Double Capture: An Alternative Protocol for Sequence Capture of Difficult Targets

For use in SeqCap EZ Library SR Applications

Primary Author:
Dan Burgess, PhD

Contributing Authors:
Jennifer Wendt, Dawn Green, Carl Kashuk, and Michael Brockman, PhD

Roche NimbleGen Inc., Madison, WI

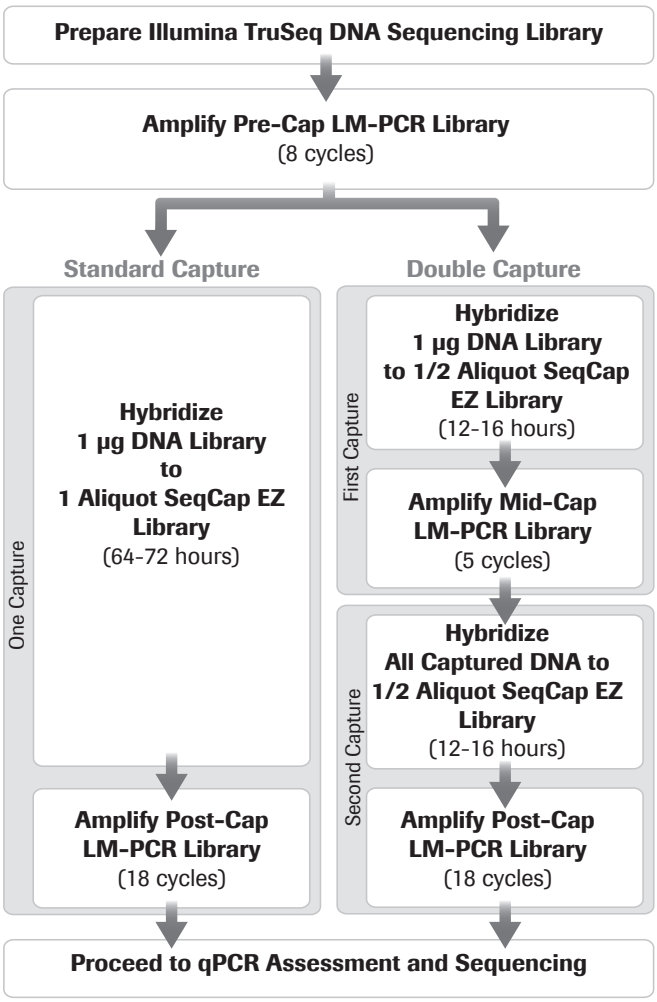
Introduction

In Sequence Capture experiments, enrichment specificity is defined as the fraction of nucleic acid fragments obtained at the end of an experiment that were explicitly targeted for capture at the beginning. This is often measured by proxy as the percentage of sequence reads, or sequenced bases, from an experiment that align with the targeted portion of a reference sequence. Along with sequence coverage uniformity across the target, enrichment specificity is a major determinant of overall process efficiency. For example, an experiment yielding 30% of sequence reads mapping to the intended target (*i.e.*, a 30% on-target rate) would require approximately twice as much sequencing, to get the same amount of useful data, as an experiment with a 60% on-target rate for that same target.

Sequence Capture specificity is determined by several factors, including probe design, hybridization and washing conditions, and repetitive element and library adapter blocking strategies. Even though much effort has gone into optimizing these experimental elements, some targets remain difficult to capture with high specificity. Smaller targets (<100 kb in size) are noteworthy in this regard, frequently exhibiting on-target rates below 50%. Capture data from larger targets (>100 kb) only seldom indicate low on-target rates, but do suggest that sequence structure and genomic context might influence this metric as well. Because it is difficult to predict which individual targets will capture with low specificity, we tested alternative protocols designed to boost capture performance for use when these cases arise. We present one protocol here, referred to as the double capture protocol, that has been markedly effective for increasing performance when working with some difficult-to-capture targets.

Methods

The workflow of the double capture protocol is straightforward. An aliquot of SeqCap EZ library probes sufficient for a single standard capture reaction is divided in half. It is used to perform two sequential captures, where the enriched library output of the first capture is used as input into the second capture. Both hybridization steps are incubated overnight (12-16 hours) so the double capture protocol is completed in less time than the standard capture protocol (Figure 1). Detailed instructions are provided in Appendix 1.



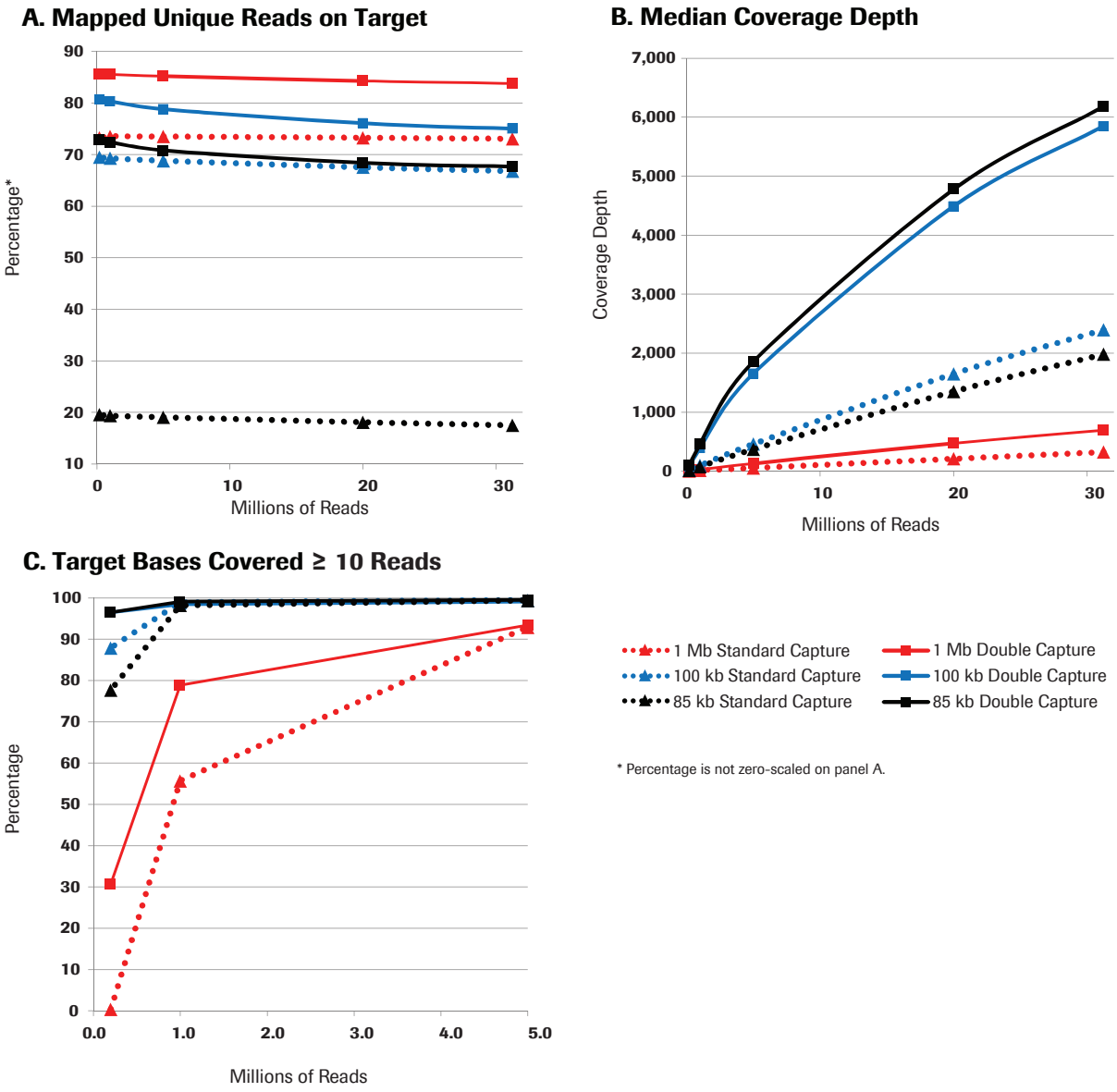
▲ **Figure 1: Capture workflows.** Comparison of the standard capture and double capture workflows for SeqCap EZ Library.

Results

Double Capture from Human Genomic Targets

Three different SeqCap EZ Libraries were produced targeting distinct regions of the human genome with cumulative sizes of 1 Mb (mixed exons and contiguous regions), 100 kb (exons only) and 85 kb (exons only). The probes were used to capture Illumina TruSeq DNA libraries constructed from human genomic DNA (NA12762; Coriell Institute Biorepository), using the standard capture protocol (*NimbleGen SeqCap EZ Library SR User's Guide v3.0*) and the double capture protocol described in this technical note. All experiments were performed in triplicate. Sample libraries with different adapter indexes were captured individually but subsequently pooled for sequencing (Illumina HiSeq 2x76bp). The reads were demultiplexed, randomly sampled to normalize for sequencing output variability between experiments, and mapped. Duplicate read pairs were removed (read pairs with the highest quality scores were retained) and remaining reads were used to determine coverage and call SNPs (single nucleotide polymorphisms). Mapping and SNP calling were performed using the SOAP2 software package (Beijing Genomics Institute). SNPs within the primary target were filtered for a minimum depth of 4X. Heterozygotes were called with MAF (minor allele frequency) thresholds of 15% and 85%. SNP detection rates were determined by comparing called SNPs against previously genotyped SNPs for sample NA12762 (International HapMap Project). The results are shown in Figures 2 - 4. Results reported represent the average of 3 replicate experiments.

The percent reads on-target was improved through use of the double capture protocol for each design tested, but the gains were variable (Figure 2A). The on-target rate for captures of the 1 Mb design increased from 73% (standard capture) to 85% (double capture), captures of the 100 kb design increased from 68% to 78%, and captures of the 85 kb design increased from 19% to 70%. As expected, the increased on-target rates resulted in deeper coverage depths. Median coverage increased 2.2-fold for the 1 Mb design, 2.7-fold for the 100 kb design, and 3.5-fold for the 85 kb design (Figure 2B).

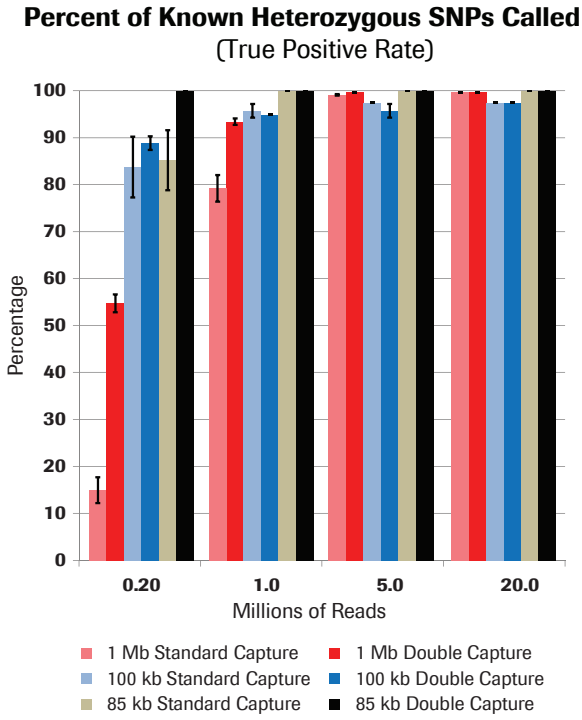


▲ **Figure 2: Performance comparison of the SeqCap EZ Library standard capture and double capture protocols.** Targets of 1 Mb, 100 kb and 85 kb were each captured using both capture methods. All experiments were performed in triplicate, with results reported representing the average of these replicate experiments. Raw sequence reads were randomly down-sampled to 31.2, 20, 5, 1, and 0.2 million reads for analysis and balanced comparison. Graphs shown illustrate A) mapped unique reads on target; B) median coverage depth; C) target bases covered ≥10 reads.

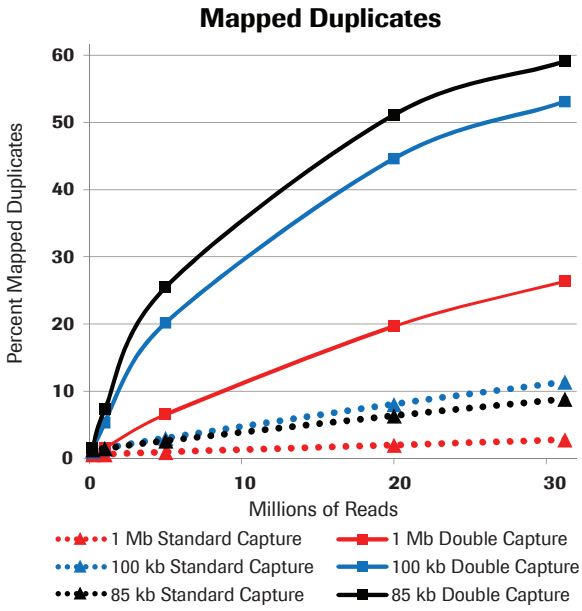
For both protocols, the percentage of targeted bases covered with a depth of ≥10 reads was saturated above 0.2 million raw reads for the 85 kb and 100 kb designs, and above 5 million reads for the 1 Mb design. Below these levels, however, the double capture protocol resulted in more targeted bases covered with at least 10 reads (Figure 2C), suggesting that this method does not sacrifice the capture of difficult targets in exchange for higher on-target rates.

The ability of Sequence Capture to facilitate identification of heterozygous (Het) SNPs with a high true-positive rate is critical for most applications. Since the experiments described here used a HapMap DNA sample (NA12762) that was previously genotyped, we assessed the capacity of the double capture protocol to identify these known SNPs when they were located in the capture target. The standard and double capture protocols performed similarly at identifying Het SNPs in all three designs for most amounts of sequence analyzed (Figure 3). However, when sequence input into the analysis became the limiting factor (*e.g.* at 1 million reads for the 1 Mb design and at 0.2 million reads for the 85 kb design) the double capture protocol was more sensitive.

Since targeted sequence enrichment is a strong selective process, it was important to measure the effect of the new double capture protocol directly on sample complexity. We examined the duplication rate among sequenced library fragments. Duplicates were defined as all mapped read pairs where the start and stop coordinates of both reads were identical to those of another mapped read pair. The double capture protocol resulted in significantly increased duplicate rates relative to the standard protocol (Figure 4). Although high rates of duplication can be a concern in some contexts, where it might indicate loss of sample complexity within the capture target and reduced sensitivity for variant detection, the high duplicate rates here are associated with good sensitivity for SNP detection. The increase in this case may therefore be a sign that the high coverage depth of captured fragments relative to the small target size of these designs already accounts for most of the unique fragments present in the original sample library.



▲ **Figure 3: Performance comparison of standard capture and double capture methods for heterozygous SNPs.** Percent known heterozygous SNPs called (error bars are ±1 standard deviation), which reveals more sensitivity for the double capture protocol.



▲ **Figure 4. Performance comparison of standard and double capture methods for mapped duplicates.** Targets of 1 Mb, 100 kb and 85 kb were each captured using both capture methods. All experiments were performed in triplicate, with results reported representing the average of these replicate experiments.

Conclusions

The double capture protocol is an alternative method that has proven useful when the SeqCap EZ Library standard protocol exhibits reduced performance, such as with very small or difficult to capture targets. Improved results have been obtained in several experiments that have been attempted, but the degree of performance gain has been variable among different designs. Although the double capture protocol consists of two sequential hybridization steps, requiring more hands-on time and reagents, the overall duration is shorter than the standard protocol (two days compared to three) and does not use any additional SeqCap EZ Library.

Appendix 1: Instructions for Double Capture using Illumina TruSeq DNA Libraries with SeqCap EZ Libraries

Description

This abbreviated protocol describes the process for the selective capture of genomic DNA (gDNA) from customer specified regions of the human genome using Illumina TruSeq sample library DNA and NimbleGen SeqCap EZ Library. Specifically, this protocol provides instructions for performing two consecutive single-day hybridizations (*i.e.* double capture) that increase the on-target read rate when used with regions that are difficult to capture or very small (<100 kb). This protocol provides details only where the double capture protocol diverges from the standard capture protocol provided in the *NimbleGen SeqCap EZ Library SR User's Guide v3.0* (or later version) and so both documents are required to complete the entire double capture workflow.

This table guides you through the SeqCap EZ Library double capture protocol. It identifies when to follow instructions in the *User's Guide** and when to follow instructions in this document.

Protocol Step	Refer to the following:
1. Before You Begin	Chapter 1 in the *User's Guide*
2. Storing the SeqCap EZ Library	"Storing the SeqCap EZ Library for Double Capture" section in this technical note, which replaces Chapter 2 in the *User's Guide*
3. Preparing the Sample Library and Performing QA	Chapter 3 in the *User's Guide*
4. Amplifying the Sample Library Using LM-PCR	Chapter 4 in the *User's Guide*
5. Hybridizing the Sample and SeqCap EZ Libraries	"Hybridizing the Samples and SeqCap EZ Libraries - First Capture and Second Capture" sections in this technical note, which replaces Chapter 5 in the *User's Guide*
6. Washing and Recovery of Captured DNA	Chapter 6 in the *User's Guide*
7. Amplifying Captured DNA Using LM-PCR	Chapter 7 in the *User's Guide*
8. Measuring Enrichment Using qPCR	Chapter 8 in the *User's Guide*
**** User's Guide = NimbleGen SeqCap EZ Library SR User's Guide v3.0 (or later)***	

Storing the SeqCap EZ Library for Double Capture

Important Note: This section replaces Chapter 2 – Storing the SeqCap EZ Library from the *NimbleGen SeqCap EZ Library SR User's Guide v3.0* (or later).

1. If frozen, thaw the SeqCap EZ Library on ice.
2. Vortex the SeqCap EZ library for 3 seconds.
3. Centrifuge the tube of SeqCap EZ Library at 10,000 x g for 30 seconds to ensure that the liquid is at the bottom of the tube before opening the tube.
4. Aliquot the SeqCap EZ Library into single-use aliquots (2.25 µl/aliquot) in 0.2ml PCR tubes. Store at -15 to -25°C until use.
5. When ready to perform the experiment, thaw the required number of single-use SeqCap EZ Library aliquots on ice.

Notes: *The SeqCap EZ Library should not undergo multiple freeze/thaw cycles. To help ensure the highest performance of the SeqCap EZ Library, Roche NimbleGen recommends aliquoting the SeqCap EZ Library into single-use volumes to prevent damage from successive freeze/thaw cycles.*

Hybridizing the Samples and SeqCap EZ Libraries – First Capture

Important Note: This section replaces Chapter 5 – Hybridizing the Sample and SeqCap EZ Libraries of the *NimbleGen SeqCap EZ Library SR User's Guide v3.0* (or later).

Step 1. Preparing for the First Hybridization

1. Turn on a heat block to +95°C and let it equilibrate to the set temperature.
2. Remove an aliquot of (2.25 µl) SeqCap EZ Library from the freezer (-15 to -25°C) and allow it to thaw on ice. Then add 2.25 µl of PCR grade water for a total of 4.5 µl solution.

Step 2. Prepare the First Hybridization Sample

1. Add 5 µl of 1 mg/ml COT DNA and 1 µg of DNA Sample Library to a new 1.5 ml tube.
2. Add 1 µl of 1,000 µM TS-INV-HE Universal Oligo 1 and 1 µl of the appropriate TS-INV-HE Index Oligo (1,000 µM) to the amplified sample library plus COT DNA.
3. Close the tube's lid and make a hole in the top of the tube's cap with an 18 - 20 gauge or smaller needle. The closed lid with a hole in the top of the tube's cap is a precaution to suppress contamination in the DNA vacuum concentrator.
4. Dry the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool in a DNA vacuum concentrator on high heat (60°C). Denaturation of the DNA with high heat is not problematic after linker ligation because the hybridization utilizes single-stranded DNA.
5. To each dried-down the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool, add:

   - 7.5 µl of 2X Hybridization Buffer (vial 5)
   - 3 µl of Hybridization Component A (vial 6)

   The tube with the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool should now contain the following components:

Component	Solution Capture
COT DNA	5 µg
Amplified DNA Sample Library	1 µg
1,000 µM TS-HE Universal Oligo 1	1,000 pmol
1,000 µM TS-INV-HE Index Oligo	1,000 pmol
2X Hybridization Buffer (vial 5)	7.5 µl
Hybridization Component A (vial 6)	3 µl
Total	10.5 µl

6. Cover the hole in the tube's cap with a sticker or small piece of laboratory tape.
7. Vortex the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool plus Hybridization Cocktail (2X Hybridization Buffer + Hybridization Component A) for 10 seconds and centrifuge at maximum speed for 10 seconds.
8. Place the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail in a 95°C heat block for 10 minutes to denature the DNA.
9. Centrifuge the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail at maximum speed for 10 seconds at room temperature.
10. Transfer the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail to the 4.5 µl aliquot of SeqCap EZ Library in a 0.2 ml PCR tube prepared in Chapter 2 of the *NimbleGen SeqCap EZ Library SR User's Guide v3.0*.

The hybridization sample should now contain the following components:

Component	Solution Capture
COT DNA	5 µg
Amplified DNA Sample Library	1 µg
1,000 µM TS-HE Universal Oligo 1	1,000 pmol
1,000 µM TS-INV-HE Index Oligo	1,000 pmol
2X Hybridization Buffer (vial 5)	7.5 µl
Hybridization Component A (vial 6)	3 µl
SeqCap EZ Library plus water	4.5 µl
Total	15 µl

11. Vortex for 3 seconds and centrifuge at maximum speed for 10 seconds.
12. Incubate in a thermocycler at 47°C overnight. The thermocycler's heated lid should be turned on and set to maintain 57°C (+10°C above the hybridization temperature).

Step 3. Washing and Recovering Captured DNA

1. After the completion of the first hybridization, follow 'Chapter 6. Washing and Recovering Captured DNA' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* for instructions regarding the binding of the DNA to Streptavidin Dynabeads and the washing of the bead bound DNA.
2. Once 'Chapter 6. Washing and Recovering Captured DNA' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* is complete, follow Step 4 (below) of this protocol for instructions regarding Mid-Capture LM-PCR.

Step 4. Amplifying Captured DNA Using LM-PCR (i.e. Mid-Capture LM-PCR)

1. Follow 'Chapter 7. Amplifying Captured DNA Using LM-PCR' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* for instructions regarding the amplification and clean-up of Post-Capture LM-PCR amplified samples with the following modifications:
   - Perform two PCR reactions, both at 100 µl total volume.
   - Utilize 5 PCR cycles (total) for the Mid-Capture LM-PCR amplification of the captured DNA.
   - Elute DNA from the Qiagen QIAquick PCR Purification Kit column using 50 µl PCR grade water.
2. The 5 amplification cycles specified for the Mid-Capture LM-PCR reaction typically produce only very small amounts of DNA which may not be detectable above background after analysis with an Agilent DNA 1000 chip. This result does not indicate a failure of the first capture. All of the amplified material will be used in the second capture.

Hybridizing the Samples and SeqCap EZ Libraries – Second Capture

Step 5. Prepare the Second Hybridization Sample

1. Turn on heat block to 95°C and let it equilibrate to the set temperature.
2. Remove an aliquot of (2.25 µl) SeqCap EZ Library from the freezer (-15 to -25°C) and allow it to thaw on ice. Then add 2.25 µl of PCR grade water for a total of 4.5 µl.
3. Add all 50 µl of the purified, amplified sample library from Step 4, plus 5 µl of 1 mg/ml COT DNA, to a new 1.5 ml tube.

Component	Solution Capture
COT DNA	5 µg
Amplified, captured sample library	50 µl

4. Add 1 µl of TS-INV-HE Oligo 1 (1,000 µM) and 1 µl of the appropriate (refer to table below) TS-INV-HE Index Oligo(s) (1,000 µM) to the amplified sample library plus Cot-1 DNA.

Illumina TruSeq DNA Adapter Index Used for library construction	**TS-INV-HE Index Oligo** To use in hybridization
AD001	TS-INV-HE Index 1 Oligo
AD002	TS-INV-HE Index 2 Oligo
AD003	TS-INV-HE Index 3 Oligo
AD004	TS-INV-HE Index 4 Oligo
AD005	TS-INV-HE Index 5 Oligo
AD006	TS-INV-HE Index 6 Oligo
AD007	TS-INV-HE Index 7 Oligo
AD008	TS-INV-HE Index 8 Oligo
AD009	TS-INV-HE Index 9 Oligo
AD010	TS-INV-HE Index 10 Oligo
AD011	TS-INV-HE Index 11 Oligo
AD012	TS-INV-HE Index 12 Oligo

5. Close the tube's lid and make a hole in the top of the tube's cap with an 18-20 gauge needle.
6. Dry the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool in a DNA vacuum concentrator on high heat (60°C). Denaturation of the DNA with high heat is not problematic after linker ligation because the hybridization utilizes single-stranded DNA.
7. Once the sample is dry, cover the hole in the tube's cap with a sticker or small piece of laboratory tape.
8. To each dried-down the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool, add:

   - 7.5 µl of 2X Hybridization Buffer (vial 5)
   - 3 µl of Hybridization Component A (vial 6)

   The tube with the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool should now contain the following components:

Component	Solution Capture
COT DNA	5 µg
Amplified DNA Sample Library	1 µg
1,000 µM TS-HE Universal Oligo 1	1,000 pmol
1,000 µM TS-INV-HE Index Oligo	1,000 pmol
2X Hybridization Buffer (vial 5)	7.5 µl
Hybridization Component A (vial 6)	3 µl
Total	10.5 µl

9. Vortex the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool plus Hybridization Cocktail (2X Hybridization Buffer + Hybridization Component A) for 10 seconds and centrifuge at maximum speed for 10 seconds.
10. Place the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail in a 95°C heat block for 10 minutes to denature the DNA.
11. Centrifuge the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail at maximum speed for 10 seconds at room temperature.
12. Transfer the Amplified DNA Sample Library/COT DNA/Hybridization Enhancing Oligo Pool/Hybridization Cocktail to the 4.5 µl aliquot of SeqCap EZ Library in a 0.2 ml PCR tube prepared in Step 5.2.

    The hybridization sample should now contain the following components:

Component	Solution Capture
COT DNA	5 µg
Amplified DNA Sample Library	1 µg
1,000 µM TS-HE Universal Oligo 1	1,000 pmol
1,000 µM TS-INV-HE Index Oligo	1,000 pmol
2X Hybridization Buffer (vial 5)	7.5 µl
Hybridization Component A (vial 6)	3 µl
SeqCap EZ Library plus water	4.5 µl
Total	15 µl

13. Vortex for 3 seconds and centrifuge at maximum speed for 10 seconds.
14. Incubate in a thermocycler at 47°C overnight. The thermocycler's heated lid should be turned on and set to maintain 57°C (+10°C above the hybridization temperature).

Step 6. Washing and Recovering Captured DNA

1. Follow 'Chapter 6. Washing and Recovering Captured DNA' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* for instructions regarding the binding of the DNA to Streptavidin Dynabeads and the washing of the bead bound DNA.
2. Once 'Chapter 6. Washing and Recovering Captured DNA' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* is complete, follow Step 7 (below) of this protocol for instructions regarding Post-Capture LM-PCR.

Step 7. Amplifying Captured DNA Using LM-PCR (i.e. Post-Capture LM-PCR)

1. Follow 'Chapter 7. Amplifying Captured DNA Using LM-PCR' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* for instructions regarding the amplification, clean-up, and QC of Post-Capture LM-PCR amplified samples.
2. If the purified, amplified sample library passes the QC metrics detailed in Chapter 7 of the *NimbleGen SeqCap EZ Library SR User Guide v3.0* then proceed with Step 8.

Step 8. Measuring Enrichment Using qPCR

1. Follow 'Chapter 8. Measuring Enrichment Using qPCR' in the *NimbleGen SeqCap EZ Library SR User Guide v3.0* for instructions regarding the analysis of the captured sample library using qPCR.
   Note: *qPCR estimates of enrichment obtained following the double capture protocol are often observed to be lower than those obtained following the standard capture protocol, although the actual enrichment may be greater. This is believed to be due to superior enrichment of the primary capture target relative to the internal control targets, and not poor enrichment of the internal control targets themselves.*
2. Proceed with sequencing.

Ordering Information

Product	Cat. No.	Pack Size
SeqCap EZ Human Exome Library v3.0	06 465 684 001	4 reactions
	06 465 692 001	48 reactions
SeqCap EZ Choice Library	06 266 282 001	12 reactions
	06 266 304 001	24 reactions
	06 266 312 001	48 reactions
	06 266 339 001	96 reactions
	06 266 347 001	384 reactions
	06 266 355 001	960 reactions
SeqCap EZ Choice XL Library	06 266 363 001	12 reactions
	06 266 371 001	24 reactions
	06 266 380 001	48 reactions
	06 266 398 001	96 reactions
	06 266 401 001	384 reactions
	06 266 517 001	960 reactions
SeqCap EZ Developer Library	06 471 684 001	12 reactions
	06 471 706 001	24 reactions
	06 471 714 001	48 reactions
	06 471 722 001	96 reactions
	06 471 749 001	384 reactions
	06 471 757 001	960 reactions
SeqCap EZ Hybridization and Wash Kit	05 634 261 001	24 reactions
	05 634 253 001	96 reactions
SeqCap EZ Developer Reagent	06 684 335 001	1 Kit

NOTICE: This product may be subject to certain use restrictions. Before using this product please refer to the Online Technical Support page (*http://technical-support.roche.com*) and search under the product number or the product name, whether this product is subject to a license disclaimer containing use restrictions.

www.nimblegen.com

NimbleGen

For life science research only.
Not for use in diagnostic procedures.

Published by:
Roche Diagnostics GmbH
Sandhofer Straße 116
68305 Mannheim
Germany

06870406001 ❶ 08/12