

Targeted Capture and High-Throughput Sequencing Using Molecular Inversion Probes (MIPs)

Stuart Cantsilieris, Holly A. Stessman, Jay Shendure, and Evan E. Eichler

Abstract

Molecular inversion probes (MIPs) in combination with massively parallel DNA sequencing represent a versatile, yet economical tool for targeted sequencing of genomic DNA. Several thousand genomic targets can be selectively captured using long oligonucleotides containing unique targeting arms and universal linkers. The ability to append sequencing adaptors and sample-specific barcodes allows large-scale pooling and subsequent high-throughput sequencing at relatively low cost per sample. Here, we describe a “wet bench” protocol detailing the capture and subsequent sequencing of >2000 genomic targets from 192 samples, representative of a single lane on the Illumina HiSeq 2000 platform.

Key words Molecular inversion probes, Massively parallel sequencing, Real-time PCR, Exonuclease cleanup and gel electrophoresis

1 Introduction

The ability to selectively enrich thousands of genomic DNA targets and sequence them in parallel has tremendously impacted the way genomes can be interrogated on a large scale [1]. Molecular inversion probes (MIPs) represent one such approach based on target circularization of single-stranded oligonucleotides consisting of a common DNA backbone flanked by target-specific sequences [2] (Fig. 1). Following hybridization of site-specific targeting arms, non-strand displacing DNA polymerase and deoxynucleotides facilitate extension (gap-closure) between targeting arms and the intervening sequence. The addition of DNA ligase completes the covalently closed circular molecule and exonuclease treatment removes linear DNA that failed to form a closed circle. PCR using universal primers complementary to the MIP backbone completes the DNA capture reaction and the library is, in principle, ready for DNA sequencing [3, 4].

These authors contributed equally to this work.

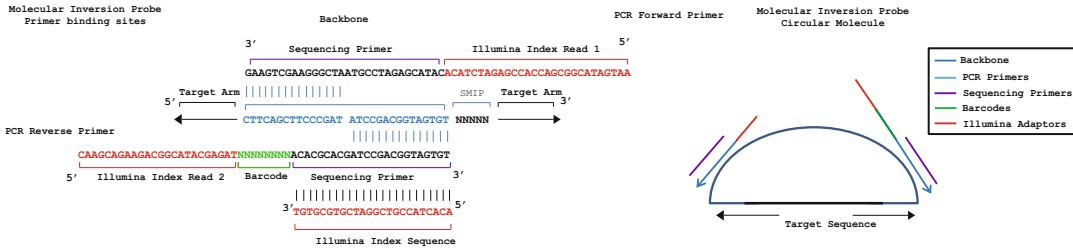


Fig. 1 Text-based primer map for molecular inversion probes. Sequence overlaps are annotated against the MIP backbone (*blue*). Positions of the forward and reverse PCR primers are annotated in *black* with Illumina index primers annotated in *red*. Forward and reverse sequencing primers (*purple*) overlap the MIP backbone and the MIP PCR primers. The 8 bp sample-specific barcode is annotated in (*green*) and the small molecular tag in (*grey*)

The success of any targeted enrichment approach is directly impacted by the performance of the DNA capture reaction. The MIP protocol has proven to be adaptable and the integration of recent technical advances has led to notable improvements in MIP performance [4–7]. MIPs demonstrate consistent capture uniformity (~98% of captured targets), capture specificity (>99% target overlap), and multiplex scalability (thousands of capture targets) [4, 5]. Improvements to MIP-design tools also allow *in silico* predictions of assay success leading to increases in capture efficiency [6]. In addition, the use of single-molecule tagging, by adding random unique barcode tags to each molecule (termed smMIPs), has also facilitated the quantitation of individual capture events, allowing for highly sensitive variant calling and precise quantitation of somatic or mosaic events [7]. The simplicity of the workflow procedure, low sample input requirements, and cost-effectiveness of the MIP protocol have proven advantageous for the detection of rare and *de-novo* mutations in large disease cohorts [5, 8]. This protocol therefore describes in detail a method for large-scale resequencing of several thousand genomic targets using MIPs.

2 Materials

Prepare all dilutions using nuclease-free water. All enzymes and mastermixes should be stored at -20°C unless otherwise noted. All reagents should be thawed and prepared on ice unless otherwise noted. All waste disposal regulations must be followed when disposing of hazardous materials.

1. Components for MIP pooling and phosphorylation.
 - (a) 70mer oligonucleotides synthesized at the 25 nanomole (nM) scale and hydrated to 100 micromole (μM) in $1\times$ TE Buffer, pH 8.0. Store at -20°C .
 - (b) T4 DNA Ligase Reaction Buffer with 10 mM ATP (New England Biolabs). Store at -20°C .

- (c) T4 Polynucleotide Kinase. Store at $-20\text{ }^{\circ}\text{C}$.
 - (d) ABgene 8-Flat-Cap Strip Tubes.
 - (e) Costar* Microcentrifuge Tubes 1.7 mL; Color: Natural (holds 1.5 mL).
 - (f) Nuclease-free water or equivalent. Store at room temperature.
2. Components for targeted capture.
- (a) Ampligase 10 \times Reaction Buffer (Epicentre). Store at $-20\text{ }^{\circ}\text{C}$.
 - (b) Ampligase® DNA Ligase (Epicentre). Store at $-20\text{ }^{\circ}\text{C}$.
 - (c) Hemo Klentaq (New England Biolabs). Store at $-20\text{ }^{\circ}\text{C}$.
 - (d) 10 mM Deoxynucleotide (dNTP) set. Product should be diluted fresh for each capture reaction 1:40 (0.25 mM). Store at $-20\text{ }^{\circ}\text{C}$.
 - (e) Nuclease-free water or equivalent. Store at room temperature.
 - (f) Phosphorylated MIP pool (*see* **Notes 1** and **2**).
 - (g) Eppendorf skirted 96-well plates or equivalent (clear).
 - (h) Thermo Scientific* ABgene* Adhesive PCR Film or equivalent.
3. Components for exonuclease treatment.
- (a) Exonuclease I (*E. coli*). Store at $-20\text{ }^{\circ}\text{C}$.
 - (b) Exonuclease III (*E. coli*). Store at $-20\text{ }^{\circ}\text{C}$.
4. Components for PCR.
- (a) iProofHigh Fidelity Master Mix (Bio-Rad). Store at $-20\text{ }^{\circ}\text{C}$.
 - (b) SYBR® Green I nucleic acid gel stain (Invitrogen). Store at $-20\text{ }^{\circ}\text{C}$. Keep away from light.
 - (c) Oligonucleotides. Synthesized at the 25 nM scale and hydrated to 100 μM in 1 \times TE Buffer, pH 8.0. Store at $-20\text{ }^{\circ}\text{C}$.
 - (d) Low-Profile 0.2 mL 8-Tube Strips Without Caps.
 - (e) Optical Flat 8-Cap Strips.
5. Clean-up protocol components.
- (a) Agencourt AMPure XP beads. Store at $4\text{ }^{\circ}\text{C}$.
 - (b) Ethanol (100%). Store at room temperature.
 - (c) DynaMag™-2 magnet (ThermoFisher Scientific).
 - (d) Buffer EB (Qiagen). Store at room temperature.
6. Agarose gel electrophoresis components.
- (a) E-Gel® EX Gel, 2% (Invitrogen).
 - (b) E-Gel® Low Range Quantitative DNA Ladder (Invitrogen). Store at $4\text{ }^{\circ}\text{C}$.

7. Sequencing components.

- (a) Qubit dsDNA High Sensitivity Assay Kit. Store at room temperature.
- (b) 0.5 mL tubes (for Qubit) or equivalent.
- (c) Illumina MiSeq Reagent Kit (300 cycles PE). Store at -20°C .

3 Methods

3.1 Oligonucleotide Pooling and Phosphorylation

1. Design MIPs using an existing pipeline [6] (*see* **Notes 1** and **2**).
2. Pool oligonucleotides at equimolar concentrations by plate, by combining 5 μL of each MIP (100 $\mu\text{M}/\mu\text{L}$) into a single 1.5 mL tube. Each individual 1.5 mL tube will represent a combined sum of 96 MIPs for a total volume of 480 μL (*see* **Note 3**).
3. Take 9.6 μL of each individual MIP pool (0.1 μL multiplied by the number of MIPs in each plate) and combine these into a single tube to generate a MIP megapool.
4. Phosphorylate the MIP megapool by combining 25 μL of the MIP megapool, 3 μL of 10 \times T4 DNA Ligase Reaction Buffer, 1 μL of T4 Polynucleotide Kinase (10 U), and 1 μL of nuclease-free water in a total reaction volume of 30 μL . Using a thermocycler, incubate the reaction mix at 37 $^{\circ}\text{C}$ for 45 min with a final denaturation step of 65 $^{\circ}\text{C}$ for 20 min. Store unphosphorylated MIPs at -20°C for future use.

3.2 Targeted MIP Capture

1. Calculate the volume of the MIP megapool required in the capture reaction based on the ratio of desired MIP copies to DNA copies. This example will assume a megapool of 2000 MIPs captured using 100 ng of total genomic DNA, for a total ratio of 800 MIP copies to 1 DNA copy.
2. Calculate the expected number of MIP copies required given an input of 100 ng of genomic DNA, e.g., $800 \times 33,000$ haploid genome copies = 2.64×10^7 MIP copies required.
3. Transform the number of MIP copies to picomoles (pmol) using Avogadro's number (6.02×10^{23}), e.g., $(2.64 \times 10^7 / 6.02 \times 10^{23}) (1 \times 10^{12}) = 4.38 \times 10^{-5}$ pmol.
Calculate the picomole per μL concentration of the MIP megapool:
e.g., $0.1 \mu\text{L} \times 100 \mu\text{M} / 2000 \text{ MIPs} = 0.005 \mu\text{M}$
 $(0.005 \times 25 \mu\text{L}) / 30 \mu\text{L} = 0.004 \text{ pmol}/\mu\text{L}$
4. Calculate the volume of 1 \times MIP megapool required in the capture reaction (*see* **Note 4**):
e.g., $4.38 \times 10^{-5} \text{ pmol} / 0.004 \text{ pmol}/\mu\text{L} = 0.011 \mu\text{L}$ per capture reaction.
5. Prepare a 15 μL capture reaction on ice by combining 2.5 μL of Ampligase 10 \times Reaction Buffer, 0.0032 μL of 0.006 mM

dNTP mix, 0.32 μL Klentaq (10 U/ μL), 0.01 μL of Ampligase (100 U/ μL), 0.0105 μL of MIP megapool, and 12.16 μL of nuclease-free water. The total volume of DH_2O can be scaled depending on your DNA concentration requirements and the volumes are based on processing 192 samples (see below).

6. Plate 10 μL of DNA into a 96-well plate format (10 μL at 10 ng/ μL). A range of 100–200 ng total DNA can be used in the final capture reaction.
7. Add 15 μL of capture reaction to each individual DNA sample.
8. Seal with adhesive PCR film (*see Note 5*).
9. Using a thermocycler, incubate the reaction mix at 95 °C for 10 min and 60 °C for 22 h. Remove plates and immediately place on chilling blocks (*see Note 6*).
10. Exonuclease treatment.
 - (a) Immediately following capture, prepare a Exonuclease clean-up master mix containing 0.5 μL of Exonuclease I, 0.5 μL of Exonuclease III, 0.2 μL Ampligase 10 \times Reaction Buffer and 0.8 μL nuclease-free water per sample.
 - (b) Add 2.0 μL of Exonuclease clean-up mix to each 25 μL capture reaction (*see Note 7*).
 - (c) Using a thermocycler, incubate the reaction at 37 °C for 45 min and 95 °C for 2 min. Cool reaction plates to 4 °C (*see Note 5*).
 - (d) Samples may be stored at 4 °C for a short term until PCR, or –20 °C for longer periods.
11. Real-Time PCR.
 - (a) Prepare a RT-PCR master mix by combining 12.5 μL of 2 \times iProof High Fidelity Master Mix, 0.125 μL of 100 μM universal MIP barcode forward primer, 0.125 μL of 100 \times SYBR[®] Green I nucleic acid gel stain and 6.125 μL of nuclease-free water.
 - (b) Add 18.75 μL of RT-PCR master mix to each well.
 - (c) Add 1.25 μL of 10 μM individual barcode primers and 5 μL of exonuclease-treated MIP capture reaction to each individual well (*see Note 8*).
 - (d) Using an RT-PCR thermocycler, amplify the reaction until the reaction begins to plateau under the following conditions: 98 °C for 30 s, followed by 20–25 cycles of 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 30 s (*see Note 9*).
12. Standard PCR.
 - (a) Prepare a PCR master mix by combining 12.5 μL of 2 \times iProof High Fidelity Master Mix, 0.125 μL of 100 μM

universal MIP barcode forward primer, and 6.25 μL of nuclease-free water.

- (b) Add 18.75 μL of RT-PCR master mix to each well.
- (c) Add 1.25 μL of 10 μM individual barcode primers and 5 μL of exonuclease-treated MIP capture reaction into each individual well.
- (d) Using a PCR thermocycler, amplify the reaction under the following conditions: 98 $^{\circ}\text{C}$ for 30 s, followed by 20–25 (established in **step 11** of the real-time PCR protocol) cycles of 98 $^{\circ}\text{C}$ for 10 s, 60 $^{\circ}\text{C}$ for 30 s, and 72 $^{\circ}\text{C}$ for 30 s with a final extension time of 72 $^{\circ}\text{C}$ for 2 min and 4 $^{\circ}\text{C}$ forever.

13. Product pooling, clean-up, and gel electrophoresis.

- (a) For each plate of DNA samples pool 5 μL of each PCR reaction into a 1.5 mL tube ($5 \mu\text{L} \times 96 = 480 \mu\text{L}$) (*see Note 3*).
- (b) Determine the correct ratio of beads to pooled MIP library by using a bead titration (*see Note 10*).
- (c) Add 0.9 μL of Agencourt AMPure XP beads per 1 μL of pooled PCR reaction, e.g., (432 μL per 480 μL of pooled PCR reaction). Vortex the tube thoroughly and pulse spin down to remove the beads from within the cap (*see Note 11*).
- (d) Incubate the sample pool with the beads for 10 min at room temperature.
- (e) Place the tube on the DynaMagTM-2 magnet, lift the cap and allow the beads to adhere to the side of the tube nearest the magnet for 5 min.
- (f) Slowly remove the supernatant using a pipette without disturbing the bead pellet. If the bead pellet is disturbed, pipette them back into the tube and wait a further 1–3 min for the beads to re-bind.
- (g) Wash the bead pellet by adding 1 mL of 70% ethanol to fully immerse the beads while the tube is still attached to the magnet. Do not disturb the bead pellet and incubate for 30 s.
- (h) Remove the supernatant and repeat step (13 g).
- (i) Remove the supernatant completely from the tube, making sure that there is no ethanol left at the bottom of the tube without disturbing the bead pellet (*see Note 12*).
- (j) Allow the beads to dry for 5 min (*see Note 13*).
- (k) Remove the tube containing the beads from the magnet and add 100 μL of EB buffer; mix well by manually pipetting up and down at least ten times. Allow the beads to sit at room temperature for 1 min (*see Note 14*).

- (l) Transfer the tube back to the magnet and incubate for at least 1 min allowing the beads to separate from the EB buffer and adhere to the side of the tube.
- (m) Transfer the supernatant, which contains the cleaned MIP library, to a new 1.5 mL tube. Individual MIP libraries can be stored at 4 °C short term or -20 °C for longer periods.
- (n) Run the MIP library on a 2% E-Gel® EX Gel by combining 2 µL of pooled MIP library with 18 µL of distilled water and loading 20 µL into the individual wells. Prepare a 100 bp DNA ladder by preparing a 1:1 ratio of E-Gel® Low Range Quantitative DNA Ladder with distilled water and load into the first or final wells in the gel (20 µL).
- (o) Run gel electrophoresis for 20 min using the E-Gel® EX Gel platform and confirm the presence of a 276 bp product (*see* **Notes 15** and **16**).

14. Massively parallel sequencing.

Quantitate and Pool MIP Libraries

- (a) Prepare individual pooled libraries for sequencing by normalizing each individual library against the concentration of the lowest library within the set pools.
- (b) Use the Qubit dsDNA High-Sensitivity assay kit to determine the concentration of each individually barcoded library [**9**, **10**].
- (c) Combine each library at equal concentration and determine the final concentration of pooled MIP library as in **step 14.b**.
- (d) The size of the MIP megapool, the number of pooled samples, and the desired depth of coverage will determine the individual sequencing requirements. The following protocol uses the Illumina MiSeq platform to test and rebalance individual MIP libraries (*see* **Note 17**).

Denature and Dilute MIP libraries

- (e) Denature and dilute MIP libraries according to the Standard Normalization Methods described in the MiSeq Denature and Dilute Libraries Guide [**11**].
- (f) Prepare a fresh 0.2 N dilution of NaOH by combining 200 µL of stock 1 N NaOH and 800 µL of DH₂O.
- (g) Dilute the MIP library to 2 nM; then add 5 µL of the library to 5 µL of 0.2 N NaOH.
- (h) Vortex the tube thoroughly and pulse spin down to remove the liquid from the lid. Incubate for 5 min at room temperature.
- (i) Prepare a 20 pmol denatured library by adding 990 µL of chilled HT1 Buffer to 10 µL of denatured MIP library.

- (j) Dilute the denatured 20 pmol MIP library according to desired MiSeq loading concentrations (6–20 pmol). 10 pmol is usually optimal for the majority of MIP libraries.

Loading the MiSeq Reagent Cartridge

- (k) Load the diluted MIP library (6–20 pmol) into the MiSeq reagent cartridge according to the MiSeq: Reagent Kit v3-Preparation Guide [12].
- (l) Prepare the forward, reverse, and index sequencing primers to a concentration of 10 μ M and load into the MiSeq reagent cartridge, according to the MiSeq: Reagent Kit v3-Preparation Guide (*see Note 18*) [12].
- (m) Set up a sequencing run according to the MiSeq System User Guide [13].

15. Assessment of MIP performance.

- (a) Assess capture uniformity by plotting the depth of coverage for individually mapped MIPs.
- (b) Normalize read counts for each individual MIP by the total number of reads mapped.
- (c) Sort in descending order and plot the ranked uniformity of MIPs in Log_{10} scale.
- (d) Rebalance poor-performing MIPs by increasing the relative concentration of MIPs that are one order of magnitude lower in abundance (*see Note 19*) (Fig. 2).
- (e) Return to methods **step 3.2** and set up the MIP capture using the rebalanced MIP megapool.

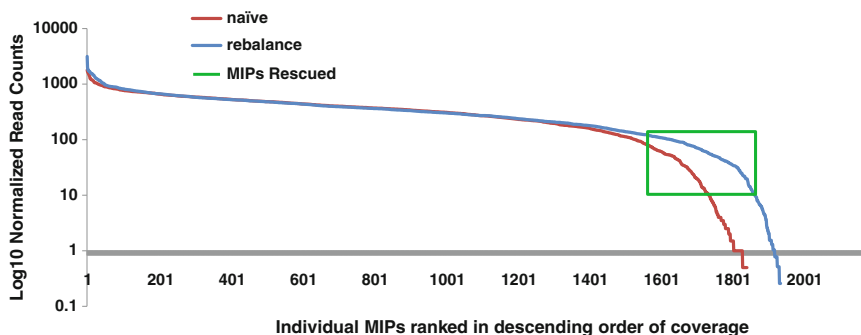


Fig. 2 Capture uniformity for 2196 MIPs “pre” (*blue*) and “post” (*red*) rebalancing. MIPs that perform poorly (one order of magnitude lower in abundance) are rebalanced at a ratio of 50:1 (bad vs. good MIPs) and a substantial number of MIPs are rescued (*green*) upon rebalancing

4 Notes

1. Download the MIPgen design and analysis suite of tools from GitHub (<https://github.com/shendurelab/MIPGEN>). Use MIPgen to design MIPs across your regions of interest. Note that there are several other dependencies for running this software (e.g., SAMtools, BWA, Tabix) successfully in your local environment.
2. MIPs can be customized to target moderate- and high-complexity DNA targets ranging from 120 to 250 base pairs in size. Low complexity and high GC regions of the genome perform poorly in this assay due primarily to the reliance of the method on PCR amplification and Illumina sequencing. Select your MIPs to be synthesized based on the SVR scores, logistic scores and failure flags (see the MIPGEN README file that accompanies this software package).
3. For ease of handling, use an 8-channel pipette to pool 5 μL of 100 μM MIPs from each well in the 96-well plate. Each tube in the 8-cap strip represents a combined sum of 12 wells or MIPs (5 $\mu\text{L} \times 12$ wells = 60 μL) which can be pooled together to generate a 96 MIP pool containing a volume of 480 μL (60 $\mu\text{L} \times 8$ strip tubes).
4. If the volume of MIP megapool is too small for manual pipetting, dilute the MIP megapool to a lower concentration e.g., 1:1000, so a higher volume can be added. Dilutions should be made fresh for each capture reaction.
5. During this step be sure to create an air-tight seal using the adhesive PCR seal. Use a 10 °C lid offset for each step of the reaction. Failure to perform this step thoroughly will cause the DNA to evaporate during the capture reaction.
6. Capture incubation times may be reduced depending on input DNA concentration. Minimum working DNA stocks should not be less than 100 ng total for the MIP capture reaction.
7. Before adding the exonuclease treatment, cool down the capture plates using cold blocks and prepare the reaction mix on ice. Dispense the exonuclease clean-up mix in equal volumes across a set of 8-Flat-Cap Strip Tubes. Use an 8-channel pipette to dispense 2 μL of exonuclease reaction mix into each capture reaction.
8. RT-PCR is performed using a universal forward primer (MIP_universal_forward: AATGATACGGCGACCACCGAGATC TACACATACGAGATCCGTAATCGGGAAGCTGAAG) and an individual reverse primer (MIP_barcode_reverse: CAAG CAGAAGACGGCATAACGAGATNNNNNNNNACAC GCACGATCCGACGGTAGTGT) containing a unique 8mer barcode sequence, which is used for subsequent pooling and sequencing.

9. DNA samples extracted and stored under different conditions will reach plateau at different points during PCR cycling. It is recommended that RT-PCR be performed on a small number of samples representative of each particular sample set so that the correct number of cycles can be established. It is common for a percentage of samples to reach plateau at different cycle points. Select the cycle in which the majority of samples are still within log linear phase before plateau. Once completed, standard PCR may subsequently be performed using the correct number of cycles per sample set.
10. Small contaminants and undesired PCR products are removed during the bead clean-up. However, the ratio of beads to PCR product may vary depending on the size of the MIP library. Here, we use a concentration of 0.9× beads to clean up the pooled MIP library. To determine the quantity of beads to use, perform a bead titration by cleaning up control libraries with varied ratios of beads to MIP library (e.g., 0.8×–1.4× beads) and evaluating by agarose gel electrophoresis. As Agencourt AMPure XP beads preferentially bind to larger DNA fragments, the desired MIP PCR product (276 bp) can be saved while removing other nonspecific PCR products.
11. Allow 100% Agencourt AMPure XP beads to come to room temperature before beginning clean-up. Vortex thoroughly to resuspend the beads into the buffer and dissolve the bead pellet at the bottom of the tube.
12. Tap the magnet gently to consolidate the ethanol at the bottom of the tube and use a p10 pipette tip to remove any residual ethanol.
13. Exceeding 5 min drying time may result in a lower DNA yield.
14. Optimize the amount of elution buffer added to individual MIP libraries to achieve the desired concentration. Smaller MIP pools can typically be eluted in lower volumes.
15. The 276 bp MIP product is specifically based on capturing 162 bp of target sequence using targeting arm lengths of 40–45 bp and single-molecule tags of 5 bp.
16. A small amount of nonspecific product (150 bp) may still remain after bead clean-up; as long as the MIP library (276 bp) represents the predominant band, this should not impact further sequencing steps.
17. Paired-end 101 bp reads are sufficient to sequence individual MIP amplicons of 276 bp, capturing 162 bases with arm lengths of 40–45 bp, 5–8 bp single-molecule tags with enough overlap for read assembly. This can be modified according to the specifics of the individual sequencing library.
18. Sequencing is performed using forward primer: 5' CATA C GAGATCCGTAATCGGGAAGCTGAAG 3', MIPseq reverse

primer: 5' ACACGCACGATCCGACGGTAGTGT 3', and MIPseqindexprimer: 5' AACTACCGTCGGATCGTGCGTGT 3'.

19. Poor-performing MIPs can be recovered by “spiking” MIPs in increased relative concentrations, termed rebalancing. Separate MIPs that perform well at 1× concentration from MIPs that require rebalancing. Phosphorylate these MIP pools separately, then pool at a ratio 10:1, 50:1, and 100:1 (poor performers: good performers). MIPs that perform particularly poorly, for example, those that generate zero sequence reads, may not be recoverable and can affect the overall performance of the MIP pool. It is recommended to do a second test run of the rebalanced MIP pool before testing large sample numbers. Check for large proportions of off-target reads indicative of rare MIPs with high off-target capture. This can be avoided by checking output files from the MIPgen design files for MIPs that have over represented arm sequences.

Acknowledgments

We thank Bradley P. Coe for his critical review of the manuscript and Tonia Brown for assistance with the manuscript preparation. We thank Brian J. O’Roak, Beth Martin, Evan A. Boyle, and Joseph B. Hiatt for their overall contributions to developing the MIP protocol. S.C. is supported by a National Health and Medical Research Council (NHMRC) CJ Martin Biomedical Fellowship (#1073726). H.A.S. is supported, in part, by the NHGRI Interdisciplinary Training in Genome Science Grant (T32HG00035). E.E.E. is an investigator of the Howard Hughes Medical Institute. J.S. is an investigator of the Howard Hughes Medical Institute.

Competing financial Interests

E.E.E. is on the scientific advisory board (SAB) of DNAnexus, Inc., and is a consultant for the Kunming University of Science and Technology (KUST) as part of the 1000 China Talent Program.

References

1. Mamanova L, Coffey AJ, Scott CE et al (2010) Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7:111–118
2. Hardenbol P, Baner J, Jain M et al (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol* 21:673–678
3. Porreca GJ, Zhang K, Li JB et al (2007) Multiplex amplification of large sets of human exons. *Nat Methods* 4:931–936
4. Turner EH, Lee C, Ng SB et al (2009) Massively parallel exon capture and library-free resequencing across 16 genomes. *Nat Methods* 6:315–316
5. O’Roak BJ, Vives L, Fu W et al (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338:1619–1622
6. Boyle EA, O’Roak BJ, Martin BK et al (2014) MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. *Bioinformatics* 30:2670–2672

7. Hiatt JB, Pritchard CC, Salipante SJ et al (2013) Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Res* 23:843–854
8. O'Roak BJ, Stessman HA, Boyle EA et al (2014) Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun* 5(5595):1–6
9. Qubit® Assays: Quick Reference Guide: Pub. no: MAN0010876
10. Qubit® 2.0 Fluorometer: MAN0003231
11. MiSeq: Denature and Dilute Libraries Guide: 15039740v1
12. MiSeq: Reagent Kit v3- Preparation Guide: Part#15044983
13. MiSeq System User Guide: part # 15027617