoff removing a little bit less raw data (dashed lines, right axis) than the Pass Filter flag (~20%; green lines), the applied filter removes slightly more false pairs in no-CAP and MP-CAP than the Pass Filter (PF) flag (sun symbols).



Accepted number of bases with quality below 15

SUPPLEMENTARY METHODS

1. Quantifying false index pairs expected based on sequencing error alone

To obtain an estimate of false pairs from random errors, direct application of the binomial distribution provides an overestimate as usually only a very small proportion of the erroneous variants will match another index. In the presented experiments, at most 18 first indexes and 13 second indexes were used. We also required perfect matches to the index sequences and thus only a specific set of errors will generate a certain other valid index sequence. As a result only 17 and 12 of the 16,383 (4⁷-1) erroneous variants, respectively, will contribute false index readouts. We can correct the estimate from the binomial distribution for this effect:

$$\sum_{x=1}^{7} \frac{d_x}{\binom{7}{x} \cdot 3^x} \cdot \binom{7}{x} p^x (1-p)^{7-x}$$

The number of erroneous variants for a specific number of errors d_x can be inferred from the edit distance matrix of the 18 first index sequences:

[-,	б,	5,	б,	б,	5,	б,	7,	7,	3,	5,	7,	б,	4,	б,	б,	7,	5]
[б,	-,	б,	б,	5,	5,	б,	5,	б,	3,	4,	б,	4,	5,	7,	5,	4,	7]
[5,	б,	-,	5,	б,	б,	б,	5,	5,	б,	5,	б,	б,	5,	3,	7,	5,	5]
[6,	б,	5,	-,	5,	5,	5,	5,	б,	б,	4,	З,	б,	5,	4,	5,	5,	7]
[б,	5,	б,	5,	-,	5,	5,	5,	б,	б,	4,	4,	3,	б,	5,	3,	4,	6]
[5,	5,	б,	5,	5,	-,	б,	5,	5,	5,	4,	5,	3,	5,	4,	б,	5,	6]
[б,	б,	б,	5,	5,	б,	-,	5,	5,	б,	б,	3,	5,	5,	4,	4,	4,	4]
[7,	5,	5,	5,	5,	5,	5,	-,	б,	5,	7,	б,	З,	б,	5,	4,	б,	7]
[7,	б,	5,	б,	б,	5,	5,	б,	-,	7,	б,	3,	5,	5,	4,	б,	7,	3]
[3,	3,	б,	б,	б,	5,	б,	5,	7,	-,	7,	7,	5,	5,	7,	4,	5,	6]
[5,	4,	5,	4,	4,	4,	б,	7,	б,	7,	-,	5,	4,	4,	б,	б,	4,	6]
[7,	б,	б,	З,	4,	5,	З,	б,	3,	7,	5,	-,	б,	б,	4,	4,	б,	4]
[б,	4,	б,	б,	3,	3,	5,	3,	5,	5,	4,	б,	-,	4,	5,	4,	5,	7]
[4,	5,	5,	5,	б,	5,	5,	б,	5,	5,	4,	б,	4,	-,	4,	5,	5,	6]
[б,	7,	3,	4,	5,	4,	4,	5,	4,	7,	б,	4,	5,	4,	-,	б,	5,	6]
[б,	5,	7,	5,	3,	б,	4,	4,	б,	4,	б,	4,	4,	5,	б,	-,	4,	6]
[7,	4,	5,	5,	4,	5,	4,	б,	7,	5,	4,	б,	5,	5,	5,	4,	-,	6]
[5,	7,	5,	7,	б,	б,	4,	7,	З,	б,	б,	4,	7,	б,	б,	б,	б,	-]

and the 13 indexes of the second index reads:

 $\begin{bmatrix} -, 4, 5, 4, 6, 7, 7, 5, 4, 7, 7, 5, 6 \\ [4, -, 3, 6, 5, 6, 4, 5, 5, 4, 6, 7, 5] \\ [5, 3, -, 6, 5, 5, 4, 7, 5, 5, 4, 6, 6] \\ [4, 6, 6, -, 6, 6, 7, 6, 6, 6, 7, 3, 7] \\ [6, 5, 5, 6, -, 5, 6, 4, 5, 7, 6, 6, 6] \\ [7, 6, 5, 6, 5, -, 5, 5, 6, 5, 5, 6, 5] \\ [7, 4, 4, 7, 6, 5, -, 6, 5, 3, 4, 7, 5] \\ [5, 5, 7, 6, 4, 5, 6, -, 4, 7, 7, 7, 5] \\ [4, 5, 5, 6, 5, 6, 5, 4, -, 6, 5, 3, 4, 6] \\ [7, 4, 5, 6, 7, 5, 3, 7, 6, -, 3, 4, 6] \\ [7, 6, 4, 7, 6, 5, 4, 7, 5, 3, -, 4, 5] \\ [5, 7, 6, 3, 6, 6, 7, 7, 5, 4, 4, -, 7] \\ [6, 5, 6, 7, 4, 5, 5, 5, 7, 6, 5, 7, -] \\ \end{bmatrix}$

The matrix of 18 indexes indicates that on average 1.2 out of the 945 3-substitution variants, 3.1 out of the 2835 4-substitution variants, 5.6 out of the 5103 5-substitution, 5.3 out of the 5103 6-substitution variants and 1.8 out of 2187 7-substitution variants generate a valid other index of the forward index set. This corresponds to a *d* vector of (0, 0, 0, 1.2, 3.1, 5.6, 5.3, 1.8). For the second index set of 13 indexes, the *d* vector from the second matrix is (0, 0, 0, 0.6, 2.0, 3.7, 3.2, 2.5).

To apply the above binomial model, we also require an estimate of the average error rate for the three lanes. Such estimate can be obtained from the weighted average of error rates corresponding to the base quality scores $(10^{QS/-10})$ of the raw reads. This way we obtain estimates of 1.022% raw sequencing error in no-CAP, 1.435% error in SP-CAP and 1.059% error in MP-CAP. The higher error rate in SP-CAP is due to the higher loading density of this lane (see Supplementary Figure S3). Considering the maximum error rate of 1.435% for *p* in the above equation results in 1.26E-07 false index read outs for the first index set and 6.30E-08 for the second index set.

Supplementary Figure S3: Distribution of base quality scores in the index reads



Adding the two rates for the forward and revere set gives an upper estimate for false pairs of 1.89E-07. This is an upper estimate as indexes are not independent in processing and both have to present a known index for the read to be considered. Direct application of the binomial model for a 1.435% error rate and without correcting for the number of valid erroneous index readouts would yield a rate of 9.90E-05 for single indexes and 1.98E-04 for index pairs.

2. Analyzing intensity values from a random set of clusters with false index pairs

We extracted the raw intensities from a single cluster that was identified with a false index pair in each of the tiles 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, and 120 in the no-CAP experiment. The intensity values for both index reads of these twelve clusters are available in the table below. Illumina's software reports intensities as intensity minus surrounding noise. Thus, if a large signal is observed around the cluster, a negative intensity can be reported. Illumina uses four fluorescent dyes to distinguish the four nucleotides A, C, G and T. Of these, two pairs (A/C and G/T) are excited using the same laser and are similar in their emission spectra. They are only partially separated using optical filters. A fluorophores are also measured in the C channel and G fluorophores are

also measured in the T channel. Thus, an A can be identified by strong signals in both the A and the C channel, while a C will only show a strong signal in the C channel. The same applies for G/T. Hence, a strong signal in A/C excludes another signal in G/T channels for pure clusters. Even though more complex intensity distributions can also hint to non-pure clusters, we simply marked the cycles where intensities of at least 200 are observed in both the A/C and G/T channels. We found at least one such observation in each of the two twelve clusters.

$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
7 111 279 1849 982 7 203 233 388 603 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:30:7501:7880 103,10 CTTGCGG,CCGGTAC Index1 A C G T Index2 A C G T 1 703 1466 139 107 1 161 731 80 49 2 777 676 84 1121 2 138 570 46 221 3 73 546 115 1209 3 70 291 940 690 4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 G
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
SOLEXA GAO2_0006_PE1_SR_DH_MW.4.30.750117880 103,10 Clicked, CCGGIAC Index1 A C G T Index2 A C G T 1 703 1466 139 107 1 161 731 80 49 2 777 676 84 1121 2 138 570 46 221 3 73 546 115 1209 3 70 291 940 690 4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW.4:40:7680:4448 109,9
1 703 1466 139 107 1 161 731 80 49 2 777 676 84 1121 2 138 570 46 221 3 73 546 115 1209 3 70 291 940 690 4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC,ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 <td< td=""></td<>
2 777 676 84 1121 2 138 570 46 221 3 73 546 115 1209 3 70 291 940 690 4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC, ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 <
3 73 546 115 1209 3 70 291 940 690 4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC, ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514
4 40 602 2108 994 4 37 238 1383 742 5 735 1480 222 175 5 268 386 450 561 6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC, ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 <td< td=""></td<>
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
6 82 143 2113 1402 6 719 654 469 451 7 89 119 2694 1264 7 238 586 656 590 SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC,ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 12
The second state The second state <th< td=""></th<>
SOLEXA-GA02_0006_PEi_SR_DH_MW:4:40:7680:4448 109,9 CCAATCC, ACCAACT Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$
Index1 A C G T Index2 A C G T 1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
1 678 1180 86 224 1 462 391 8 254 2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
2 593 1245 341 204 2 303 666 48 -1 3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
3 1163 1090 391 557 3 70 370 299 405 4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
4 1514 1557 114 2 4 530 780 0 96 5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
5 72 218 92 1370 5 721 647 52 369 6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
6 196 838 392 585 6 196 387 605 398 7 299 1214 60 156 7 161 258 503 570
7 299 1214 60 156 7 161 258 503 570
SULEXA-GAUZ_UUU6_PE1_SR_DH_MW:4:50:9834:15557 109,9 CCAATCC,ACCAACT
INCELL A C G I INCEAZ A C G I
1 55 227 63 59 1 1851 1757 50 62
1 55 227 63 59 1 1851 1757 50 63 2 119 326 38 66 2 219 1211 29 29
1 55 227 63 59 1 1851 1757 50 63 2 119 326 38 66 2 219 1211 29 29 3 305 392 76 58 3 238 1121 11 297
1 55 227 63 59 1 1851 1757 50 63 2 119 326 38 66 2 219 1211 29 29 3 305 392 76 58 3 238 1121 -11 227 4 289 437 48 10 4 1618 1595 39 64
$\begin{array}{cccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

COLEXY C	NO2 0006	סדן כס		1.60.056	6.5016	102 0	രന്നററ		አ አ ርሞ
JULEAA-G	AUZ_0000	_PEI_SK	_DR_MW·	4.00.950	Tradica 0	103,9	CIIGC	GG, ACCI	MACI
Indexi	A	C	G	T	Index2	A	C	G	T
1	393	1273	73	430	1	563	644	413	189
2	67	306	686	1101	2	68	433	147	130
3	459	445	782	1279	3	268	567	62	157
4	413	451	2282	1239	4	476	648	141	162
5	198	1218	824	425	5	450	724	329	219
6	99	182	2793	1349	6	277	615	331	199
7	100	231	1856	1217	7	268	294	250	665
/	100	231	1000	121/	1	200	294	200	005
SOLEXA-G.	A02_0006	_PE1_SR_	_DH_MW:	4:70:719	1:7553	103,11	CTTGC	GG, AAC	TCCG
Indexl	A	C	G	Т	Index2	A	С	G	Т
1	88	1148	71	66	1	664	568	61	518
2	693	683	-173	723	2	788	924	69	96
3	-111	-52	1081	1315	3	152	582	687	305
4	26	14	2614	1080	4	327	356	252	731
-	674	1200	2011	1000	-	1 4 1	550	202	110
5	674	1206	1400	54	5	141	520	- 3 3	410
6	588	566	1482	663	6	160	528	647	395
7	75	16	2407	987	7	12	35	2055	924
SOLEXA-G	A02_0006	_PEi_SR_	_DH_MW:	4:80:640	6:10544	107,9	CAGGA	AG, ACC	AACT
Index1	A	С	G	Т	Index2	A	С	G	Т
1	78	299	98	80	1	2111	1864	92	95
- 2	288	414	9.0 Q.4	27	2	557	1408	25	_9
2	500	414	660	27	2	227	1201	20	110
3	51	54	668	318	3	284	1391	16	112
4	41	57	594	247	4	1428	1442	74	280
5	268	298	92	102	5	1612	1545	42	86
6	353	364	129	94	6	261	1147	28	50
7	72	110	610	257	7	51	58	575	1572
	. 2	110	010	207		51	50	575	10/1
COLEXA	102 000¢		DII MU.	4.00.746	1.1100	107 0	01001	70 700	ላ ላ ጣ
SULEAA-G	AU2_0000	_PEI_SR	_DH_MW ·	4.90.740	- 1 0	107,9	CAGGA	AG, ACC	AACI
Indexi	A	C	G	Л.	Index2	A	C	G	Л.
1	114	431	49	54	1	1594	1421	-7	20
2	580	494	- 8	80	2	463	1050	61	48
2									
3	123	98	699	226	3	190	947	-5	197
3	123 121	98 146	699 653	226 241	3 4	190 843	947 776	-5 49	197 315
2 3 4 5	123 121 450	98 146 349	699 653 44	226 241 129	3 4 5	190 843 868	947 776	-5 49 4	197 315 58
2 3 4 5	123 121 450	98 146 349	699 653 44	226 241 129	3 4 5	190 843 868	947 776 1042	-5 49 4	197 315 58
2 3 4 5 6	123 121 450 425	98 146 349 447	699 653 44 72	226 241 129 27	3 4 5 6	190 843 868 160	947 776 1042 818	-5 49 4 16	197 315 58 47
2 3 4 5 6 7	123 121 450 425 87	98 146 349 447 81	699 653 44 72 504	226 241 129 27 261	3 4 5 6 7	190 843 868 160 19	947 776 1042 818 11	-5 49 16 615	197 315 58 47 1119
3 4 5 6 7	123 121 450 425 87	98 146 349 447 81	699 653 44 72 504	226 241 129 27 261	3 4 5 6 7	190 843 868 160 19	947 776 1042 818 11	-5 49 4 16 615	197 315 58 47 1119
3 4 5 6 7 SOLEXA-G	123 121 450 425 87 A02_0006	98 146 349 447 81 _PEi_SR	699 653 44 72 504 _DH_MW:	226 241 129 27 261 4:100:56	3 4 5 6 7	190 843 868 160 19	947 776 1042 818 11 TGGAC	-5 49 4 16 615 GT,ACT	197 315 58 47 1119 ATCA
3 4 5 6 7 SOLEXA-G. Index1	123 121 450 425 87 A02_0006 A	98 146 349 447 81 _PEi_SR_ C	699 653 44 72 504 _DH_MW: G	226 241 129 27 261 4:100:56 T	3 4 5 6 7 584:7314 Index2	190 843 868 160 19 105,13 A	947 776 1042 818 11 TGGAC C	-5 49 4 16 615 <u>GT, ACT</u>	197 315 58 47 1119 ATCA T
3 4 5 6 7 SOLEXA-G. Index1	123 121 450 425 87 A02_0006 A -1	98 146 349 447 81 _PEi_SR_ C -22	699 653 44 72 504 	226 241 129 27 261 4:100:56 T 559	3 4 5 6 7 584:7314 Index2 1	190 843 868 160 19 105,13 A 1567	947 776 1042 818 11 TGGAC C 1600	-5 49 4 16 615 GT,ACTZ G 111	197 315 58 47 1119 ATCA T 88
3 4 5 6 7 SOLEXA-G. Index1 1 2	123 121 450 425 87 A02_0006 A -1 -113	98 146 349 447 81 <u>PEi_SR</u> C -22 31	699 653 44 72 504 	226 241 129 27 261 <u>4:100:56</u> T 559 436	3 4 5 6 7 584:7314 Index2 1 2	190 843 868 160 19 105,13 A 1567 175	947 776 1042 818 11 TGGAC C 1600 1074	-5 49 4 16 615 GT,ACTZ G 111 79	197 315 58 47 1119 ATCA T 88 22
3 4 5 6 7 SOLEXA-G. Index1 1 2 3	123 121 450 425 87 A02_0006 A -1 -113 37	98 146 349 447 81 	699 653 44 72 504 	226 241 129 27 261	3 4 5 6 7 	190 843 868 160 19 105,13 A 1567 175 142	947 776 1042 818 11 TGGAC C 1600 1074 440	-5 49 4 16 615 <u>GT, ACTZ</u> G 111 79 -20	197 315 58 47 1119 ATCA T 88 22 999
3 4 5 6 7 SOLEXA-G. Index1 1 2 3	123 121 450 425 87 A02_0006 A -1 -113 37 E 62	98 146 349 447 81 	699 653 44 72 504 	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4	190 843 868 160 19 105,13 A 1567 175 142	947 776 1042 818 11 TGGAC C 1600 1074 440	-5 49 4 16 615 GT, ACTZ G 111 79 -20 22	197 315 58 47 1119 ATCA T 88 22 999 20
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4	123 121 450 425 87 A02_0006 A -1 -113 37 562 ~~	98 146 349 447 81 <u>PEi_SR</u> C -22 31 39 590	699 653 44 72 504 <u>DH_MW:</u> G 62 835 958 24 24	226 241 129 27 261 4:100:56 T 559 436 347 24	3 4 5 6 7 1 1 1 2 3 4 -	190 843 868 160 19 105,13 A 1567 175 142 1557	947 776 1042 818 11 TGGAC C 1600 1074 440 1524	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33	197 315 58 47 1119 ATCA T 88 22 999 -38
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5	123 121 450 425 87 A02_0006 A -1 -113 37 562 25	98 146 349 447 81 	699 653 44 72 504 DH_MW: G 62 835 958 24 117	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5	190 843 868 160 19 105,13 A 1567 175 142 1557 517	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519	-5 49 4 16 615 GT,ACTZ G 111 79 -20 33 86	197 315 58 47 1119 ATCA T 88 22 999 -38 1015
3 4 5 6 7 SOLEXA-G Index1 1 2 3 4 5 6	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66	98 146 349 447 81 	699 653 44 72 504 	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076	-5 49 4 16 615 GT,ACTZ G 111 79 -20 33 86 30	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47	98 146 349 447 81 	699 653 44 72 504 <u>DH_MW:</u> 6 6 2 835 958 24 117 1187 22	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066	-5 49 4 16 615 G 111 79 -20 33 86 30 107	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7	$ \begin{array}{r} 123 \\ 121 \\ 450 \\ 425 \\ 87 \\ \end{array} $ A02_0006 A -1 -113 37 562 \\ 25 \\ -66 \\ 47 \\ \end{array}	98 146 349 447 81 <u>PEi_SR</u> C -22 31 39 590 445 -57 93	699 653 44 72 504 <u>DH_MW:</u> 6 62 835 958 24 117 1187 22	226 241 129 27 261	3 4 5 6 7 	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066	-5 49 4 16 615 GT,ACTI G 111 79 -20 33 86 300 107	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G.	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006	98 146 349 447 81 <u>PEi_SR</u> C -22 31 39 590 445 -57 93 _PEi_SR	699 653 44 72 504 <u>_DH_MW:</u> 6 62 835 958 24 117 1187 22 DH_MW:	226 241 129 27 261 261	3 4 5 6 7 1 1 2 3 4 5 6 7 2 3:5197	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 GGCGG	-5 49 4 16 615 GT,ACTZ G 111 79 -20 33 86 30 107 AG,ACTZ	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A	98 146 349 447 81 	699 653 44 72 504 <u>DH_MW:</u> 6 2 835 958 24 117 1187 22 <u>DH_MW:</u> G	226 241 129 27 261 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 .23:5197 Index2	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 066 GGCGG C	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G	197 315 58 47 1119 ATCA T 88 22 999 -38 22 999 -38 1015 59 382 ATCA T
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39	98 146 349 447 81 	699 653 44 72 504 <u>DH_MW</u> : 6 22 835 958 24 117 1187 22 <u>DH_MW</u> : 6 866	226 241 129 27 261 4:100:56 7 559 436 347 24 -9 531 196 4:110:61 7 7	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 23:5197 Index2 1	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 GGCGG C 1500	-5 49 4 16 615 GT,ACTZ G 111 79 -20 33 86 30 107 AG,ACTZ G 575	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 EE	98 146 349 447 81 	699 653 44 72 504 	226 241 129 27 261 4:100:56 T 559 436 347 24 -9 531 196 4:110:61 T 392 200	3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 GGCGG C C 1500 1428	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33 86 30 107 AG,ACT2 G 575 26	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 382 ATCA T
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 SOLEXA-G.	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55	98 146 349 447 81 <u>PEi_SR</u> C -22 31 39 590 445 -57 93 <u>PEi_SR</u> C -7 87 C	699 653 44 72 504 6 6 835 958 24 117 1187 22 _DH_MW: G 866 718	226 241 129 27 261	3 4 5 6 7 1 1 2 3 4 5 6 7 23:5197 Index2 1 2 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C C 1500 1439	-5 49 4 16 615 GT,ACT G 111 79 -20 33 86 30 107 AG,ACT G 575 36	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94	98 146 349 447 81 -22 31 39 590 445 -57 93 -7 87 217	699 653 44 72 504 <u>DH_MW:</u> 6 2835 958 24 117 1187 22 <u>DH_MW:</u> 6 866 718 68	226 241 129 27 261 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 .23:5197 Index2 1 2 3	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 GGCGG C 1500 1439 31	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G 575 36 104	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 4 5 6 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43	98 146 349 447 81 _PEi_SR -22 31 39 590 445 -57 93 _PEi_SR C -7 87 217 18	699 653 44 72 504 DH_MW: G 62 835 958 24 117 1187 22 DH_MW: G 866 718 68 675	226 241 129 27 261 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 23:5197 Index2 Index2 1 2 3 4 3 4	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 0 GGCGG C 1500 1439 31 1457	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACT; G 575 36 104 16	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 382 ATCA T 341 -14 1642 31
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G Index1 1 2 3 4 5 5 1 ndex1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83	98 146 349 447 81 	699 653 44 72 504 <u>DH_MW:</u> 6 22 335 958 24 117 1187 22 <u>DH_MW:</u> 6 866 718 68 675 624	226 241 129 27 261 4:100:56 7 559 436 347 24 -9 531 196 4:110:61 T 392 309 48 324 318	3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33 86 30 107 AG,ACT2 G 575 36 104 16 734	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 382 ATCA T 341 -14 1642 31 1610
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. 6 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290	98 146 349 447 81 <u>PEi_SR</u> -22 31 39 590 445 -57 93 <u>PEi_SR</u> C -7 87 217 18 37 279	699 653 44 72 504 6 6 835 958 24 117 1187 22 _DH_MW: G 866 718 68 675 624 236	226 241 129 27 261	3 4 5 6 7 284:7314 Index2 1 2 3 4 5 6 7 23:5197 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 1487 77 498	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315	-5 49 4 16 615 GT,ACTI G 111 79 -20 33 86 30 107 AG,ACTI G 575 36 104 16	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98	98 146 349 447 81 <u>PEi_SR</u> C -22 31 39 590 445 -57 93 <u>PEi_SR</u> C -7 87 217 18 37 279 120	699 653 44 72 504 6 6 2835 958 24 117 1187 22 	226 241 129 27 261	3 4 5 6 7 1 1 2 3 4 5 6 7 2 3:5197 1 1 2 3 4 5 6 7 1 2 3 4 5 6 7 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 6 1315 1502	-5 49 4 16 615 GI,ACTZ G 111 79 -20 33 86 30 107 AG,ACTZ G 575 36 104 16 734 16 107	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1642 31 1642 31 1042 31 105 59 38 105 59 38 105 59 38 105 59 38 105 59 38 105 59 38 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 59 105 105 105 105 105 105 105 105
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98	98 146 349 447 81 -22 31 39 590 445 -57 93 -7 87 217 18 37 279 120	699 653 44 72 504 <u>DH_MW:</u> 6 2835 958 24 117 1187 22 <u>DH_MW:</u> 6 866 718 686 718 688 675 624 236 564	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 .23:5197 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G 575 36 104 16 734 16 107	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G.	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006	98 146 349 447 81 _PEi_SR C -22 31 39 590 445 -57 93 _PEi_SR C -7 87 217 18 37 279 120 PEi_SP	699 653 44 72 504 <u>DH_MW:</u> 6 2835 958 24 117 1187 22 <u>DH_MW:</u> 6 866 718 68 675 624 236 564 DH_MW:	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 Index2 Index2 Index2 1 2 3 4 5 6 7 7 321:17030	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 0 GGCGG C 1500 1439 31 1457 86 1315 1502	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33 86 30 107 AG,ACT2 G 575 36 104 16 734 16 107 TG,ACC	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. 1 2 3 4 5 6 7 SOLEXA-G.	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006	98 146 349 447 81 PEi_SR C -22 31 39 590 445 -57 93 PEi_SR C -7 87 217 18 37 279 120 PEi_SR	699 653 44 72 504 <u>DH_MW</u> : <u>G</u> 6 22 355 958 24 117 1187 22 <u>DH_MW</u> : <u>G</u> 866 718 68 675 624 236 564 <u>236</u> 564	226 241 129 27 261 4:100:56 7 559 436 347 24 -9 531 196 4:110:61 7 392 309 48 324 318 120 203 4:120:48	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 1 23:5197 Index2 1 2 3 4 5 6 7 2 3:4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7 7 1 2 3 4 5 6 7 1 2 3 3 4 5 6 7 7 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 1 2 3 3 4 5 5 6 7 7 1 1 1 2 3 3 4 5 5 6 7 7 1 1 1 2 3 3 4 5 5 6 7 7 1 1 1 2 3 3 4 5 5 6 7 7 1 1 1 2 1 1 2 3 1 1 1 1 2 3 1 1 1 1 1 2 3 1 1 1 1	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551 106,9 2	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33 86 30 107 AG,ACT2 G 575 36 104 16 734 16 107 TG,ACC2	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G Index1 1 2 3 4 5 6 7 SOLEXA-G Index1	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A	98 146 349 447 81 PEi_SR -22 31 39 590 445 -57 93 PEi_SR C -7 87 217 18 37 217 18 37 279 120 PEi_SR	699 653 44 72 504 6 6 835 958 24 117 1187 22 	226 241 129 27 261	3 4 5 6 7 1 1 2 3 4 5 6 7 2 2 3:5197 1 1 2 3 4 5 6 7 2 2 3:5197 1 1 2 3 4 5 6 7 2 3 1 2 3 4 5 6 7 7 2 3 4 5 6 7 7 2 3 4 5 6 7 7 7 7 7 8 4:514 7 1 2 3 4 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 104,13 A 104,13 A 1756 267 37 1487 77 498 1551 106,9 A	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C	-5 49 4 16 615 GT,ACTI G 111 79 -20 33 86 30 107 AG,ACTI G 575 36 104 16 107 TG,ACCI G	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 SOLEXA-G. Index1 1 1	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277	98 146 349 447 81 -22 31 39 590 445 -57 93 -77 87 217 18 37 279 120 -PEi_SR C -22 31 39 590 445 -57 93	699 653 44 72 504 6 6 2835 958 24 117 1187 22 <u>DH_MW:</u> G 866 718 68 675 624 236 564 <u>236</u> 564 <u>236</u> 564	226 241 129 27 261	3 4 5 6 7 .284:7314 Index2 1 2 3 4 5 6 7 .23:5197 Index2 1 2 3 4 5 6 7 .23:5197 Index2 1 2 3 4 5 6 7 .23:5197 Index2 1 2 3 4 5 6 7 .23:5197 Index2 Index3 In	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551 106,9 A 1325	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G 575 36 104 16 734 16 107 TG, ACCZ G 216	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1642 31 1642 31 1642 31 103 AACT T 90
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277 394	98 146 349 447 81 C -22 31 39 590 445 -57 93 <u>PEi_SR</u> C -7 87 217 18 37 279 120 <u>PEi_SR</u> C <u>9</u> 279 120	699 653 44 72 504 G 62 835 958 24 117 1187 22 <u>_DH_MW</u> : G 866 718 68 675 624 236 564 236 564 <u></u> G 40 -38	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 1 23:5197 Index2 1 2 3 4 5 6 7 321:17030 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 1487 77 1487 551 106,9 A 1325 308	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACT2 G 575 36 104 16 734 16 107 TG, ACC2 G 216 -11	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T 90 14
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. 3 4 5 6 7 1 SOLEXA-G. 3 3 4 5 6 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 6 7 7 1 1 2 3 3 4 5 5 6 7 7 1 1 2 2 3 3 4 5 5 6 7 7 1 1 1 2 2 3 3 4 5 5 6 7 7 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 1 2 2 3 1 1 2 2 3 1 1 1 2 2 3 3 4 1 5 6 7 7 1 1 2 2 3 1 1 1 2 2 3 1 1 2 2 3 1 7 1 1 2 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 1 2 2 3 3 3 4 5 5 5 5 1 1 1 2 2 3 3 4 5 5 5 5 1 2 3 1 1 1 2 2 3 3 1 1 2 2 3 3 1 1 2 2 3 3 1 2 2 3 3 1 2 2 3 3 1 1 2 2 3 3 2 3 1 2 2 3 1 1 2 2 3 3 2 3 1 2 3 3 2 3 1 2 2 3 3 1 2 2 3 1 2 2 3 3 2 3 2	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277 394 -22	98 146 349 447 81 _PEi_SR C -22 31 39 590 445 -57 93 _PEi_SR C -7 87 217 18 37 217 18 37 279 120 _PEi_SR C _7 87 217 18 37 279 120 	699 653 44 72 504 <u>DH_MW</u> : <u>6</u> 6 28 35 958 24 117 1187 22 <u>DH_MW</u> : <u>6</u> 866 718 68 675 624 236 564 <u>236</u> 564 <u>236</u> 564 236 564	$\begin{array}{c} 226\\ 241\\ 129\\ 27\\ 261\\ \hline \\ \\ \hline \\ \\ \hline \\ \\ \\ \hline \\ \\ \\ \\ \\ \\ \\ \\$	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 7 23:5197 Index2 1 2 3 4 5 6 7 7 23:5197 Index2 1 2 3 4 5 6 7 7 321:17030 Index2 2 3 4 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551 106,9 A 1325 308 286	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944 893	-5 49 4 16 615 GT,ACT2 G 111 79 -20 33 86 30 107 AG,ACT2 G 575 36 104 16 734 16 107 TG,ACC2 G 216 -11 414	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T 90 14 285
3 4 5 6 7 Index1 1 2 3 4 5 6 7 SOLEXA-G Index1 1 2 3 4 5 6 7 SOLEXA-G Index1 1 2 3 4 5 6 7 SOLEXA-G 1 2 3 4 5 6 7 1 1 2 3 4 5 5 6 7 1 1 2 3 1 1 2 3 4 5 5 6 7 7 1 1 2 3 1 1 2 3 4 5 6 7 7 1 1 2 3 4 5 6 7 7 1 1 2 3 4 5 6 7 7 1 1 2 3 4 5 6 7 7 1 1 2 3 4 5 6 7 7 1 1 2 3 4 5 5 6 7 7 1 1 2 3 3 4 5 7 7 1 1 2 3 3 4 5 5 6 7 7 1 1 2 3 3 4 5 5 6 7 7 1 1 1 2 3 3 4 5 5 6 7 7 1 1 2 3 3 4 5 5 6 7 7 1 1 2 3 3 4 5 5 6 7 7 7 7 7 7 1 1 2 3 3 4 5 7 7 7 7 1 1 2 3 3 4 5 5 6 7 7 7 7 1 1 2 3 3 4 5 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277 394 -22 110	98 146 349 447 81 PEi_SR C -22 31 39 590 445 -57 93 PEi_SR C -7 87 217 18 37 279 120 PEi_SR C 361 418 123 276	699 653 44 72 504 6 6 835 958 24 117 1187 22 	226 241 129 27 261	3 4 5 6 7 284:7314 1 2 3 4 5 6 7 23:5197 Index2 1 2 3 4 5 6 7 223:5197 Index2 1 2 3 4 5 6 7 7 223:5197 Index2 1 2 3 4 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 104,13 A 104,13 A 104,13 A 104,13 A 106,9 A 1325 308 286 1019	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944 893 967	-5 49 4 16 615 GT,ACTI G 111 79 -20 33 86 30 107 AG,ACTI G 575 36 104 16 107 TG,ACCI G 216 -11 414 367	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T 90 144 285 383
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. 3 4 5 6 7 SOLEXA-G. 3 4 5 6 7 8 5 8 5 7 8 7 8 7 8 7 8 7 8 7 8 7 8 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -39 55 94 43 -83 290 98 A02_0006 A 277 394 -22 110 323	98 146 349 447 81 -22 31 39 590 445 -57 93 -27 93 -22 31 39 590 445 -57 93 -27 120 -7 87 217 18 37 279 120 -PEi_SR C -361 418 123 276 269	699 653 44 72 504 6 6 2835 958 24 117 1187 22 0 <u>DH_MW:</u> G 866 718 68 675 624 2366 564 2366 2364 2364 2364 2364 2364 2364 23	226 241 129 27 261	3 4 5 6 7 1 1 2 3 4 5 6 7 2 3 3 4 5 6 7 2 3 3 4 5 6 7 2 3 3 4 5 6 7 2 3 1 2 3 4 5 6 7 7 2 3 1 2 3 4 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 104,13 A 104,13 A 104,13 A 10551 106,9 A 1325 308 286 019 874	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944 893 967 748	-5 49 4 16 615 GT,ACTZ G 111 79 -20 33 86 30 107 AG,ACTZ G 575 36 104 16 734 16 107 TG,ACCZ G 216 -11 414 367 9 216 -11	197 315 58 477 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1642 31 1642 31 1642 31 1642 31 103 AACT T 90 14 285 382 263
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277 394 -22 110 323 207	98 146 349 447 81 -22 31 39 590 445 -57 93 <u>PEi_SR</u> C -7 87 217 18 37 279 120 <u>PEi_SR</u> C 361 418 123 276 269 190	699 653 44 72 504 G 62 835 958 24 117 1187 22 DH_MW: G 866 718 68 675 624 236 564 236 564 236 564 236 564 236 564 236 564 236	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 23:5197 Index2 1 2 3 4 5 6 7 321:17030 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 172 517 258 1130 104,13 A 1756 267 37 1487 77 1487 77 1487 1551 106,9 A 1325 308 286 1019 874 365	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944 893 967 748 707	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G 575 36 104 16 734 16 107 C 734 16 107 TG, ACCZ G 216 -11 414 367 434	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T 90 14 285 383 263 81
3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G. Index1 1 2 3 4 5 6 7 SOLEXA-G.	123 121 450 425 87 A02_0006 A -1 -113 37 562 25 -66 47 A02_0006 A -39 55 94 43 -83 290 98 A02_0006 A 277 394 -22 110 323 207 30	98 146 349 447 81 _PEi_SR C -22 31 39 590 445 -57 93 _PEi_SR C C -7 87 217 18 37 279 120 _PEi_SR C C 361 418 123 276 269 190	699 653 44 72 504 G 62 835 958 24 117 1187 22 <u>DH_MW</u> : G 866 718 68 675 624 236 564 236 24 236 26 26 26 26 26 26 26 26 26 26 26 26 26	226 241 129 27 261	3 4 5 6 7 584:7314 Index2 1 2 3 4 5 6 7 7 23:5197 Index2 1 2 3 4 5 6 7 321:17030 Index2 1 2 3 4 5 6 7	190 843 868 160 19 105,13 A 1567 175 142 1557 517 258 1130 104,13 A 1756 267 37 1487 77 498 1551 106,9 A 1325 308 286 1019 874 365 44	947 776 1042 818 11 TGGAC C 1600 1074 440 1524 519 1076 1066 0 GGCGG C 1500 1439 31 1457 86 1315 1502 AACCA C 1181 944 893 967 748 708	-5 49 4 16 615 G 111 79 -20 33 86 30 107 AG, ACTZ G 575 36 104 16 734 16 107 TG, ACCZ G 216 -11 414 367 479 34	197 315 58 47 1119 ATCA T 88 22 999 -38 1015 59 382 ATCA T 341 -14 1642 31 1610 67 103 AACT T 90 14 285 383 263 81 826

3. Reading each index twice from short-insert molecules

We searched the sequence data of the three experiments for short-insert molecules where the complete and error-free adapter sequences with perfectly matching indexes were obtained in both the forward and reverse read. We identified a total of 3,574,203, 1,699,585 and 3,451,555 such clusters in no-CAP, SP-CAP and MP-CAP. For no-CAP, in 411 out of 3,574,203 (11.499E-5) of these observations, the indexes identified in the read out of the first index did not agree with its second read out. In 245 out of 3,574,203 (6.855E-5) the indexes read out from the second index did not agree. We obtained similar rates for SP-CAP (Index1: 209/1,699,585 = 12.297E-5, Index2: 138/1,699,585 = 8.120E-5) and MP-CAP (Index1: 305/3,451,555 = 8.837E-5, Index2: 275/3,451,555 = 7.967E-5). In cases where the actual index reads provided conflicting information on sample origin in the no-CAP experiment (n=311), the index read from the template read produced a valid index pair. Similar results were obtained for the two other experiments (SP-CAP 141 out of 149 observations, MP-CAP 99 out of 109).

4. Quantifying jumping PCR and other effects

We checked the frequency of all putative index pairs, also including (i) index sequences that have not participated in the experiment at all, but were synthesized in the same batch of indexed oligonucleotides, and (ii) indexes that participated in some step of the experiment but did not participate in the final pooling (see Supplementary Figures 4-6 below). In all three experiments we see a clear increase of correct pairs compared to background when filtering the raw data based on the index reads, i.e. excluding all clusters where at least one base in any of the two index reads has a base quality score below 15. The quality filter affects false index pairs stronger than correct pairs.

When comparing row/column medians to the median of the medians for rows/columns (Supplementary Figures 4-6 below), in some rows and columns we see overrepresentation of false index pairs compared to background. Considering the median values assumes that always a minority of index pairs is affected by cross-contamination and that libraries are pooled in equimolar ratio. The row and column median of medians are 26.5/24, 61/60, and 1208.25/1050 for no-CAP, SP-CAP, and MP-CAP, respectively. IS4 represents the index sequence obtained from the P5-adapter of single-indexed libraries (i.e. using primer IS4 for the preparation of single index libraries instead of a second indexed oligonucleotide). This sequence is expected for the $\omega X174$ library, which was spiked into all sequencing lanes. IS4 shows a low reduction of counts when applying the quality filter to all three datasets and has much higher counts compared to the other rows. The IS4 medians are 1667, 2444 and 9055.5 for no-CAP, SP-CAP and MP-CAP respectively. This indicates contamination of the preparation chemicals (e.g. the PCR buffer) with the IS4 oligonucleotide from the Meyer and Kircher protocol (1). From no-CAP and SP-CAP, we estimate 0.11% and 0.25% contamination with IS4 correspondingly. Other examples of putative contamination include the second index 1 in no-CAP (5.8x higher than the median of medians) and the first index 106 in SP-CAP (10.6x higher than the median of medians). The figures clearly show individual pairs, e.g.

11/103 and 11/105 in no-CAP, 97/3 and 10/105 in SP-CAP, and 97/3 in MP-CAP that are overrepresented compared to background.

Indexes that were used in control reactions but not sequenced are seen 2-56 times more frequently in no-CAP than the ones not used in the experiment at all. This suggests handling contamination as one source of error. Assuming that higher frequency false index pairs (frequency five times higher than background) are due to contamination, we estimate that 0.04% of the 0.06% false pairs in no-CAP, 0.10% of the 0.14% false pairs in SP-CAP and 0.04% of the 0.43% false pairs in MP-CAP are due to cross-contamination of indexed oligonucleotides or libraries (Supplementary Tables 6.1-3 on the following pages).

Supplementary Figure 4: Counts of all index pairs before and after applying a minimum base quality score cutoff of 15 to the index reads. Shown are the result for experiment no-CAP, where libraries were amplified independently and pooled just prior to sequencing. The first index is plotted on the vertical axis, the second on the horizontal axis. Indexes with green labels did not participate in the experiment at all. Pink indexes were used during library preparation, but not included in the library pool for sequencing. The control library is identified by the combination of *control* and *IS4* (blue).



Supplementary Figure 5: Results for perfect index pairings in experiment SP-CAP, where all libraries were enriched and amplified individually and just pooled prior to



Supplementary Figure 6: Results for perfect index pairings in experiment MP-CAP, where all libraries were pooled prior to enrichment and amplification.



Supplementary Table 6.1: Index cross-contamination in no-CAP quantified from higher frequency false index pairs (five times above the medians of the row and column medians). The upper part of the table gives the counts for all false index pairs, while in the lower table (next page) the higher frequency values have been replaced by the average of the medians of row/column medians.

97 60 1309 466 118 116 129 27 22 13 24 39 98 5 85 171 22 43 9 130 12 5 49 8	60 22
98 5 85 171 22 43 9 130 12 5 49 8	22
99 15 15 19 61 12 67 33 13 8 10 7	15
100 534 30 33 <u>33</u> 17 17 21 21 8 23 18	21
101 47 57 30 43 11 316 22 38 3 45 8	38
102 16 37 22 20 24 <u>12</u> 16 25 4 10 7	16
103 1705 19 79 40 21 28 6 262 12 1372 10	28
104 1 28 3 20 131 3 2 4 0 0 2	3
105 717 15 62 16 16 16 9 9 <u>2</u> 892 9	16
106 314 96 52 148 67 24 389 51 22 82 17	67
107 303 18 32 10 30 8 12 161 46 3 10	18
109 <u>19</u> <u>91</u> <u>35</u> <u>76</u> <u>22</u> <u>17</u> <u>8</u> <u>22</u> <u>43</u> <u>18</u> <u>49</u>	22
110 12 796 337 17 13 12 4 13 5 24 13	13
111 5 85 31 16 15 11 30 11 6 20 28	16
112 139 75 65 69 30 15 96 41 12 28 15	41
113 259 83 32 20 14 23 37 75 5 39 126	37
114 293 1 439 5 2 3 4 0 0 3 1	3
ColMedian 139 30 52 40 24 15.5 12 24.5 22 5 26 10	
Total count	16049
Total number of correct pairs	26901906
	0.060%
Index 1 2 3 4 5 6 7 8 9 10 11 13	
97 60 22.63 22.65 118 116 22.63 27 22 13 24 39	
98 5 85 22.05 22 43 9 22.05 12 5 49 8	
99 15 15 19 61 12 67 33 13 8 10 7	
100 22,85 30 33 33 1/ 1/ 21 21 8 23 18 101 47 57 20 42 11 22 20 29 2 45 8	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
102 10 5/ 22 20 24 12 10 25 4 10 /	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
104 1 26 3 20 2203 5 2 4 0 0 2 2020 0 105 105 215 215 215 16 16 16 16 0 0 2 2020 0 0 0 2 2020 0	
105 22.05 13 02 10 10 10 7 7 2 22.05 7 106 23 62 06 52 23 63 67 24 23 65 12 2 22 27 7	
100 22.05 90 32 22.05 0/ 24 22.05 31 22 82 1/ 107 22 42 18 22 10 20 8 12 22 46 2	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
111 2 3 75 65 69 30 15 16 17 0 11 0 20 26	
113 12 13 12 13 12 13 12 13 12 13 12 13 12 13 12 13 12 13 12 13 12 13 12 13 13 12 13 13 13 13 13 13 13 13	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	

Total count

Cross-contamination corrected

4777 **0.018%** **Supplementary Table 6.2:** Quantification of index cross-contamination in SP-CAP from higher frequency false index pairs.

Index	1	2	3	4	5	6	7	8	9	10	11	13	RowMd
97		68	2536	277	64	22	76	14	65	150	69	42	68
98	34		87	58	48	18	39	132	126	104	251	59	59
99	12	26		15	25	0	23	4	5	10	0	2	10
100	230	25	25		44	10	429	11	312	78	82	37	44
101	28	78	29	61		7	362	29	55	105	47	13	47
102	9	28	12	8	21		15	4	15	21	8	8	12
103	828	41	119	410	249	16		17	409	95	524	30	119
104	6	19	7	21	88	0	7		9	23	3	10	9
105	359	160	122	348	24	21	269	30		5623	468	17	160
106	2223	780	490	646	474	148	3026	238	9666		583	353	583
107	154	328	84	74	33	16	101	86	65	120		427	86
109	194	193	49	77	35	61	32	27	103	103	456		77
110		80	693	259	4	2	4	6	20	21	10	99	20
111	119		91	15	10	10	4	14	11	79	44	43	15
112	171	128		65	62	34	14	37	74	172	58	47	62
113	247	59	57		15	29	26	18	60	38	48	305	48
114	208	17	509	16		9	3	7	17	24	12	13	16
ColMedian	171	68	87	65	35	16	29	17.5	62.5	87	53	39.5	
Total count Total number	r of corr	ect pairs											42761 30989801
													0.138%
Index	1	2	3	4	5	6	7	8	9	10	11	13	0.138%
Index 97	1	2 68	3 52.9	4 277	5 64	6 22	7 76	8 14	9 65	10 150	11 69	13 42	0.138%
Index 97 98	1 34	2 68	3 52.9 87	4 277 58	5 64 48	6 22 18	7 76 39	8 14 132	9 65 126	10 150 104	11 69 251	13 42 59	0.138%
Index 97 98 99	1 34 12	2 68 26	3 52.9 87	4 277 58 15	5 64 48 25	6 22 18 0	7 76 39 23	8 14 132 4	9 65 126 5	10 150 104 10	11 69 251 0	13 42 59 2	0.138%
Index 97 98 99 100	1 34 12 230	2 68 26 25	3 52.9 87 25	4 277 58 15	5 64 48 25 44	6 22 18 0 10	7 76 39 23 52.9	8 14 132 4 11	9 65 126 5 52.9	10 150 104 10 78	11 69 251 0 82	13 42 59 2 37	0.138%
Index 97 98 99 100 101	1 34 12 230 28	2 68 26 25 78	3 52.9 87 25 29	4 277 58 15 61	5 64 48 25 44	6 22 18 0 10 7	7 76 39 23 52.9 52.9	8 14 132 4 11 29	9 65 126 5 52.9 55	10 150 104 10 78 105	11 69 251 0 82 47	13 42 59 2 37 13	0.138%
Index 97 98 99 100 101 102	1 34 12 230 28 9	2 68 26 25 78 28	3 52.9 87 25 29 12	4 277 58 15 61 8	5 64 48 25 44 21	6 22 18 0 10 7	7 76 39 23 52.9 52.9 15	8 14 132 4 11 29 4	9 65 126 5 52.9 55 15	10 150 104 10 78 105 21	11 69 251 0 82 47 8	13 42 59 2 37 13 8	0.138%
Index 97 98 99 100 101 102 103	1 34 12 230 28 9 52.9	2 68 26 25 78 28 41	3 52.9 87 25 29 12 119	4 277 58 15 61 8 52.9	5 64 48 25 44 21 249	6 22 18 0 10 7 16	7 76 39 23 52.9 52.9 15	8 14 132 4 11 29 4 17	9 65 126 5 52.9 55 15 52.9	10 150 104 10 78 105 21 95	11 69 251 0 82 47 8 52.9	13 42 59 2 37 13 8 30	0.138%
Index 97 98 99 100 101 102 103 104	1 34 12 230 28 9 52.9 6	2 68 26 25 78 28 41 19	3 52.9 87 25 29 12 119 7	4 277 58 15 61 8 52.9 21	5 64 48 25 44 21 249 88	6 22 18 0 10 7 16 0	7 76 39 23 52.9 52.9 52.9 15 7	8 14 132 4 11 29 4 17	9 65 126 5 52.9 55 15 52.9 9	10 150 104 10 78 105 21 95 23	11 69 251 0 82 47 8 52.9 3	13 42 59 2 37 13 8 30 10	0.138%
Index 97 98 99 100 101 102 103 104 105	1 34 12 230 28 9 52.9 6 52.9	2 68 26 25 78 28 41 19 160	3 52.9 87 25 29 12 119 7 122	4 277 58 15 61 8 52.9 21 52.9	5 64 48 25 44 21 249 88 24	6 22 18 0 10 7 10 7	7 76 39 23 52.9 52.9 15 7 7 269	8 14 132 4 11 29 4 17 30	9 65 126 5 52.9 55 15 52.9 9	10 150 104 10 78 105 21 95 23 52 .9	11 69 251 0 82 47 8 52.9 3 52.9	13 42 59 2 37 13 8 30 10 17	0.138%
Index 97 98 99 100 101 102 103 104 105 106	1 34 12 230 28 9 52.9 6 52.9 52.9	2 68 25 78 28 41 19 160 52.9	3 52.9 87 25 29 12 119 7 122 52.9	4 277 58 15 61 8 52.9 21 52.9 52.9	5 64 48 25 44 21 249 88 24 52.9	6 22 18 0 10 7 16 0 21 148	7 76 39 23 52.9 52.9 15 7 269 52.9	8 14 132 4 11 29 4 17 30 238	9 65 126 5 52.9 55 15 52.9 9 52.9	10 150 104 10 78 105 21 95 23 52.9	11 69 251 0 82 47 8 52.9 3 52.9 52.9	13 42 59 2 37 13 8 30 10 17 52.9	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107	1 34 12 230 28 9 52.9 6 52.9 52.9 154	2 68 26 25 78 28 41 19 160 52.9 52.9	3 52.9 87 25 29 12 119 7 122 52.9 84	4 277 58 15 61 8 52.9 21 52.9 52.9 74	5 64 48 25 44 21 249 88 24 52.9 33	6 22 18 0 10 7 16 0 21 148 16	7 76 39 23 52.9 52.9 15 7 269 52.9 101	8 14 132 4 11 29 4 17 30 238 86	9 65 126 5 52.9 55 15 52.9 9 52.9 9	10 150 104 10 78 105 21 95 23 52.9	11 69 251 0 82 47 8 52.9 3 52.9 52.9	13 42 59 2 37 13 8 30 10 17 52.9 52.9	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109	1 34 12 230 28 9 52.9 6 52.9 52.9 154 194	2 68 25 78 28 41 19 160 52.9 52.9 193	3 52.9 87 25 29 12 119 7 122 52.9 84 49	4 277 58 15 61 8 52.9 21 52.9 52.9 74 77	5 64 48 25 44 21 249 88 24 52.9 33 35	6 22 18 0 10 7 16 0 21 148 16 61	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32	8 14 132 4 11 29 4 17 30 238 86 27	9 65 126 5 52.9 55 15 52.9 9 52.9 9 52.9 65 103	10 150 104 10 78 105 21 95 23 52.9 120 103	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9	13 42 59 2 37 13 8 30 10 17 52.9 52.9	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109 110	1 34 12 230 28 9 52.9 6 52.9 52.9 154 194	2 68 25 78 28 41 19 160 52.9 52.9 193 80	3 52.9 87 25 29 12 119 7 122 52.9 84 49 52.9	4 277 58 15 61 8 52.9 21 52.9 52.9 74 77 259	5 64 48 25 44 21 249 88 24 52 .9 33 35 4	6 22 18 0 10 7 16 0 21 148 16 61 2	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32 4	8 14 132 4 11 29 4 17 30 238 86 27 6	9 65 126 5 52.9 55.9 9 52.9 9 52.9 65 103 20	10 150 104 10 78 105 21 95 23 52.9 120 103 21	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9	13 42 59 2 37 13 8 30 10 17 52.9 52.9 99	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109 110	1 34 12 230 28 9 52.9 6 52.9 52.9 154 194 119	2 68 26 25 78 28 41 19 160 52.9 52.9 193 80	3 52.9 87 25 29 12 119 7 122 52.9 84 49 52.9 91	4 277 58 15 61 8 52.9 21 52.9 52.9 74 77 259 15	5 64 48 25 44 21 249 88 24 52 .9 33 35 4 10	6 22 18 0 10 7 16 0 21 148 16 61 2 10	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32 4 4	8 14 132 4 11 29 4 17 30 238 86 27 6 14 	9 65 126 5 52.9 55 52.9 9 52.9 9 52.9 65 103 20 11	10 150 104 10 78 105 21 95 23 52.9 120 103 21 79	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9 10 44	13 42 59 2 37 13 8 30 10 17 52.9 52.9 99 43	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109 110 111 112	1 34 12 230 28 9 52.9 6 52.9 52.9 154 194 119 171 247	2 68 26 25 78 28 41 19 160 52.9 52.9 193 80 128	3 52.9 87 25 29 12 119 7 122 52.9 84 49 52.9 91	4 277 58 15 61 8 52.9 21 52.9 52.9 74 77 259 15 65	5 64 48 25 44 21 249 88 24 52.9 33 35 4 10 62	6 22 18 0 10 7 16 0 21 148 16 61 2 10 34	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32 4 4 4 14	8 14 132 4 11 29 4 17 30 238 86 27 6 14 37 6 14 37 12 12 12 12 12 12 12 12 12 12	9 65 126 5 52.9 55 52.9 9 52.9 9 52.9 65 103 20 11 74	10 150 104 10 78 105 21 95 23 52.9 120 103 21 79 172 22	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9 10 44 58	13 42 59 2 37 13 8 30 10 17 52.9 52.9 99 43 43	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109 110 111 112 113	1 34 12 230 28 9 52.9 6 52.9 154 194 119 171 247 208	2 68 26 25 78 28 41 19 160 52.9 52.9 193 80 128 59	3 52.9 87 25 29 12 119 7 122 52.9 84 49 52.9 91	4 277 58 15 61 8 52.9 21 52.9 74 77 259 15 65	5 64 48 25 44 21 249 88 24 52.9 33 35 4 10 62 15	6 22 18 0 10 7 16 0 21 148 16 61 2 10 34 29	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32 4 4 4 14 26	8 14 132 4 11 29 4 17 30 238 86 27 6 14 37 18 7	9 65 126 5 52.9 55 52.9 9 52.9 65 103 20 11 74 60	10 150 104 10 78 105 21 95 23 52.9 120 103 21 79 172 38	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9 10 44 58 48	13 42 59 2 37 13 8 30 10 17 52.9 52.9 99 43 47 52.9	0.138%
Index 97 98 99 100 101 102 103 104 105 106 107 109 110 111 112 113 114	1 34 12 230 28 9 52.9 6 52.9 152.9 154 194 119 171 247 208	2 68 25 78 28 41 19 160 52.9 52.9 193 80 128 59 17	3 52.9 87 25 29 12 119 7 122 52.9 84 49 52.9 91 57 57 52.9	4 277 58 15 61 8 52.9 21 52.9 52.9 74 77 259 15 65 16	5 64 48 25 44 21 249 88 24 52.9 33 35 4 10 62 15	6 22 18 0 10 7 16 0 21 148 16 61 2 10 34 29 9	7 76 39 23 52.9 52.9 15 7 269 52.9 101 32 4 4 4 14 26 3	8 14 132 4 11 29 4 17 30 238 86 27 6 14 37 18 7	9 65 126 5 52.9 55 52.9 9 52.9 65 103 20 11 74 60 17	10 150 104 10 78 105 21 95 23 52.9 120 103 21 79 172 38 24	11 69 251 0 82 47 8 52.9 3 52.9 52.9 52.9 10 44 58 48 12	13 42 59 2 37 13 8 30 10 17 52.9 52.9 99 43 47 52.9 13	0.138%

Supplementary Table 6.3: Quantification of index cross-contamination in MP-CAP from higher frequency false index pairs.

Index	1	2	3	4	5	6	7	8	9	10	11	RowMd
97		727	6949	1328	1206	1325	819	1885	184	894	1196	1201
98	621		781	1625	1167	965	941	1826	192	1135	1377	1050
99	19	63		52	46	51	30	68	17	60	58	51.5
100	1283	1277	976		2269	1694	1055	1794	290	1092	1931	1280
101	471	1141	363	1761		1211	1707	1596	253	944	1562	1176
102	425	747	345	1480	1284		837	1374	224	786	1037	811.5
103	2933	996	1326	1989	1215	1385		2103	363	833	5659	1355.5
104	327	704	425	951	1344	842	730		212	697	1098	717
105	1012	156	262	219	246	244	174	447		233	2522	245
106	1087	1667	880	1578	1733	1211	1645	1902	285		1891	1611.5
107	1142	1068	652	1916	1467	1626	1054	1976	461	1013		1105
ColMedian	816.5	871.5	716.5	1529	1249.5	1211	889	1810	238.5	863.5	1469.5	

Total count

Total number of correct pairs

118717 27636356 **0.428%**

Index	1	2	3	4	5	6	7	8	9	10	11	
97		727	997	1328	1206	1325	819	1885	184	894	1196	
98	621		781	1625	1167	965	941	1826	192	1135	1377	
99	19	63		52	46	51	30	68	17	60	58	
100	1283	1277	976		2269	1694	1055	1794	290	1092	1931	
101	471	1141	363	1761		1211	1707	1596	253	944	1562	
102	425	747	345	1480	1284		837	1374	224	786	1037	
103	2933	996	1326	1989	1215	1385		2103	363	833	997	
104	327	704	425	951	1344	842	730		212	697	1098	
105	1012	156	262	219	246	244	174	447		233	2522	
106	1087	1667	880	1578	1733	1211	1645	1902	285		1891	
107	1142	1068	652	1916	1467	1626	1054	1976	461	1013		
Total count Cross-contamin	ation correct	ed										108103 0.390%

5. Estimating false-assignment rates based on the occurrence of unused indexes

In standard multiplex sequencing experiments with single-indexed libraries, falseassignment rates can only be estimated by quantifying the occurrence of unused index sequences (2). These unused indexes are expected to appear if indexes are converted into each other due to errors in synthesis, amplification and sequencing or if there is crosscontamination among index PCR primers and indexed libraries. When we restrict our analysis to the forward index read, we determine false-assignment rates of 0.02% in no-CAP, 0.68% in SP-CAP and 0.004% in MP-CAP from five unused first indexes. We also performed the reciprocal analysis and analyzed perfect reverse index reads for which one index primer was not used. Here we obtained false-assignment rates of 0.001% for no-CAP, SP-CAP and MP-CAP, respectively. These rates can be combined to an estimate for false pairs by considering the average number of sequences observed for a first and a second index as well the expected number of sequences observed for a sample:

 $\frac{\overline{\#SeqIndex1} \cdot r_{1} + \overline{\#SeqIndex2} \cdot r_{2}}{\left(\sum \#SeqIndex1 + \sum \#SeqIndex2\right)/(2 \cdot \#Samples)}$

Using this formula, we obtain the joint estimates of:

	no-CAP	SP-CAP	MP-CAP
Unused forward indexes (r1)	0.020%	0.681%	0.004%
Unused reverse index (r ₂)	0.001%	0.001%	0.001%
Number of samples	17	17	11
Average per used first index	2125542	3381013	2888431
Average per used second index	2201444	2458359	2887183
Sum used first indexes	31364391	41792106	31772738
Sum used second indexes	31566651	40572152	31759009
Joint estimate unused indexes	0.024%	0.818%	0.005%

The false-assignment rates estimated from unused indexes is highest in SP-CAP. Since amplification and sequencing errors will only rarely convert one index sequence into another (see main text), these rates must almost exclusively reflect cross-contamination among indexed oligonucleotides or libraries.

5. Identification of false sample assignments in single indexed data

To quantify false index assignments in regular single index libraries, we used the 2x101+7 PE sequencing data from 7 present-day humans presented by Reich et al.(3). From the seven lanes, we separately aligned the forward and reverse read of all raw clusters to the φ X174 reference genome using BWA(4) and identified false index assignments as reads with a φ X alignment showing the designated sample index for the specific library. Supplementary Table 2 provides the results for the forward read with and without applying a minimum quality score filter of 15 to the index read out. In a second experiment of five lanes with mRNA libraries, generated using the Illumina TruSeq RNA Sample Prep Kit and sequenced on a single read run with 76+7 cycles, we obtain on average 0.14% to 0.17% φ X174 contamination for each sample index, which reduces to 0.05-0.06% after applying the index quality filter (Supplementary Table 3).

SUPPLEMENTARY REFERENCES

- Meyer, M. and Kircher, M. (2010) Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb Protoc*, 2010, pdb prot5448.
- 2. Meyer, M., Stenzel, U. and Hofreiter, M. (2008) Parallel tagged sequencing on the 454 platform. *Nat Protoc*, **3**, 267-278.
- 3. Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L. *et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, **468**, 1053-1060.
- 4. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760.