



# Rapid flow-sorting to simultaneously resolve multiplex massively parallel sequencing products

Julia Sandberg<sup>1</sup>, Beata Werne<sup>1</sup>, Mark Dessing<sup>2</sup> & Joakim Lundeberg<sup>1</sup>

<sup>1</sup>Science for Life Laboratory, Department of Gene Technology, Royal Institute of Technology (KTH), Box 1031, SE-17121, Solna, Sweden, <sup>2</sup>BD Biosciences, Tullastrasse 8-12, 69126 Heidelberg, Germany.

Sample preparation for Roche/454, ABI/SOLiD and Life Technologies/Ion Torrent sequencing are based on amplification of library fragments on the surface of beads prior to sequencing. Commonly, libraries are barcoded and pooled, to maximise the sequence output of each sequence run. Here, we describe a novel approach for normalization of multiplex next generation sequencing libraries after emulsion PCR. Briefly, amplified libraries carrying unique barcodes are prepared by fluorescent tagging of complementary sequences and then resolved by high-speed flow cytometric sorting of labeled emulsion PCR beads. The protocol is simple and provides an even sequence distribution of multiplex libraries when sequencing the flow-sorted beads. Moreover, since many empty and mixed emulsion PCR beads are removed, the approach gives rise to a substantial increase in sequence quality and mean read length, as compared to that obtained by standard enrichment protocols.

Current methods for massively parallel sequencing produce enormous amounts of data, and as the throughputs of sequencing instruments continue to rise the possibilities and need to expand the degree of multiplexing follows<sup>1-4</sup>. Also, it has been proposed that multiplexing of sequencing libraries should be included in NGS experimental designs to minimize potential problems such as miss-assignment and be used as a means to validate the output<sup>5,6</sup>. Multiplexing procedures require accurate quantification of the individual targets before pooling, and a number of techniques are used for this purpose. However, despite these precautions highly biased fractions of constituent libraries have been reported<sup>7,8</sup>.

Suggested explanations for this bias include the effects of variations in GC-content, fragment length distributions and DNA quality on amplification efficiency<sup>9</sup>. However, regardless of the cause(s) it is clearly important to determine the optimal proportions of DNA from different libraries to include in high-throughput next-generation sequencing runs. In emulsion PCR based sequencing systems such as Roche/454, Life Technologies/SOLiD and Life Technologies/Ion Torrent, this can be done by separately sequencing representative small-scale emulsion PCR samples with different amounts of input DNA or (more often nowadays) by counting the number of beads enriched from such titration samples<sup>10,11</sup>.

Clearly, as multiplicity increases, the feasibility of performing individual titration runs for each barcoded library decreases. As an alternative to titration, amplification efficiencies of individual libraries can be evaluated prior to pooling and emulsion PCR, by quantitative<sup>12,13</sup> or digital<sup>14</sup> PCR. However, these approaches increase the workload in a linear fashion and do not completely mimic multiplex amplification, notably optimized amounts of particular libraries determined by qPCR may be far from ideal when the libraries are being simultaneously, competitively amplified<sup>7</sup>. Multiplex amplicon sequencing is also subject to biases, similar to those that affect shotgun procedures, which is of particular concern when the strategy is applied to samples representing a few, or single, cells<sup>2,15</sup>.

To address these problems we have developed a protocol for simultaneous normalization of multiplex products carrying different barcodes based on the Roche/454 sequencing platform. The procedure extends a technique we previously described — flow cytometric analysis and sorting of emulsion PCR beads<sup>10,16</sup> — by applying color-coded labeling to differentiate between multiplex reaction products. For this purpose, barcoded fragments from multiple targets are fluorescently tagged, thus enabling normalization of sequence read frequencies of the libraries, with simultaneous enrichment of DNA-covered beads. We report here a proof-of-principle experiment demonstrating that an amplified sample containing four multiplexed libraries with uneven frequencies can be flow-sorted at high speed, to obtain an even distribution of sequence reads from all four DNA libraries. Moreover, we show that the flow cytometric sorting protocol results in considerable increases in both sequence read quality and average read lengths.

## SUBJECT AREAS:

GENOMICS

DEVELOPMENT

MOLECULAR SEQUENCE DATA

METHODS

Received  
6 May 2011

Accepted  
22 September 2011

Published  
6 October 2011

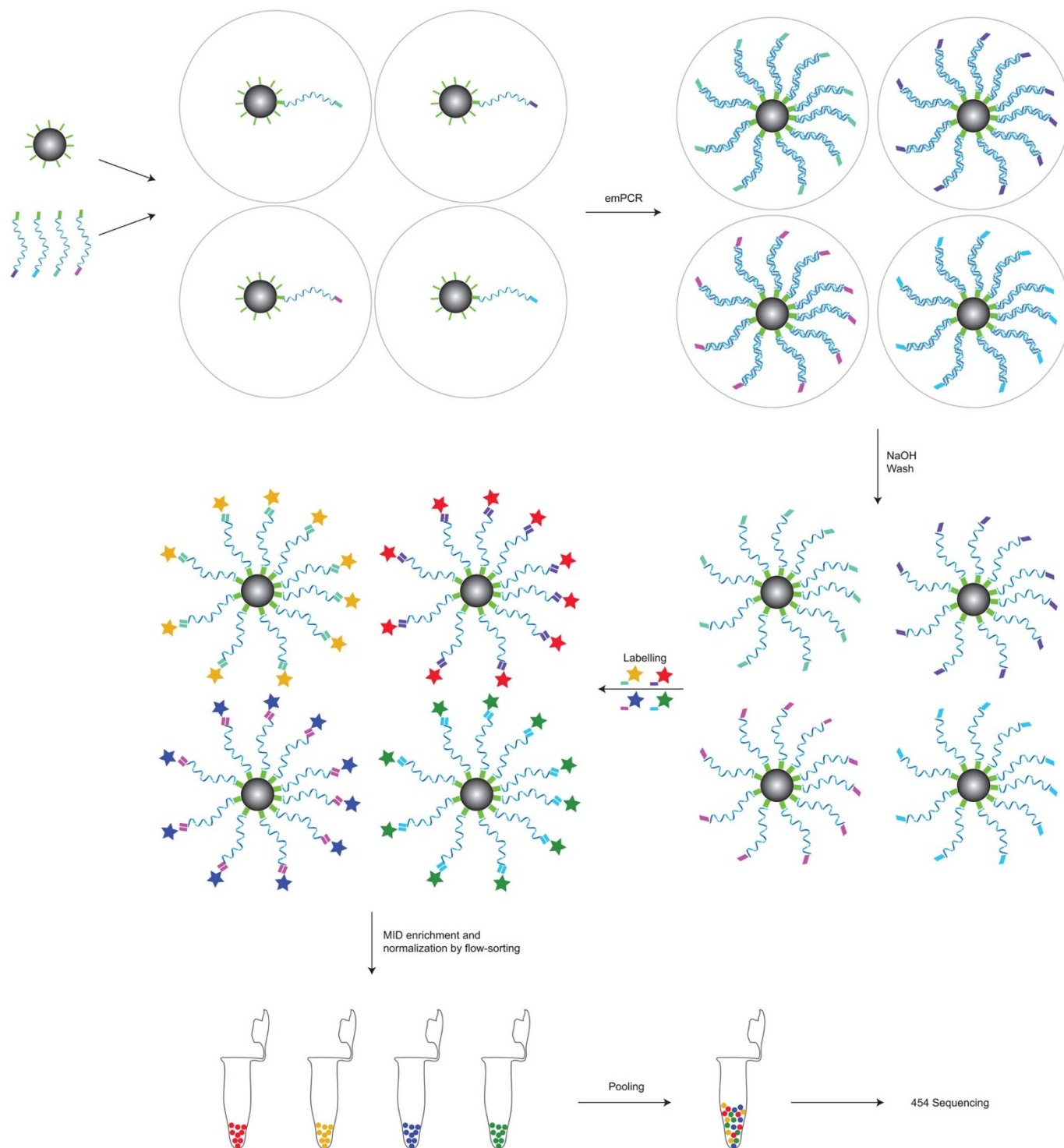
Correspondence and  
requests for materials  
should be addressed to  
J.L. (joakim.lundeberg@  
scilifelab.se)



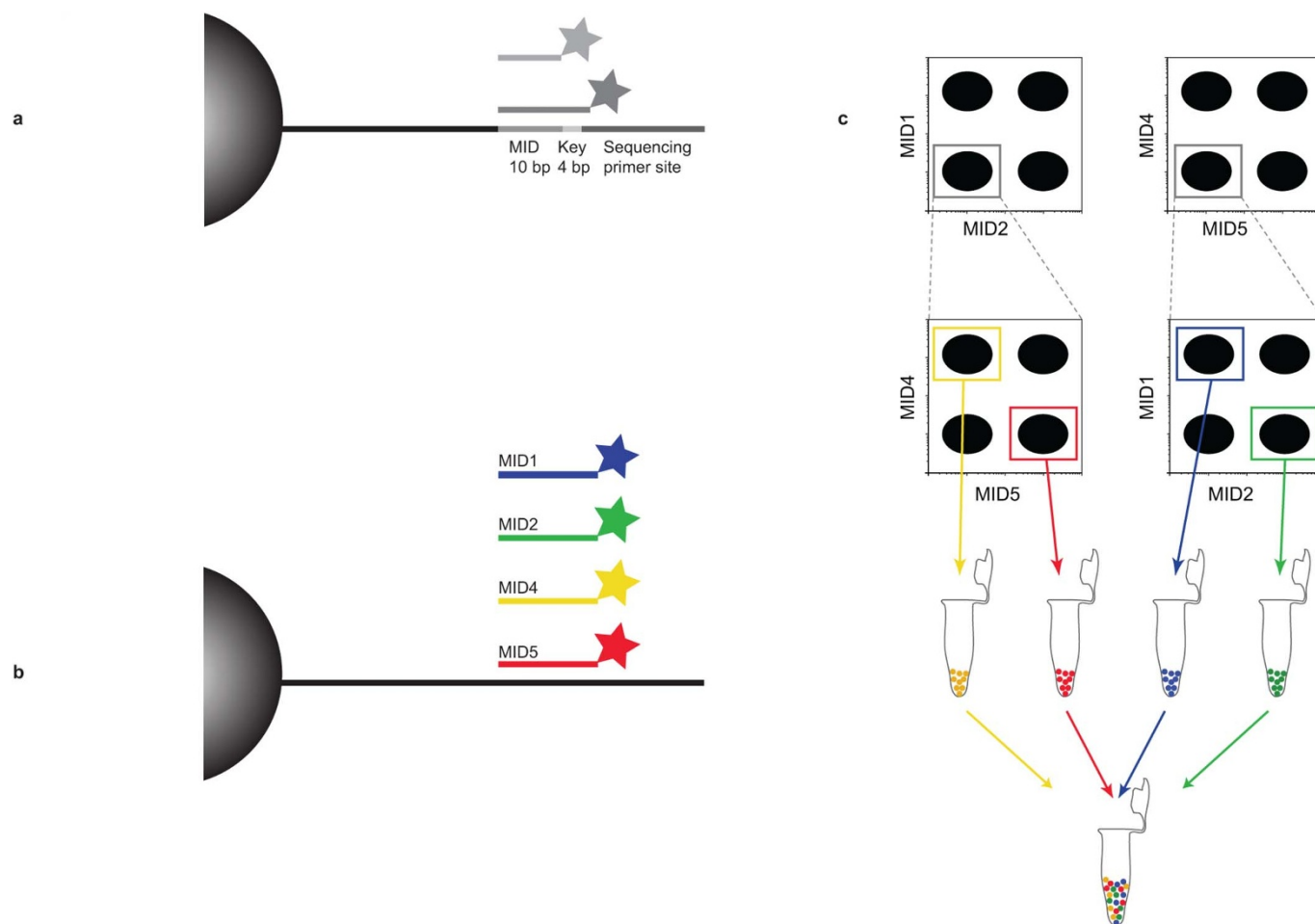
## Results

We designed a protocol for barcode-specific enrichment of emulsion PCR beads, enabling normalization of different libraries. The technique is outlined in **Figure 1**. Complementary oligonucleotide probes for samples carrying specific Multiplex Identifier (MID)-tags for Roche/454 sequencing were designed. The length of these tags is 10 bp, which is too short to give efficient labeling for flow cytometry. Thus, to obtain probes that were sufficiently long for robust

hybridization we used probes that overlapped the common 4 bp key sequence and 2 bp of the general sequencing primer site, as illustrated in **Figure 2a**. This gave us 16 bp long unique, color-coded oligonucleotides that could be used for labeling barcoded DNA-capture beads, with good resolution. MID3 was omitted based on earlier findings that this MID sequence considerably decreased the amplification efficiency of its library<sup>7</sup>. Hence, probes were designed to match MID 1, 2, 4 and 5. Color-coded probes are depicted in **Figure 2b**.



**Figure 1 | Schematic illustration of the workflow for barcode normalization by flow-sorting.** DNA libraries are pooled and then PCR-amplified on the surface of DNA-capture beads in emulsion. The emulsion is broken, beads are collected and amplified DNA is made single-stranded by alkali washing. Beads are labeled with color-coded probes that hybridize to unique barcode sequences. Equal numbers of beads with each barcode are then collected by flow cytometric sorting. Beads are pooled, washed and sequenced by Roche/454 sequencing.



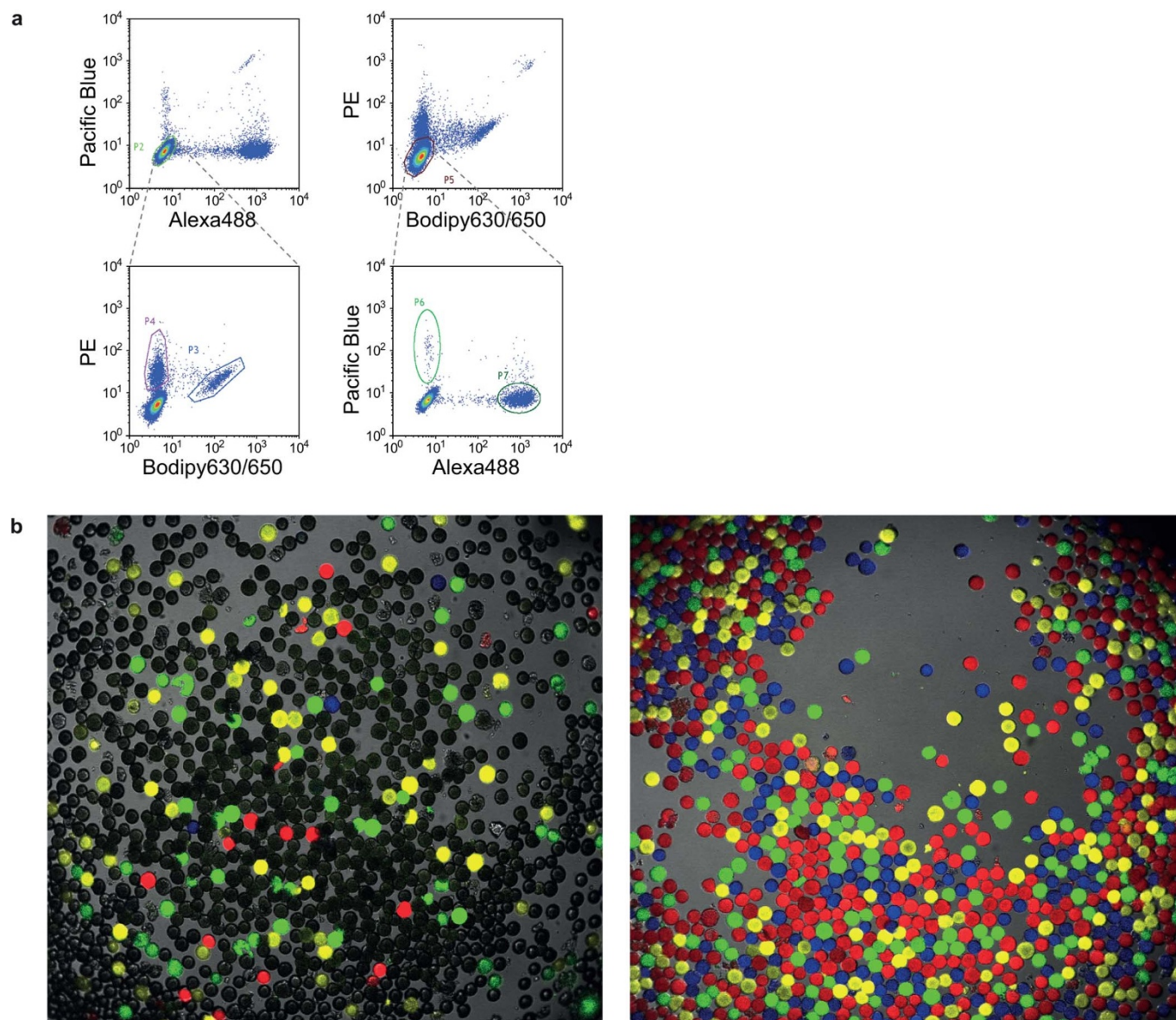
**Figure 2 | Labeling and gating strategy for barcode normalization by flow-sorting.** (a) Schematic representation of oligonucleotides used for labeling beads after emulsion PCR. Sixteen bp probes gave better resolution in flow cytometry than shorter probes (10 bp), and were therefore used in the experiments described here. (b) Schematic representation of MID-specific color-coded probes used to differentiate beads carrying amplified fragments of different barcoded libraries. (c) Schematic drawing of flow-sorting setup for a sample containing four different MID libraries. To collect only beads carrying library fragments with one barcode labeling, gates were set in a Boolean fashion so that beads yielding positive signals in only one fluorescent channel were collected.

Labeled beads were flow-sorted in a Boolean manner, using linked sorting gates to ensure that collected beads were positively labeled with one fluorescent dye and yielded only background signals in the other fluorescent channels (Fig. 2c). This was done to remove beads carrying more than one amplified barcode on the surface, and thus avoid loading these unreadable mixed beads in the sequencing plate. Equal amounts of each type of MID-tag bead were collected by flow-sorting (Fig. 3a), and re-analysis of aliquots of the sorted bead samples showed purities of approximately 98%. Collected beads were pooled and labels were removed by alkali washing. Figure 3b show micrographs of labeled beads before and after flow-sorting.

Flow-sorted and ‘standard-enriched’ beads were sequenced in parallel using the Roche/454 platform, and the barcode frequencies determined by sequencing and flow cytometry, were compared. Flow cytometric analysis of the original, labeled, un-enriched emulsion PCR sample showed that it contained extremely biased proportions of the four MID libraries, since MID 1, 2, 4 and 5 were detected on 1.4, 52.2, 32.9 and 13.5% of DNA-covered beads, respectively (Fig. 4a I, left plot). Sequencing of beads enriched using a standard Roche/454 protocol yielded a sequence read distribution that was very similar to that obtained by flow cytometric analysis, with MID 1, 2, 4 and 5 tag proportions of 1.4, 48.6, 38.5 and 11.6%, respectively (Fig. 4a I, middle plot). In contrast, sequencing of the flow-sorted sample indicated that it contained similar proportions of

the four libraries, with MID 1, 2, 4 and 5 tag frequencies of 27.9, 24.2, 24.2 and 23.8%, respectively (Fig. 4a I, right plot). The extremely skewed starting distribution had thus been normalized and the frequency of the very rare MID1 members had been increased from 1.4% to 27.9%.

Important differences in sequence quality were also observed between the flow-sorted and the ‘standard-enriched’ sample. The flow-sorting protocol resulted in a significantly increased proportion of high quality reads (58.8% in the ‘standard-enriched’ sample versus 83.3% in the flow-sorted sample) as illustrated in Figure 4aII. To investigate if it was in fact the increased relative amount of the low cpb libraries in the flow-sorted sample that gave rise to the increased sequence quality in the flow-sorted sample individual Phred quality scores of the four MID libraries were compared, for the flow-sorted and for the standard enriched sample. We saw an elevated sequence quality for all four MID libraries in the flow-sorted sample, however there was no large difference evident between the different MID libraries. In the ‘standard-enriched’ sample average Phred quality scores of 27.7, 28.0, 27.3 and 27.1 were seen for MID 1, 2, 4 and 5, respectively. In the flow-sorted sample the corresponding values were 29.2, 29.2, 28.7 and 28.3. To make sure that the observed differences in sequence quality were not due to variation between the sequencing lanes, sequence qualities of control beads were compared. The proportion of control bead sequence reads at 400 bp, which had a



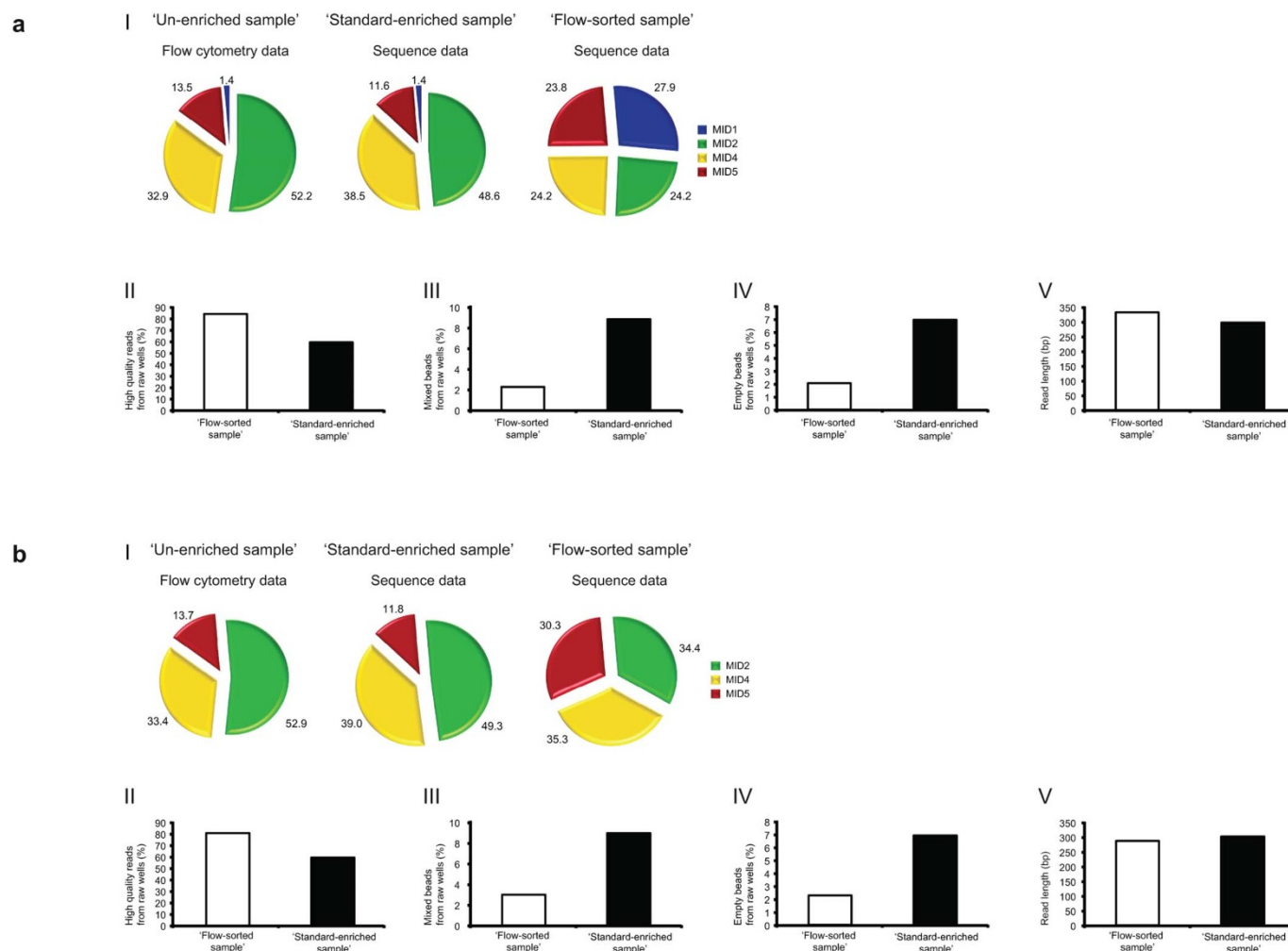
**Figure 3 | Flow-sorting of labeled beads.** (a) Flow cytometry data from sorting of barcode-specifically labeled libraries. Equal amounts of single channel positive beads were sorted out for subsequent sequencing, using gates P3, P4, P6 and P7. (b) Confocal micrographs of beads before (**left panel**) and after (**right panel**) flow cytometric normalization.

sequence that mapped with 98% accuracy to the control sequences, were 15, 19 and 22% respectively for the flow-sorted 4-plex, 3-plex and ‘standard-enriched’ sample. It appears that the sequencing reaction was in fact better for the lane containing the ‘standard-enriched’ sample, thus the change in sequence quality cannot be explained by lane-to-lane variation. Instead, we propose that the increase in sequence quality obtained by flow-sorting was most likely due to a combination of several observed differences between the two samples, as explained in more detail below.

As expected, frequencies of beads carrying more than one amplified template (mixed beads) were substantially decreased in the flow-sorted sample. Two kinds of mixed beads could be expected to form during emulsion PCR; ‘single-barcode mixed beads’ carrying more than one different fragment on the surface, each fragment having the same barcode and ‘multiple-barcode mixed beads’ carrying more than one different fragment and the fragments having different barcodes. The latter type, the ‘multiple-barcode mixed beads’, could be identified by multiple color labeling of these beads and could thus be removed by our protocol. The distribution of

‘single-barcode mixed beads’ and ‘multiple-barcode mixed beads’ within the total population of mixed beads depends on the fraction of each barcode in the first place, but overall makes up a large portion of the mixed bead population. We observed a decrease in mixed beads from 8.8% in the ‘standard-enriched’ sample to 2.2% in the flow-sorted sample (Fig. 4a III) corresponding to a 75% decrease.

Moreover, the number of beads not covered with amplified DNA (empty beads) in the sequencing reaction was reduced from 6.8% in the ‘standard-enriched’ sample to 2.0% in the flow-sorted sample (Fig. 4a IV). This agrees with our previous findings, that flow cytometric sorting yields samples with higher purity than ‘standard-enrichment’<sup>10</sup>. Further, an increase in mean read length from 298 bp in the ‘standard-enriched’ sample to 330 bp in the flow-sorted sample (Fig. 4a V) was obtained. The increase in read length may be due to a combination of two features in the sorted sample; a smaller proportion of empty beads and the fact that highly fluorescent beads were collected (thus carrying a large number of amplified fragments on the surface) leading to stronger signals in the sequencing reaction



**Figure 4 | Comparison of 'standard-enriched' and flow-sorted samples.** (a) First experiment, including four barcodes. (I) Proportions of the four DNA libraries in (left chart) an un-enriched sample as determined by flow cytometric analysis, (middle chart) a 'standard-enriched' sample as determined by Roche/454 sequencing and (right chart) in a flow-sorted sample, as determined by Roche/454 sequencing. (II–V) Comparison of sequence quality in 'standard-enriched' and flow-sorted samples. Bar plots illustrate (II) percentages of high quality sequence reads from raw wells, (III) percentages of mixed sequence reads from raw wells, (IV) percentages of empty beads from raw wells and (V) average read lengths in 'standard-enriched' and flow-sorted sample. (b) Additional experiment, including three barcodes. An additional experiment was carried out using three DNA libraries with MID2, MID4 and MID5 barcodes. (I) Proportions of the three MID-libraries (left chart) before enrichment as determined by flow cytometric analysis, (middle chart) in a 'standard-enriched' sample as determined by Roche/454 sequencing and (right chart) in a flow-sorted sample as determined by Roche/454 sequencing. (II–V) Comparison of sequence quality of 'standard-enriched' and flow-sorted samples. Bar plots illustrate (II) percentages of high quality sequence reads from raw wells, (III) percentages of mixed sequence reads from raw wells, (IV) percentages of empty beads from raw wells and (V) average read lengths, in the two samples.

and (hence) greater numbers of sequencing cycles in which the threshold signal to noise ratio was exceeded.

These findings were supported in the subsequent experiment where the scarce MID1 sample was omitted and MID2, MID4 and MID5 were used. Again, the protocol resulted in equal frequencies of all barcodes, as illustrated in Figure 4b I, and a similar pattern of improvement in sequence quality was also observed, however the readlength was slightly lower (283 bp). (Fig. 4b II–IV).

## Discussion

Current next generation sequencing platforms are increasing their sequencing capacity in a steady pace. Consequently, some sequencing units have difficulties making use of the full capacity of individual sequencing runs, which results in samples being over-sequenced beyond the diversity in the initial sequencing library. This has been addressed by the launching of a set of instruments with lower capacity (such as Roche's GS Junior, Illumina's MiSeq and

Life Technologies' SOLiD PI and Ion Torrent PGM instrument) that are more directed towards low to medium throughput applications. Yet, another widely used option is to multiplex samples to combine several projects on a common sequencing lane. Hereby optimal use of the sequencing capacity can be achieved and more focus need to be devoted to these multiplexing approaches.

Here, we report for the first time a strategy to continuously resolve and normalize multiplex products for massively parallel sequencing. By barcode tagging and high-speed flow-sorting we obtained nearly equal distributions of individual libraries starting with a highly biased mixed library. In addition, we also observed a considerable increase in sequence quality in the flow-sorted samples. This is most likely due to the 75% reduction in frequencies of mixed beads and 70% reductions in frequency of empty beads, observed in flow-sorted samples.

The described protocol causes minimal loss of beads and could be used in combination with qPCR-guided pooling or as a stand-alone procedure. By applying a double or triple-sorting approach using a current 6-way sorting instrument, the described protocol can be



scaled up to a multiplicity of 11 or 16 barcodes, respectively and the sample preparation protocol could still be done in one hour<sup>17</sup>. We have shown that we can design detection probes based on index sequences in the 454/Roche system and we foresee that our set-up can be adopted by index system from manufacturers of other bead-based sequencing systems.

This approach could be of particular interest for deep sequencing of multiplex amplicons from scarce material with limited possibilities for re-sampling. Our method offers a possibility to balance-out amplification biases to allow for equal sequence coverage and statistically meaningful comparisons of multiple targets<sup>18,19</sup>, such as in multiplex mutation analysis of cancer genes or mutational hotspots.

In conclusion, we describe a rapid procedure to resolve and normalize multiplex products for massively parallel sequencing that can be used for several of the current sequencing platforms and we also report that the resulting sequences are of higher quality than those obtained by standard procedures.

## Methods

**Library preparation and emulsion PCR.** Nebulized *Chironomus tentans* DNA was CA-purified using Dynabeads® MyOne™ carboxylic acid paramagnetic beads (CA-beads; Invitrogen, Carlsbad, CA, USA), pooled and prepared pooled, then prepared by automated library preparation as previously described<sup>7</sup> using Multiplex Identifier (MID) adaptors (MID 1, 2, 4 and 5) and GS FLX Titanium kits (Roche, Basel, Switzerland). Optimal DNA-to-bead ratios of 2, 11, 12 and 5 copies per bead for MID 1, 2, 4 and 5, respectively, had previously been determined by sequencing-titration according to the manufacturer's manual, and the libraries were pooled based on these values. The pooled DNA was then subjected to emulsion PCR in eight SV (small volume) reactions using GS FLX Titanium emPCR reagents, the emulsions were broken and beads were collected, following the manufacturer's recommendations. Normally, after emulsion PCR, 80–90% of the beads in the sample have no amplified template on their surface (naked beads), while the rest are covered in amplified material (DNA-covered beads) and these are usually enriched by using streptavidin-covered magnetic enrichment beads, according to the manufacturer's instructions.

Here, after emulsion PCR the recovered beads were divided equally between two tubes. One tube was prepared for sequencing by 'standard-enrichment' of the DNA-covered beads according to the manual and the replicate bead sample was labeled and enriched by two separate flow-sorting experiments, both procedures are described below.

**Enrichment using streptavidin-covered magnetic beads.** For comparison, enrichment of DNA covered beads from one sample was carried out according to the manual. Briefly, DNA was made single-stranded by alkali washing and then enrichment primer, which is biotinylated and specific to the outermost part of the amplified DNA on the beads, was hybridized to the beads. Streptavidin-covered magnetic enrichment beads and a magnet were utilized to capture the biotinylated DNA-covered beads whereas empty beads were washed away. All washing steps were carried out on a tabletop spinner by spinning tubes for 10 seconds, turning tubes 180°, spinning again for 10 seconds and then discarding the supernatant.

**Bead labeling for flow-sorting.** To obtain sufficiently long probes for robust hybridization and labeling, oligonucleotide probes that overlapped not only the 10 bp barcode sequence but also the neighboring 4 bp common key sequence and 2 bp of the sequencing primer site were designed. This resulted in the following 16 bp oligonucleotide probes: FL\_MID1, 5' Pacific Blue-ACTCAGACGAGTGGC 3', (Thermo Fisher Scientific, Waltham, USA); FL\_MID2, 5' Alexa488-ACTCAGACGCTCGACA 3' (Eurofins MWG Operon, Ebersberg, Germany); FL\_MID4, 5' biotin-ACTCAGACTGTAG (Eurofins MWG Operon); and FL\_MID5, 5' BODIPY630/650-ACTCAGATCAGACACG (Eurofins MWG Operon). We have in previous publications demonstrated that FACS enrichment does not affect the GC content or sample complexity, in FACS sorted bead samples<sup>10,16</sup>.

Beads were prepared for fluorescent labeling by washing twice in 100 µl 0.1 M NaOH and twice in 100 µl 1xAnnealing Buffer (1xAB, Roche), as described above. They were then re-suspended in 100 µl 1xAB and labeled by adding 48 µl of a mixture of FL\_MID1, FL\_MID2, FL\_MID4 and FL\_MID5 probes (10 µM each). The sample was incubated at 95°C for 1 min and 65°C for 1 min, then cooled by incubation at room temperature for 2 min and finally placed on ice for 5 min. Beads were then washed and re-suspended in 100 µl 1xAB and 9 µl streptavidin R-phycoerythrin conjugate (Invitrogen, Paisley, UK, 1mg/ml) was added. After 30 min of incubation at room temperature in the dark, samples were washed in phosphate-buffered saline (PBS) containing 0.1% Pluronic F108 NF surfactant (PBSP) (BASF Corporation, Mount Olive, NJ, USA) and transferred to 5 ml polystyrene round-bottom FACS tubes (BD Biosciences, Franklin Lakes, NJ, USA). Flow-sorting was done in a final volume of 7 ml PBSP to prevent the beads aggregating.

To obtain reference points for setting the flow cytometry gates, single dye positive and negative control samples were prepared for each of the four fluorescent dyes used. Negative control samples were prepared by simply labeling unreacted beads with the FL\_MID- probes described above, to obtain indications of the level of unpecific

labeling of the beads. Positive control samples were prepared by labeling unreacted DNA capture beads using a probe that was specific to the capture probe on the surface of these unreacted beads (5' modification-AGGCACACAGGGGATAGG 3'). Probes were conjugated with the abovementioned fluorescent dyes, ordered from the same companies, as stated above. To each sample, 6 µl DNA capture beads, 6 µl probe (10 µM) and 25 µl 1xAB were added. In addition, a multi-color negative control sample was prepared by adding 6 µl unreacted DNA capture beads, 25 µl 1xAB and 6 µl of the mixture of FL\_MID1, FL\_MID2, FL\_MID4 and FL\_MID5 probes (10 µM each). The nine samples were heated and washed as described above, but here 2 µl streptavidin R-phycoerythrin conjugate (Invitrogen, 1 mg/ml) was added, labeling was done in a final volume of 25 µl 1xAB, and the final volume for flow cytometric sorting was 500 µl PBSP. The single-dye positive and negative control samples are only needed once, to obtain indications of the differences in signal intensities between the positive and negative samples. In subsequent flow-sorting experiments the multi-color negative control sample is sufficient to adjust the gates in flow cytometry.

**MID-specific flow-sorting.** Beads were sorted using a BD Influx Cell Sorter (BD Biosciences, San Jose, CA) fitted with a 200 µm nozzle, operating at 9 psi pressure and 6.9 kHz, with a sort rate of 2000 events per second. Doublets were excluded by plotting forward scatter pulse area against forward scatter pulse width. Initially, gates were adjusted using the single-dye positive and negative control samples for each fluorophore followed by the multi-color negative control sample. Sorting of the true sample was done using linked gates to ensure that sorted beads were positively labeled with only one dye. Equal numbers of the different bead populations were flow-sorted into separate eppendorf tubes containing 100 µl PBSP. Since there are so many amplified DNA copies on each bead (e.g. thousands to millions), and the detection limit for most flow cytometer instruments is 10–100 fluorophores, the positive fluorescent labeling is orders of magnitude stronger than that from negative beads and can easily be distinguished from negative labeling.

In a first experiment 7253 beads of each type were collected by flow-sorting. Due to the low frequency of Pacific Blue-labeled (MID1) beads, another sample was prepared by omitting the MID1 beads and collecting 42747 each of the MID2-, MID4- and MID5-positive beads. Aliquots of the sorted bead samples were re-analyzed for purity evaluation using a LSRFortessa cell analyzer (BD Biosciences).

**Microscopy of labeled beads.** Images of labeled beads were acquired using a Leica SP5 confocal microscope (Mannheim, Germany) equipped with a 10x objective.

**Sample preparation and sequencing.** Flow-sorted and 'standard-enriched' beads were sequenced using the GS FLX 454 (Roche) platform, according to the manufacturer's instructions. To remove labeling probes from flow-sorted beads these were washed three times in 0.1M NaOH and then three times in 1xAB with centrifugation steps between washes, as described above. In previous publications we have demonstrated that alkali wash is sufficient to remove the fluorescent labeling<sup>10</sup>. Sequencing primer was annealed to the beads, then the 'standard-enriched' sample and the two flow-sorted samples were loaded onto separate lanes of a PicoTiter plate and sequenced following the manufacturer's instructions.

- Lennon, N. J. *et al.* A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. *Genome biology* **11**, R15 (2010).
- Parameswaran, P. *et al.* A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res* **35** (2007).
- Andersson, A. F. *et al.* Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS One* **3** (2008).
- Neiman, M., Lundin, S., Savolainen, P. & Ahmadian, A. Decoding a substantial set of samples in parallel by massive sequencing. *PLoS One* **6** (2011).
- Auer, P. L. & Doerge, R. W. Statistical design and analysis of RNA sequencing data. *Genetics* **185**, 405–416 (2010).
- Binladen, J. *et al.* The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS One* **2**, e197 (2007).
- Lundin, S., Stranneheim, H., Pettersson, E., Klevebring, D. & Lundeberg, J. Increased throughput by parallelization of library preparation for massive sequencing. *PLoS One* **5**, e10029, doi:10.1371/journal.pone.0010029 (2010).
- Steuernagel, B. *et al.* De novo 454 sequencing of barcoded BAC pools for comprehensive gene survey and genome analysis in the complex genome of barley. *BMC Genomics* **10**, 547 (2009).
- Shendure, J. *et al.* Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005).
- Sandberg, J., Stahl, P. L., Ahmadian, A., Bjursell, M. K. & Lundeberg, J. Flow cytometry for enrichment and titration in massively parallel DNA sequencing. *Nucleic Acids Res* **37** (2009).
- Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005).
- Meyer, M. *et al.* From micrograms to picograms: quantitative PCR reduces the material demands of high-throughput sequencing. *Nucleic Acids Res* **36** (2008).
- Zheng, Z. *et al.* Titration-free massively parallel pyrosequencing using trace amounts of starting material. *Nucleic Acids Res* **38** (2010).
- White, R. A., 3rd, Blainey, P. C., Fan, H. C. & Quake, S. R. Digital PCR provides sensitive and absolute calibration for high throughput sequencing. *BMC Genomics* **10**, 116 (2009).



15. Galan, M., Guivier, E., Caraux, G., Charbonnel, N. & Cosson, J. F. A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. *BMC Genomics* **11**, 296 (2010).
16. Sandberg, J., Neiman, M., Ahmadian, A. & Lundeberg, J. Gene-specific FACS sorting method for target selection in high-throughput amplicon sequencing. *BMC Genomics* **11**, 140 (2010).
17. Peretto, S. P., Chattopadhyay, P. K. & Roederer, M. Seventeen-colour flow cytometry: unravelling the immune system. *Nature reviews Immunology* **4**, 648–655 (2004).
18. Stahl, P. L. *et al.* Sun-induced nonsynonymous p53 mutations are extensively accumulated and tolerated in normal appearing human skin. *J Invest Dermatol* **131**, 504–508 (2011).
19. Wang, D. & Bodovitz, S. Single cell analysis: the new frontier in 'omics'. *Trends Biotechnol* **28**, 281–290 (2010).

## Acknowledgements

The authors wish to thank Afshin Ahmadian, Patrik Stahl and Sverker Lundin, for helpful discussions, Martin Hjemare for assistance with microscopy and Kristina Holmberg, Christian Natanaelsson and Miriam Landfors for carrying out the sequencing work. We would also like to thank the Science for Life Laboratory and SNISS for providing the infrastructure and essential help with sequencing.

## Author contributions

JL and JS designed the study. BW performed emulsion PCR. MD and JS performed FACS sorting. JS analyzed the data. JL and JS wrote the manuscript. All authors read and approved the final manuscript.

## Additional information

**Funding:** This work was supported by the Swedish Scientific Council and Knut and Alice Wallenberg foundation.

**Competing financial interests:** MD is a full-time employee at BD Biosciences. Development and optimization of the technique was done with no awareness or involvement of MD. The last flow-sorting experiment reported in this paper, was carried out using a BD Influx Cell Sorter with the assistance of MD. The other authors have declared no conflicts of interest.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

**How to cite this article:** Sandberg, J., Werne, B., Dessing, M. & Lundeberg, J. Rapid flow-sorting to simultaneously resolve multiplex massively parallel sequencing products. *Sci. Rep.* **1**, 108; DOI:10.1038/srep00108 (2011).