# Method

# Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture

Nadin Rohland[1] and David Reich

*Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02139, USA*

Improvements in technology have reduced the cost of DNA sequencing to the point that the limiting factor for many experiments is the time and reagent cost of sample preparation. We present an approach in which 192 sequencing libraries can be produced in a single day of technician time at a cost of about $15 per sample. These libraries are effective not only for low-pass whole-genome sequencing, but also for simultaneously enriching them in pools of approximately 100 individually barcoded samples for a subset of the genome without substantial loss in efficiency of target capture. We illustrate the power and effectiveness of this approach on about 2000 samples from a prostate cancer study.

[Supplemental material is available for this article.]

Improvements in technology have reduced the sequencing cost per base by more than a 100,000-fold in the last decade (Lander 2011). The amount of sequence data that is needed per sample, for example, for studying small target regions or low-coverage sequencing of whole genomes is often less than the commercial cost of "library" preparation, so that library preparation is now often the limiting cost for many projects. To reduce library preparation costs, researchers can purchase kits and produce libraries in their own laboratories or use published library preparation protocols (Mamanova et al. 2010; Meyer and Kircher 2010; Fisher et al. 2011). However, this approach has two limitations. First, available kits have limited throughput so that scaling to thousands of samples is difficult without automation. Second, an important application of next-generation sequencing technology is to enrich sample libraries for a targeted subsection of the genome (like all the exons) (Albert et al. 2007; Hodges et al. 2007; Gnirke et al. 2009), and then to sequence this enriched pool of DNA, but such experiments are expensive because of the high costs of target capture reagents. One way to save funds is to pool samples prior to target enrichment (after barcoding to allow them to be distinguished after the data are gathered). Although the recently introduced Nextera DNA Sample Prep Kit (Illumina) together with "dual indexing" ($12 \times 8$ indices and two index reads) allows higher sample throughput for library preparation (Adey et al. 2010) and pooling of up to 96 libraries, the long indexed adapter may interfere during pooled hybrid selection (see below).

We report a method for barcoded library preparation that allows highly multiplexed pooled target selection (hybrid selection or hybrid capture). We demonstrate its usefulness by generating libraries for more than 2000 samples from a prostate cancer study that we have enriched for a 2.2-Mb subset of the genome of interest for prostate cancer. We also demonstrate the effectiveness of libraries produced with this strategy for whole-genome sequencing, both by generating 40 human libraries and sequencing them to fivefold coverage, and by generating 12 microbial libraries and sequencing them to 150-fold coverage. Our method was engineered for high-throughput sample preparation and low cost, and thus we implemented fewer quality-control steps and were willing

to accept a higher rate of duplicated reads compared with methods that have been optimized to maximize library complexity and quality (Meyer and Kircher 2010; Fisher et al. 2011). Because of this, our method is not ideal for deep sequencing of large genomes (e.g., human genome at $30\times$), where sequencing costs are high enough that it makes sense to use a library that has as low a duplication rate as possible. However, our method is advantageous for projects in which a modest amount of sequencing is needed per sample, so that the savings in sample preparation outweigh costs due to sequencing duplicated molecules or failed libraries. Projects that fall into this category include low-pass sequencing of human genomes, microbial sequencing, and target capture of human exomes and smaller genomic targets.

Our method reduces costs and increases throughput by parallelizing the library preparation in 96-well plates, reducing enzyme volumes at a cost-intensive step, using inexpensive paramagnetic beads for size selection and buffer exchange steps (DeAngelis et al. 1995; Lennon et al. 2010; Meyer and Kircher 2010), and automation (Farias-Hesson et al. 2010; Lennon et al. 2010; Lundin et al. 2010; Fisher et al. 2011). To permit highly multiplexed sample pooling prior to target enrichment or sequencing, we attach "internal" barcodes directly to sheared DNA from a sample that is being sequenced, and flank the barcoded DNA fragments by partial sequencing adapters that are short enough that they do not strongly interfere during enrichment (the adapters are then extended after the enrichment step). By combining these individual libraries in pools and enriching them for a subset of the genome, we show that we obtain data that are effective for polymorphism discovery, without substantial loss in capture efficiency.

## Outline

Our method is based on a blunt-end-ligation method originally developed for the 454 Life Sciences (Roche) platform (Stiller et al. 2009), which we have extensively modified for the Illumina platform to reduce costs and increase sample processing speed, by parallelizing the procedure in 96-well format and automating the labor-intensive cleanup steps (Figs. 1, 2A; Methods; Supplemental Notes). Some of the modifications adapt ideas from the literature, such as DNA fragmentation on the Covaris E210 instrument in 96-well PCR plates (Lennon et al. 2010) or replacing the gel-based size selection by a bead-based, automatable, size selection (Lennon et al.

**Figure 1.** Experimental workflow of the library preparation protocol for 95 samples for pooled hybrid capture.

2010; Borgstrom et al. 2011). Another change is to replace a commonly used commercial kit (AMPure XP kit) for SPRI-cleanup steps with a homemade mix. An important feature of our libraries compared with almost all other Illumina library preparation methods (Cummings et al. 2010; Mamanova et al. 2010; Meyer and Kircher 2010; Teer et al. 2010) is that we add a 6-bp "internal" barcoded adapter to each fragment (Craig et al. 2008). These adapters are ligated directly to the DNA fragments, leading to "truncated" libraries with 34- and 33-bp overhanging adapters at the end of each DNA fragment. Adapters at this stage in our library preparation are sufficiently short that they interfere with each other minimally during hybrid capture, compared with what we have found when long adapters are used (64 and 61 bp on either side, including the 6-bp internal barcode). The truncated adapter sites are then extended to full length after hybrid capture allowing the libraries to be sequenced (Fig. 2A). To assess how this strategy works in different-sized pools (between 14 and 95), we applied it to 2.2 Mb of the genome of interest for prostate cancer, where it reduces the capture reagent that is required by two orders of magnitude while still producing highly useful data. Sequencing these libraries shows that we can perform pooled target capture on at least 95 barcoded samples simultaneously without substantial reduction in capture efficiency.

The fact that we are using internal strategy in which barcoded oligonucleotides are ligated directly to fragmented DNA is a nonstandard strategy, which deserves further discussion. First, when combined with indexing (introducing a second barcode via PCR after pooling) (Fig. 2A; Meyer and Kircher 2010), an almost unlimited number of samples can be pooled and sequenced in one lane. We are currently using this strategy in our prostate cancer study to test library quality and to assess the number of sequence-able molecules per library prior to equimolar pooling for hybrid capture. Second, a potential concern of our strategy of directly ligating barcodes is that differences in ligation efficiency for different barcodes in principle could cause some barcodes to perform less efficiently than others. However, to date, we have used each of 138 barcodes at least 15 times and have not found evidence of particular barcodes performing worse than others as measured by the number of sequenced molecules per library. Third, the blunt-end ligation used in our protocol results in a loss of 50% of the input DNA because two different adapters have to be attached to either side. This is not a concern for low-coverage and small-target studies using input DNA amounts of 500 ng or higher but is not an ideal strategy for samples with less input material. Fourth, chimeras of blunted DNA molecules can be created during blunt-end ligation. In our protocol, the formation of chimeras is reduced by using adapter oligonucleotides in such vast excess to the sample DNA that the chance of ligating barcodes to the DNA is much higher than ligating two sample molecules (while the adapters can form dimers, these are removed during bead cleanup). Fifth, when using our internal barcodes, it is important to pool samples in each lane in such a way that the base composition of the barcodes is balanced, because the Illumina base-calling software assumes balanced nucleotide composition especially during the first few cycles. This is of particular importance when only a few barcoded samples are being pooled. To prevent base-calling problems in such unbalanced pools, a PhiX library can be spiked into the library to increase diversity.

We performed a rough calculation breaking the cost for our method down into (a) reagents, (b) technician time, and (c) capital equipment (Table 1). The reagents and consumables cost is about $9 per sample without taking into account discounts that would be available for a project that produced large numbers of libraries. The
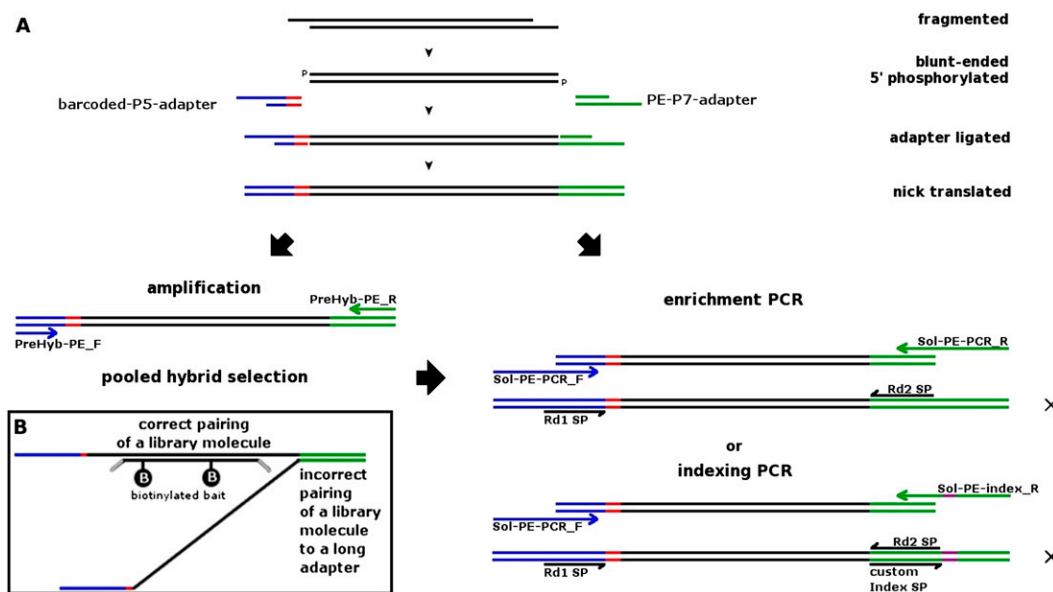
**Figure 2.** (*A*) Schematic overview of the library preparation procedure using the Illumina PE adapter (internal barcode in red). After a cascade of enzymatic reactions and cleanup steps, enrichment PCR can be performed to complete the adapter sites for Illumina PE sequencing (Rd1 SP, Rd2 SP are PE sequencing primers). Alternatively, libraries can be pooled for hybrid selection (if desired), and then enrichment PCR can be performed after hybrid selection. To achieve an even higher magnitude of pooling for sequencing, "indexing PCR" can be performed instead of "enrichment PCR," whereby unique indices (in purple) are introduced to the adapter, and a custom index sequencing primer (index-PE-sequencing-Primer) is used to read out the index in a separate read. Finished libraries that have all the adapters necessary to allow sequencing are marked with an X. (*B*) Schematic figure of "daisy-chaining" during pooled solution hybrid capture, which may explain why a large proportion of molecules are empirically observed to be off-target when using long adapters. Library molecules exhibiting the target sequences hybridize to biotinylated baits, but unwanted library molecules can also hybridize to the universal adapter sites. The adapters of our "truncated" libraries (including barcode: 34 and 33 bp) are about half the length of regular "long" adapters (64 and 61 bases), and thus may be less prone to binding DNA fragments that do not belong to the target region.

cost for technician time is $3 per sample, assuming that an individual makes 480 libraries on five plates per week. Capital costs are difficult to compute (because some laboratories may already have the necessary equipment), but if one computes the cost of a Covaris LE220 instrument, a PCR machine, and an Agilent Bravo liquid handling platform, and divides by the cost of 100,000 libraries produced over the 2–3-yr lifetime of these instruments, this would add about $3 more to the cost per sample. This accounting does not include administrative overhead, space rental, process management, quality control on the preparation of reagents, bioinformatic support, data analysis, and research and development, all of which could add significantly to cost.

## Application 1: Enrichment of more than 2000 human samples by solution hybrid capture

For many applications, it is of interest to enrich a DNA sample for a subset of the genome; for example, in medical genetics, a candidate region for disease risk, or all exons. The target-enriched (captured) sample can then be sequenced. To perform studies with statistical power to detect subtle genetic effects with genome-wide significance, however, it is often necessary to study thousands of samples (Kryukov et al. 2009; Lango Allen et al. 2010), which can be prohibitively expensive given current sample preparation and target enrichment costs. We designed our protocol with the aim of allowing barcoded and pooled samples to be captured simultaneously. Specifically, our libraries have internal barcodes that are tailored to pooled hybrid capture, whereas most other libraries have external barcodes in the long adapters. It has been hypothesized that hybridization experiments using libraries that already

have long adapters do not work efficiently in pooled hybridizations because a proportion of library molecules not only hybridize to the "baits" but also catch unwanted off-target molecules with the long adapter ("daisy-chaining") (Mamanova et al. 2010; Nijman et al. 2010), thus reducing capture efficiency (Fig. 2B). In the Supplemental Notes ("Influence of Adapter Length in Pooled Hybrid Capture"), we present experiments showing that the number of reads mapping to the target region increased from 29% to 73% when we shortened the adapters (Supplemental Table S1), providing evidence for the hypothesis that interference between barcoded adapters is lowered by short adapters. Our results show empirically that short adapters improve hybridization efficiency.

To investigate the empirical performance of our libraries in the context of target capture, we produced libraries for 189 human samples starting from 0.2–4.8 μg of DNA (98% <1 μg for fragmentation), prepared in two 96-well plates as in Supplemental Figure S1. We combined the samples into differently sized pools of libraries (14, 28, 52, and 95) and then enriched the pooled libraries using a custom Agilent SureSelect Target Enrichment Kit in the volume recommended for a single sample (the target was a 2.2-Mb subset of the genome containing loci relevant to prostate cancer). We sequenced the three smaller pools together on one lane of the Genome Analyzer II instrument (36 bp, single reads) and the 95-sample pool on one lane of a HiSeq2000 instrument (50 bp, paired-end reads). We aligned the reads to the human genome using BWA (Li and Durbin 2009), after removing the first six bases of the first read that we used to identify the sample. We removed PCR duplicates using Picard's (http://picard.sourceforge.net) MarkDuplicates and computed hybrid selection statistics with Picard's CalculateHsMetrics. For

**Table 1.** Cost and time assumptions for library preparation

| Task | Item | Price per sample | Sample processing time for 192 samples | Technician hands-on time for 192 samples |
|---|---|---|---|---|
| Covaris shearing | Plate | $ 0.04 | 44 h | 2 h |
| Cleanup | Beads and ethanol | $ 0.54 | 4 h | 2 h |
| Blunt end repair | Kit | $ 0.75 | 1.5 h | 0.5 h |
| Barcoded adapter ligation | Kit and oligonucleotides | $ 3.30 | 1.2 h | 0.5 h |
| Nick fill-in reaction | Enzyme and buffer | $ 0.48 | 1 h | 0.5 h |
| Amplification | Kit and oligonucleotides | $ 1.58 | 1–2 h | 0.5 h |
| Copy number determination | qPCR reagents or sequencing cost[a] | $ 0.67 | | |
| Consumables | Plates and pipette tips | $ 1.40 | | |
| | Subtotal | $ 8.76 | | 6 h |
| Technician salary | Total assuming 480/week[b] | $ 3.00 | | |
| Capital equipment | Amortized over 100,000 libraries[c] | $ 3.00 | | |
| | Total for library preparation | $ 14.76 | | |

[a]qPCR for two measurements per sample, or sequencing one lane SR36 and indexing read, divided by 2152 libraries.
[b]$3/sample for personnel time (assuming salary and benefits of $70,000 per year and processing five 96-well plates/week).
[c]$3/sample for capital equipment (assuming purchase of a Covaris LE220 instrument, a PCR machine, and an Agilent Bravo liquid handling platform, and dividing over 100,000 libraries).

To demonstrate that pooled target capture using our libraries is amenable to an experiment on the scale that is relevant to medical genetic association studies, we used the library preparation method to prepare 2152 DNA samples from one population (African-Americans) in the space of 2 mo. We normalized these samples to the lowest concentrated sample in each pool, combined them into 15 pools of between 138 and 144 samples, and enriched these 15 pools for the 2.2-Mb target. We sequenced the captured products on a HiSeq 2000 instrument using 75-bp paired-end reads to an average coverage of 4.1 in nonduplicated reads (data not shown). The duplication rate of the reads was on average 72%, an elevation above the levels reported in Table 2 and Supplemental Table S2 that we hypothesize is due to dilution to the lowest-complexity library within the pools. We were able to solve this problem by replacing the dilution with a cherry-picking approach that combines samples of similar complexity. We tested this approach by pooling 81 prostate cancer libraries with similar complexity (allowing

the 95-sample pool (unnormalized before hybrid capture), $f_2 = 93\%$ of samples had a mean target coverage of within a factor of 2 of the median, $f_{1.5} = 67\%$ within a factor of 1.5 of the median, and the coefficient of variation (standard deviation divided by mean coverage) was CV = 0.40. For the three smaller pools where normalization was performed, coverage was in general more uniform: For the pool of 14, $f_2 = 93\%$, $f_{1.5} = 86\%$, CV = 0.66; for the pool of 28, $f_2 = 100\%$, $f_{1.5} = 96\%$, CV = 0.19; and for the pool of 52, $f_2 = 100\%$, $f_{1.5} = 94\%$, CV = 0.19 (Supplemental Table S2). In the 95-sample experiment, the percentage of selected bases, defined as "on bait" or within 250 bp of either side of the baits, was 70%–79% across samples (Table 2; Supplemental Table S2), comparable to the literature for single-sample selections (Supplemental Table S3). Results on the 95-sample pool are as good as the 14-, 28-, and 52-sample pools.

no more than a 5× difference in molecule count per library), resulting in a duplication rate of 24% on average at 7× coverage (Supplemental Table S2e).

The experiment was highly sensitive for detecting polymorphisms in the targeted regions. After restricting to sites with at least one-fourth of the average coverage, we discovered 35,211 polymorphisms at high confidence (10,000:1 probability of being real based on their quality score from BWA). This is more than double the 16,457 sites discovered by the 1000 Genomes Project in 167 African ancestry samples over the same nucleotides (February 2011 data release) (The 1000 Genomes Project Consortium 2010). Exploring this in more detail, we found that we rediscovered 99.7% of sites in the 1000 Genomes Project with minor allele frequency >5% and 83% of 1000 Genomes Project sites with lower frequency in the African samples. As a second measure of the quality of our data, we

**Table 2.** Sequencing results

| Application | Number of libraries | Input DNA (μg) | Normalization strategy | PF reads per library | % Reads aligning to reference genome | % Duplicated reads (removed) | Mean target coverage per library[a] | % Selected bases[b] | % Target with 2× coverage[a] |
|---|---|---|---|---|---|---|---|---|---|
| Human hybrid selection[c] | 14 | 0.6–0.9 | Dilution | $2.8 \times 10^5$ | 73 | 53.6 | 0.9 | 78 | 23 |
| Human hybrid selection[c] | 28 | 0.2–0.9 | Dilution | $3.3 \times 10^5$ | 72 | 56.4 | 1.1 | 76 | 31 |
| Human hybrid selection[c] | 52 | 0.3–0.9 | Dilution | $2.7 \times 10^5$ | 74 | 51.1 | 1.1 | 78 | 29 |
| Human hybrid selection[d] | 95 | 0.2–4.8 | Unnormalized | $1 \times 10^6$ | 89 | 37.5 | 7.4 | 74 | 79 |
| Human hybrid selection[d] | 81 | 0.6–2.6 | Cherry picking | $5.6 \times 10^5$ | 92 | 24.4 | 7.1 | 92 | 87 |
| Human whole-genome shotgun[e] | 40 | 0.75 | | $7.1 \times 10^7$ | 95 | 14.4 | 5.4 | n/a | n/a |
| Microbial sequencing[d] | 12 | 1 | | $7.2 \times 10^6$ | 97 | 1 | 147 | n/a | n/a |

[a]Target for the hybrid selection experiment is defined as the regions where baits were designed.
[b]"Selected bases" is defined as in Picard as 250 bp on either side of the bait (target).
[c]36 cycles of single-read sequencing on GAII.
[d]50 cycles of paired-end sequencing on HiSeq2000.
[e]100 cycles of paired-end sequencing on HiSeq2000; four libraries were prepared for each of 10 samples.

compared $n = 1642$ African-American samples that had previously been genotyped on an Illumina 1M array at 1367 SNPs that overlapped between that array and the 2.2-Mb target region of the capture experiments. We found that 99.77% of the mapped deduplicated reads are consistent with the "gold standard" results from genotyping. As a third measure of data quality, we checked for a potential reference bias by counting the reads matching the reference and variant allele at the 1367 SNPs where we knew the true genotypes. As shown in Supplemental Figure S2, there is a slight bias ($N_{ref}/N_{tot} = 1{,}289{,}080/2{,}537{,}488 = 50.8\%$) for the reference allele, which is sufficiently small that we do not expect it to cause a major problem for most applications such as identification of heterozygous sites.

## Application 2: Whole-genome sequencing of 40 human libraries to 5× coverage

Whole-genome shotgun sequencing (WGS) of mammalian genomes to high coverage (e.g., 30×) is still a process that is dominated by sequencing costs. However, lighter sequencing is of interest for some applications. For example, Genomewide Association Studies (GWAS), which have discovered more than 1300 associations to human phenotypes (Manolio 2010), cost hundreds of dollars per sample on SNP arrays, which is less than commercial costs of library preparation, and hence sequence-based GWAS are not economical. However, the situation would change if library production costs were lower. If libraries were inexpensive, sequencing the genome to light coverage followed by imputing missing data using a reference panel of more deeply sequenced or genotyped samples, in theory would allow more cost-effective GWAS (Li et al. 2011). With sufficiently low library production costs, sequencing may begin to compete seriously with SNP array–based analysis for medical genetic association studies, as is already occurring in studies of gene expression analysis, where RNA-seq is in the process of replacing array-based methods (Majewski and Pastinen 2010).

To test if our method can produce libraries appropriate for whole-genome sequencing, we prepared 40 libraries using an earlier version of our protocol that used microTUBES for shearing instead of plates and a slightly different enrichment PCR procedure (Supplemental Fig. S3). (A more up-to-date protocol, which involves shearing in plates and which we used to produce libraries for the prostate cancer study, further reduces costs by about $5 per sample.) Table 2 and Supplemental Table S4 show the results of sequencing these libraries to an average of 5.4× coverage using 100-bp paired-end reads on 58 lanes on Illumina HiSeq 2000 instruments. A high proportion (95%) of the reads align to the human reference genome (hg19) using BWA (Li and Durbin 2009), and duplicates were removed. We found that 99.86% of the mapped reads are concordant with the "gold standard" SNP array data previously collected on these samples (Li et al. 2008) (sequences with quality ≥30 for the 40 libraries were compared at 585,481 SNPs). Thus, we have demonstrated that our protocol can produce libraries that are useful for low-pass whole-genome human sequencing.

## Application 3: Sequencing of 12 *Escherichia coli* strains to 150× coverage

An important application of high-throughput sequencing is the study of microbial genomes, for example, in an epidemiological context where it is valuable to study strains from many patients to study the spread of an epidemic, or in the same individual to study

the evolution of an infection. Microbial genomes are small so that the required amount of sequencing per sample can be small, and thus the limiting cost is often sample preparation. To explore the utility of our library preparation protocol for microbial sequencing, we produced libraries for 12 *E. coli* strains for a project led by M. Lajoie, F. Isaacs, and G. Church (whom we thank for allowing us to report the data) (Isaacs et al. 2011). We produced these libraries as a single row on a 96-well plate with an input DNA amount of 1 μg together with human libraries that we were producing for another study following the protocol in Supplemental Figure S4. Table 2 and Supplemental Table S5 report the results of the sequencing of these 12 libraries on a single lane of a HiSeq 2000 (50-bp paired-end reads). We analyzed the data after separating the libraries by sample using internal barcodes and mapping to the *E. coli* reference (strain K12 substrain MG1655, Refseq NC_000913) using BWA (Li and Durbin 2009). Overall, 97% of reads mapped, with an average of 147-fold coverage and 1% duplicated reads.

## Discussion

We have reported a high-throughput library preparation method for next-generation sequencing, which has been designed to allow an academic laboratory to generate thousands of barcoded libraries at a cost that is one to two orders of magnitude less than the commercial cost of library preparation. These libraries are appropriate for whole-genome sequencing of large and small genomes. A particularly important feature of these libraries is that they are effective for pooling approximately a hundred samples together and enriching them for a subset of the genome of interest. We have proven that the method is practical at a scale that is relevant to medical genetics by generating more than 2000 libraries for a prostate cancer study, enriching them for more than 2 Mb of interest, and obtaining sequencing data that are concordant with previously reported genotype calls.

From an engineering point of view, our method was designed with a different set of goals than have driven most previous library preparation methods. In most methods, the emphasis has been on producing libraries with maximal complexity (as measured by the number of unique molecules) and length uniformity (as measured by the tightness of the distribution of insert sizes) given the large amount of sequencing that was planned for each library. Our goal is different: to increase throughput and decrease reagent cost, while building libraries that are appropriate for pooled target capture. In this study, we empirically show that the human libraries produced by our method are complex enough that when shotgun-sequenced to a coverage of around 5×, they give duplication rates of 9%–20%. This duplication rate is somewhat higher than some published protocols, and the problem of duplication becomes greater as coverage increases, so that for deep-sequencing studies (e.g., whole-genome sequencing at 30×) in which thousands of dollars are invested per sample, it may be more economical to use a more expensive library preparation protocol that minimizes duplication rates. One reason for an increased duplication rate in our libraries is our distribution of fragment insert sizes. Because size selection with beads is not as tight as gel-based size selection, fragment insert sizes of the libraries produced with our protocol are variable. Longer fragments are more prone to duplicated reads ("optical duplicates"), in which the Illumina software identifies one cluster as two adjacent clusters. Another reason for an increased duplication rate is the low input DNA amount per ligation reaction (0.75 μg for each of the four ligation

# DOCKET ALARM

# Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

## Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

## API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

### LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

### FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

### E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.

fastcase
Smarter legal research.