

Video Article

Flow-sorting and Exome Sequencing of the Reed-Sternberg Cells of Classical Hodgkin Lymphoma

Jonathan B. Reichel¹, Jason McCormick², Jonathan R. Fromm³, Olivier Elemento⁴, Ethel Cesarman⁵, Mikhail Roshal⁶

¹Innovation Laboratory, Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center

²Flow-Sorting Core Facility, Weill Cornell Medical College

³Department of Laboratory Medicine, University of Washington

⁴Department of Physiology and Biophysics, Institute for Computational Biomedicine, Weill Cornell Medical College

⁵Department of Pathology and Laboratory Medicine, Weill Cornell Medical College

⁶Department of Laboratory Medicine, Memorial Sloan Kettering Cancer Center

Correspondence to: Mikhail Roshal at roshalm@mskcc.org

URL: <https://www.jove.com/video/54399>

DOI: [doi:10.3791/54399](https://doi.org/10.3791/54399)

Keywords: Genetics, Issue 124, Classical Hodgkin lymphoma, cHL, Reed-Sternberg, FACS, exome, next-generation sequencing library, flow cell-sorting

Date Published: 6/10/2017

Citation: Reichel, J.B., McCormick, J., Fromm, J.R., Elemento, O., Cesarman, E., Roshal, M. Flow-sorting and Exome Sequencing of the Reed-Sternberg Cells of Classical Hodgkin Lymphoma. *J. Vis. Exp.* (124), e54399, doi:10.3791/54399 (2017).

Abstract

The Hodgkin Reed-Sternberg cells of classical Hodgkin lymphoma are sparsely distributed within a background of inflammatory lymphocytes and typically comprise less than 1% of the tumor mass. Material derived from bulk tumor contains tumor content at a concentration insufficient for characterization. Therefore, fluorescence activated cell sorting using eight antibodies, as well as side- and forward-scatter, is described here as a method of rapidly separating and concentrating with high purity thousands of HRS cells from the tumor for subsequent study. At the same time, because standard protocols for exome sequencing typically require 100-1,000 ng of input DNA, which is often too high, even with flow sorting, we also provide an optimized, low-input library construction protocol capable of producing high-quality data from as little as 10 ng of input DNA. This combination is capable of producing next-generation libraries suitable for hybridization capture of whole-exome baits or more focused targeted panels, as desired. Exome sequencing of the HRS cells, when compared against healthy intratumor T or B cells, can identify somatic alterations, including mutations, insertions and deletions, and copy number alterations. These findings elucidate the molecular biology of HRS cells and may reveal avenues for targeted drug treatments.

Video Link

The video component of this article can be found at <https://www.jove.com/video/54399/>

Introduction

Advancements in cancer genomics as a result of next-generation sequencing have led to significant breakthroughs in the identification of therapeutic targets and in prognostication for many hematologic and non-hematologic neoplasms. New individualized treatment strategies based on specific genomic alterations are rapidly being introduced in many tumor types (reviewed in references^{1,2}). Despite significant progress in lymphoma genomics, the genome of the neoplastic HRS cells in classical Hodgkin lymphoma (CHL) had been underexplored. The investigations have been hampered by the scarcity of neoplastic HRS cells within a reactive microenvironment, making it difficult to isolate purified HRS cell populations³.

The method for isolating viable HRS cells from primary tumors was developed by Fromm *et al.*^{4,5,6}. The method uses an eight-antibody cocktail, consisting of CD30, CD15, CD40, CD95, CD45 CD20, CD5, and CD64, to unequivocally identify HRS cells from a CHL tumor suspension. Using this methodology, we are able to isolate at least 1,000 viable HRS cells from fresh or frozen cell suspensions from tumor biopsies consisting of at least 10⁷ cells (approximately 10 mg of tissue). The purity is greater than 90% by flow cytometric analysis and is estimated to be at least 80% by exome genomic analysis of ten consecutive cases.

We have refined a flow cytometric cell isolation technique that has greatly eased the process, allowing for the rapid isolation of thousands of viable HRS cells from primary CHL tumors⁷. We have utilized the technique to produce what is believed to be the first whole-exome sequence of the tumor cells in primary cases of Hodgkin lymphoma. Our studies demonstrate the feasibility of high-throughput, genome-wide studies of individual CHL cases and have already led to the identification of novel genomic alterations with the potential to explain aspects of CHL pathogenesis.

We further developed a pipeline to utilize the extracted DNA for high-throughput genomic studies. In order to achieve reliable results from as few as 1,000 sorted HRS cells (the minimum obtained from sequential cases), we further developed a modified next-generation DNA library construction procedure⁸ that allowed us to increase adaptor ligation efficiency and to generate DNA fragment libraries without excessive

amplification. This method allows for the analysis of routine clinical samples and the detection of recurrent mutations and chromosomal alterations⁷.

Protocol

1. Tissue Processing and Freezing

1. Collect lymph node tissue in phosphate-buffered saline (PBS) or Roswell Park Memorial Institute medium (RPMI) and process within 24 h of collection. Transfer excised lymph node tissue⁹ to a Petri dish containing 10 mL of RPMI with 2% fetal calf serum (FCS) and finely mince it with a fresh scalpel blade. Use the back of a 10-mL syringe plunger to further grind/dissociate the tissue.
2. Transfer the liquid to a 50-mL conical tube through a 100 μ m cell strainer. Rinse the Petri dish and the filter with an additional 10 mL of RPMI 2% FCS.
3. Obtain a cell count using either an automated cell counter or a hemocytometer.
NOTE: Generally, at least 2×10^7 cells would be expected from approximately 5 mm³ of CHL lymph node tissue. A viability of over 80% is expected, but it may vary by sample.
4. Spin down the cells at 400 x g for 10 min and aspirate the supernatant. Place the tube containing cell pellet on ice.
5. Pre-chill freezing medium on ice. Resuspend the cells in cold freezing medium at a concentration of 2×10^7 /mL and resuspend it by pipetting. Do not vortex. Incubate the suspension on ice for 10 min.
6. Aliquot the sample 1 mL/cryo vial (pre-chilled on ice). Transfer the vials to a -80 °C freezer for 1-2 days and transfer them again to liquid nitrogen.

2. Preparing Cell Suspensions for Cell Sorting

NOTE: Each lot of antibody must be properly titered using 10 million cells in a 300- μ L staining volume. Peripheral blood may be used for all antibodies except CD30, for which the KMH2 cell line spiked into peripheral blood may be used for titration¹⁰. We generally begin with the manufacturer-recommended volume of antibody and perform two two-fold dilutions and one two-fold increase (four data points) for each lot of antibody titration. For example, if the manufacturer recommends a 10- μ L volume, we perform the titration using 2, 5, 10, and 20 μ L volumes.

1. Set a water bath to 37 °C. Premix the titered amount of antibodies in a dark glass vial and add PBS + 2% BSA for a total cocktail volume of 100 μ L.
NOTE: Although we recommend titering the antibodies, the following volumes may be used as a starting point: CD64, 20 μ L; CD95, 5 μ L; CD30, 20 μ L; CD5, 10 μ L; CD20, 10 μ L; CD15, 20 μ L; CD40, 5 μ L; and CD45, 10 μ L.
2. Transfer the vial from liquid nitrogen into an ice bucket containing dry ice to prevent thawing. Pre-warm 50 mL of thawing medium containing RPMI/20% FCS/DNase (100 μ g/mL) in a 50-mL conical tube in a 37 °C water bath.
3. Transfer 45 mL of thawing medium to a fresh tube and keep it at 37 °C. Rapidly thaw the cells by holding a cryogenic vial in a 37 °C water bath until only a very small frozen portion remains.
4. Pour the vial contents into the tube containing 45 mL of thawing medium. Rinse the empty cryogenic vial 2 times with 1 mL of thawing medium and combine the rinses.
5. Incubate the cells at room temperature for 15 min to allow for DNase digestion and cell re-equilibration. Spin down the cells at 500 x g for 10 min and aspirate the supernatant.
6. Resuspend the cells in approximately 200 μ L of the thawing medium held back (5 mL) and let it equilibrate to room temperature for 2-3 min. Recovery of >70% of frozen viable cells is expected.
NOTE: Optionally, for greater purity of sorted HRS cells that may be rosetted by attached T cells, a cocktail of unlabeled antibodies may be added at this step (see the optional protocol).
7. Add 100 μ L of an antibody cocktail and incubate for 15 min at room temperature (RT), protected from light. Add 3 mL of sorting medium, spin down the cells at 500 x g for 10 min, and aspirate the supernatant.
8. Resuspend the cells in 1 mL of sorting medium and transfer them to a 5 mL flow tube top strainer.
9. Rinse both the 50 mL conical tube and the cell strainer with an additional 1 mL of sorting medium and place cells on ice.

3. (Optional Protocol) T Cell Rosette Blocking

NOTE: HRS cells are rosetted by T cells in tissue sections and cell suspension, and these T cells may potentially contaminate the sorted HRS fraction. These interactions are mediated by CD54 and CD58 on the HRS cell binding to LFA-1 and CD2 on the T cells^{4,11}. These interactions can be blocked with unlabeled antibodies to these adhesion molecules.

1. Aliquot 100,000 to 500,000 cells in 100 μ L of RPMI.
2. Incubate the cell suspension with unlabeled antibodies to CD2, CD54, CD58, and LFA-1 (10 μ L each) on ice for 1 h. The cell suspension can now be labeled with fluorescent antibodies.

4. HRS-, B-, and T-cell Isolation Using Cell Sorting

NOTE: Although we used a special research order instrument using 5 lasers (see the Materials spreadsheet), any sorter with the capability of detecting the fluorochromes used in the antibody panel should be sufficient. The execution of the steps below requires a familiarity with software¹² function and a basic knowledge of cell sorter operations. Please refer to the online software manual for detailed instructions.

1. **Cytometer setup:**

1. Turn on the computer and login. Power on the BSC (and BSC outlets), and then power on the cytometer. Wait at least 90 s for the internal CPU of the cytometer to start, and then open the laser control software and verify that all lasers are powered on. Launch the cytometer software¹³ and login.
 2. Within the cytometer software, click on "Cytometer → View Configurations." When the dialog box for the configuration sub-program opens, highlight the 130- μ m custom configuration and click "set configuration" and "OK." Exit the configuration sub-program.
 3. Observe the dialog box in the cytometer software and click "use CS&T settings." Install a 130- μ m nozzle into the instrument and turn the stream on by clicking the red "X" button in the stream window. Allow the instrument to warm up for at least 30 min.
 4. Run a performance check on the instrument using the desired method (a performance-tracking software module is included with the sorter described here, see the software manual (pp. 117 - 122) and reference standard (refer to the Materials for the standard used in this setup).
 1. Click "Cytometer → CST," ensure that the "Characterize" field is set to "Check Performance" on the pull-down menu, and click "Run." When prompted by software, load a tube of reference particles onto the sample stage and click "OK."
 2. After the run completes, click "Finish" and close the software module. After the cytometer software has finished reconnecting to the cytometer, click "use CS&T settings" in the dialog box that appears.
 5. Determine the correct drop delay using the automatic drop delay feature of the cytometer software (see pp. 154 - 161 of the software manual).
 1. Open the "Drop Delay" experiment in the "Browser" window of the cytometer software, install a tube of calibration particles on the sample stage, and click "Load" in the "Acquisition Dashboard" window. Turn on the automatic stream monitoring feature by clicking the "Sweet Spot" button in the "Stream" window. Leave this feature on at all times for all following steps.
 2. Turn on the deflection plate voltage by clicking the "Voltage" button in the side-stream window, and then turn on test sorting by clicking the "Test Sort" button immediately adjacent to the "Voltage" button.
 3. Adjust all side-stream settings to zero except for the left side stream. Adjust the left side stream setting such that two stream-spots are visible in the side-stream window. Click the "Optical Filter" button and verify that the left side-stream spot will fall within the left box that appears in the black area of the side-stream window.
 4. Adjust the left side-stream setting if needed. Turn off test sorting by clicking on the "Test Sort" button again. In the "Browser" window, expand the "Global Worksheets" element of the "Drop Delay" experiment by clicking on the "+."
 5. Double-click on "Sort Layout_001" to open the sort layout, verify by visual inspection that the left sort population has "P1" assigned to it, and click "Sort" in the "Sort Layout" window. In the "Sort Layout" window, click "Auto Delay" and then "Run."
 6. When the run completes, click "Exit." Install a tube of sterile deionized water on the sample stage and click "Load" in the "Acquisition Dashboard" window. Run the tube of water for at least 5 min to clear residual test particles from the instrument before proceeding.
 6. Run compensation controls (using compensation beads from the Materials spreadsheet or equivalent) using a built-in compensation setup in the sorter software (see pp. 131 - 137 of the software manual for additional details).
 1. Create a new experiment by clicking "Experiment New Experiment." In the "Parameters" tab of the "Instrument Status" window, delete unused parameters, if any. Click "Experiment Compensation Setup Create Compensation Controls."
 2. In the "Browser" window, expand the "Compensation Controls" specimen by clicking the "+" sign. Run compensation control tubes without recording the data, and if necessary, adjust the detector voltages (in the "Parameters" tab of the "Instrument Status" window) so that positively-stained bead populations for each fluorochrome are between channel 10,000 and 100,000 and are the brightest in their primary detection channel.
 3. Write down the voltages needed for each parameter for each tube. Highlight the "Unstained Control" tube under the "Compensation Controls" specimen by single left-clicking it. Load the unstained control tube onto the sample stage and click "Load" in the "Acquisition Control" window.
 4. Manually enter the determined voltages by running individual compensation controls into the detector voltage fields for all parameters, and then click "Record" in the "Acquisition Control" window. Run all remaining compensation controls, recording the data without changing any detector settings. Install a tube of sterile, deionized water for 5 min to clear any residual material from the instrument before proceeding.
2. **HRS-cell gating:**
1. Acquire and record at least 100,000 events for initial gating while adjusting the flow rate to acquire 3,000-4,000 events/s (see step 4.1). NOTE: It may be necessary to add additional sorting medium to dilute the cells if the cell concentration is too high. Stop the acquisition.
 2. Gate HRS cells using the steps outlined in **Figure 1**
NOTE: In most CHL cases, between 0.01% and 0.1% of cells will be HRS cells.
3. **B- and T-cell gating:**
1. Identify somatic controls (B and T cells) by the gating of CD20 and CD5, respectively (gating lymphocytes by CD45/SSH), followed by CD20 versus CD5 (see **Figure 1**).
 2. Target collection streams to collection tubes in a pre-chilled two-way or four-way collection rack. Fill either flow tubes or 15-mL centrifuge-style tubes at least halfway with collection medium.
 3. Assign HRS and control populations to proper collection tubes in the sort setup following vendor instructions. Restart the cell acquisition and start the sort.
 4. Collect all HRS cells and up to 1 million B and T cells. Target the collection streams to collection tubes in a pre-chilled four-way (or two-way) collection rack.

5. DNA Extraction

1. Pellet collected cells by centrifugation in 1.5 mL conical tubes at 3,000 x g for 10 min and resuspend once with 1 mL of PBS to wash the cells.

2. Pellet once more at 3,000 x g for 10 min and remove the supernatant; be very careful not to disturb the tiny pellet.
3. Add 150 μ L of lysis buffer (or an appropriate volume for the kit used) to the washed cells and mix by pipetting up and down.
NOTE: One may store cell lysate at -70°C at this point, if needed.
4. Construct a column assembly by placing the filter column inside the tube and add the lysate from step 5.5 to the column. Spin the assembly at 13,000 x g for 3 min.
5. Remove the minicolumn from the assembly and discard the liquid in the collection tube. Replace the minicolumn in the collection tube.
6. Add 650 μ L of column wash solution to each assembly. Centrifuge for 1 min at 13,000 x g. Discard the liquid from the collection tube. Repeat this step for a total of 4 washes.
7. Discard the liquid from the collection tube and reassemble the minicolumn assembly. Centrifuge for 2 min at 13,000 x g to dry the binding matrix.
8. Transfer the minicolumn to a new 1.5 mL tube and add 25 μ L of 10 mM Tris-Cl, which is preferred for subsequent sonication step, or Nuclease-Free Water heated to 65°C . Incubate for 2 min at room temperature and centrifuge the assembly at 13,000 x g for 1 min. Repeat once more with 25 μ L for a total of 50 μ L.
9. Mix the DNA elution by pipetting up and down and quantify using fluorimetry¹⁴.

6. Library Construction

1. Preparation before beginning:

1. Set up the sonicator at water level 12, Intensity 5, Cycles/burst 200, and Temp 7.
2. Based on the available DNA quantity determined by fluorimetry and on the experimental design, determine the amount of DNA to use for library construction.

NOTE: Keep in mind that if too little DNA is used and library quality is compromised, there may be little application for material that is left behind for validation purposes. For substandard input quantities, the greater the input mass, the higher the complexity of the resulting sequencing library. Some guidelines for this protocol are as follows: 10 ng should yield good results, and 50 ng may be considered a rough maximum.

3. Decide upon the adapter:insert molar ratio based loosely on the values in **Table 1**.
4. Calculate the amount of adapter to use in the following fashion:

NOTE: Number of moles of DNA input, n

$$i. n = 1.54e-12 * \text{input mass (in ng)} / (\text{mean fragment size})$$

ii. when the mean fragment size is 200 bp, this simplifies to:

$$n = 7.7e-15 * \text{input mass (in ng)}$$

Number of moles of adapter to include, a

$$i. a = n * r$$

Volume of adapter stock to use in adapter ligation step, v

$$i. v \text{ (in } \mu\text{L)} = a / [10^{-12} * \text{adapter stock concentration (in } \mu\text{M)}]$$

NOTE: If the adapter concentration is low, to ensure that the necessary amount of adapter is included, adapter solution may be used to dilute the end-repair reagents in lieu of water.

Example: For 100 ng of input DNA with a mean fragment size of 200 bp, for a desired molar adapter:insert ratio of 15:1 and an adapter stock concentration of 2 μM , the recommended volume of adapter to use is 5.8 μL .

5. Assign indexed barcodes to the samples.

NOTE: Libraries that will be pooled together into a hybridization reaction or a lane on a sequencer's flow cell must not contain any redundant indices.

2. Library construction – DNA shearing:

1. Add the full amount of DNA to be used to a sonication tube. If the volume of sample containing the input DNA alone is less than 50 μL , add Buffer EB up to a 50 μL total volume and mix.
2. Sonicate for 30 s.
3. Remove the tube and perform a quick-spin in a mini centrifuge just enough to collect any spray from the upper portion of the walls of the microtube.
4. Repeat steps 6.2.2-6.2.3 for a total of seven 30-s sonication sessions for a total of 210 s of sonication.

NOTE: Feel free to experiment with dividing the 210 s into fewer sessions.

3. Library construction - End repair, A-tailing, and adapter ligation:

NOTE: Avoid doing any size selection before the library PCR amplification step.

1. Follow the manufacturer's instructions¹⁹ for end-repair and A-tailing after the sonication of the sample DNA.
2. After A-tailing, use the appropriate number of moles of adapter (calculated in step 6.1.3) in the reaction. Combine the adapter, A-tailed DNA fragments, enzyme, and buffer and incubate overnight at 20°C for approximately 16 h.

4. Library construction - Library amplification:

1. Perform a bead cleanup of the adapter ligation reaction. After adding beads to the PCR product, wait 5 min at room temperature. Place it against a magnetic stand and remove the liquids, wash twice in 200 μL of 80% ethanol, dry the beads long enough to remove most of the liquid without over-drying, and elute the DNA off of the beads by pipetting 25 μL of nuclease-free water, as suggested, onto the beads while the tube remains against a magnetic stand.

NOTE: it may be possible to experiment with preserving the beads in polyethylene glycol (PEG) until after the amplification instead of discarding them, but this has not been tested.

2. Include 0.6 μL of a 1:1,000 dilution of Green dye per 50 μL of PCR master mix. Alternatively, use a real-time PCR-compatible intercalating dye of choice in the appropriate volume for the equipment.
3. Program the PCR reaction at 98°C for 45 s for the initial denaturation, followed by a cycle of denaturation, annealing, and extension at 98°C for 15 s, 60°C for 30 s, and 72°C for 30 s, or choose the appropriate program if using an alternative polymerase enzyme.

1. Set the machine to take fluorescence data at 72 °C for each cycle. Program a final extension at 72 °C for 1 min followed by a hold at 4 °C for an indefinite amount of time. The PCR primers for the amplification of the adapter-ligated library using reagents in the supplementary table are: Oligo 1, AATGATACGGCGACCACCGAGA and Oligo 2, CAAGCAGAAGACGGCATACGAG
4. Amplify the library using the PCR conditions described above while observing the fluorescence intensity values in real-time with the qPCR software, stopping just before the end of the exponential growth phase.
5. After amplification, do a standard bead cleanup (see step 6.4.1) using 0.8x the volume of the amplification reaction recovered, typically 0.8 x 50 = 40 µL of beads. Add the beads to the PCR product and wait 5 min at room temperature.
6. Place it against a magnetic stand and remove the liquids, wash twice in 200 µL of 80% ethanol, dry the beads long enough to remove most of the liquid without over-drying, and elute by adding nuclease-free water to the dry beads.
7. Quantify the resulting DNA using fluorimetry. Visualize the library fragments for size; see the section on *Data QC expectations* for more details.

7. Exome Hybridization

1. Combine four libraries with distinct adapter bar codes.
NOTE: Mass by fluorimetry and size by gel are used to calculate the molarity, and then libraries are combined in equimolar amounts for a total of a 1,000-ng mass of pooled library. It is best to keep all tumor-normal pairs together rather than separating them into separate pools.
2. Apply the exome capture protocol¹⁵ and do 8 PCR cycles after the capture cleanup. Other choices of capture targets may be possible.

8. Multiplexed Sequencing

1. Sequence a single hybridization capture containing four multiplexed libraries in a single lane on the sequencing platform referenced in the Materials spreadsheet¹⁶.
NOTE: Alternative configurations are possible for advanced users who wish to further plan and optimize their target read depth of coverage.

9. Analysis (Can be Substituted with Alternative Pipeline(s) if Desired)

1. **Snps and small indels:**
 1. Map raw data to the human reference genome, UCSC hg19, using Burrows-Wheeler Aligner (BWA)¹⁷ or an alternative algorithm of choice. Filter or mark reads with a mapping quality score below 20 and PCR duplicates using Samtools¹⁸ or Picard¹⁹.
 2. Detect somatic nucleotide variants and small indels in HRS samples compared to the T-cell somatic controls using Strelka²⁰ or a variant caller of choice. Apply snpEff²¹ to annotate the Strelka output. If desired, systematically inspect variant loci for artifacts using the Integrated Genome Viewer (IGV)^{22,23}.

2. **Copy number alterations:**

1. Calculate the log-transformed ratio "*ltr*" for every exome target interval "*i*" of intra-library normalized read counts in the tumor against those of the normal in the following manner:

$$ltr(i) = \log_2 \left(\frac{c_t + 1}{l_t} \right) - \log_2 \left(\frac{c_n + 1}{l_n} \right),$$

- NOTE: *c* is the number of reads mapping to a given capture interval, *l* is the total library size, *t* denotes tumor, and *n* denotes normal.
2. Filter out intervals with insufficient coverage ($C_t + C_n < 100$ reads) for further analysis. Conduct pan-interval segmentation using DNACopy v.1.0²⁴ from Bioconductor in R.

NOTE: Consider segments where the absolute value of the mean *ltr* is below 0.5 to be copy-neutral. Remaining segments may be designated as copy number gains, if the sign of the mean *ltr* is positive (in other words, there are significantly more reads in the tumor sample than in the normal sample after normalization), or copy number losses, if the sign of the mean *ltr* is negative.

Representative Results

A bioanalyzer plot should be taken after library amplification and 0.8x bead cleanup. One should see a "normal-like" distribution of fragment sizes in the desired range (**Figure 2a**). Deviations from this shape, such as a visible "shoulder" in the curve, indicate the presence of a high or low molecular weight artifact. For example, **Figure 2b-2d** shows examples of libraries containing visible artifacts that should ideally not be sequenced. If a library is significantly compromised, it may be worthwhile to repeat the library construction if DNA is available and/or the cell sorting to start fresh.

An adapter dimer may sometimes carry over through the first 0.8x bead cleanup. If this occurs, it will be visible as a sharp peak centered around 125 - 130 bp (see **Figure 2c**). In this case, it is advisable to repeat an additional 0.8x bead cleanup and to repeat the bioanalyzer to ensure the successful removal of the dimer.

Using the specifications described in this protocol and multiplexing the sequencing at the level of four samples per lane in the sequencer listed in the Materials spreadsheet, a median depth of coverage of 50 - 100x across the target (after the bioinformatics removal of PCR duplicate reads) is achievable (see **Figure 3a**). In addition, libraries that were not overamplified and that resulted in libraries with adequate coverage should produce clean copy number plots. During early work, this library construction protocol was tested over three orders of magnitude of input mass, and variable results were discovered (see the Discussion section). Copy number plots generated using a library generated from 1 ng, a sub-optimal amount of input DNA, resulted in spurious copy number gains and losses (**Figure 3b**, bottom).

Approximately 70% of cases of CHL carry B2M and A20 mutations and/or deletions⁷. B2M mutations predominantly involve a translational start site, a first exon stop site, and a first exon stop site. A20 mutations are a mix of copy number losses and indels. The mutations appear to be clonal and can be used to estimate relative tumor content. If T cells are used as somatic controls, HRS sequencing will show significant copy number gains in the positions corresponding to T-cell receptor alpha/delta (14q11 - 12) and beta positions (7q32 - 35), as well as significant losses in the B-cell receptor heavy chain (14q32) and kappa light chain positions (2p11) (see **Figure 4**). Overall sequencing shows a range from 100 to 400 somatic mutations per case. Copy number variation is highly variable from case to case, with some cases showing numerous segmental gains and losses and other cases showing only a few alterations.

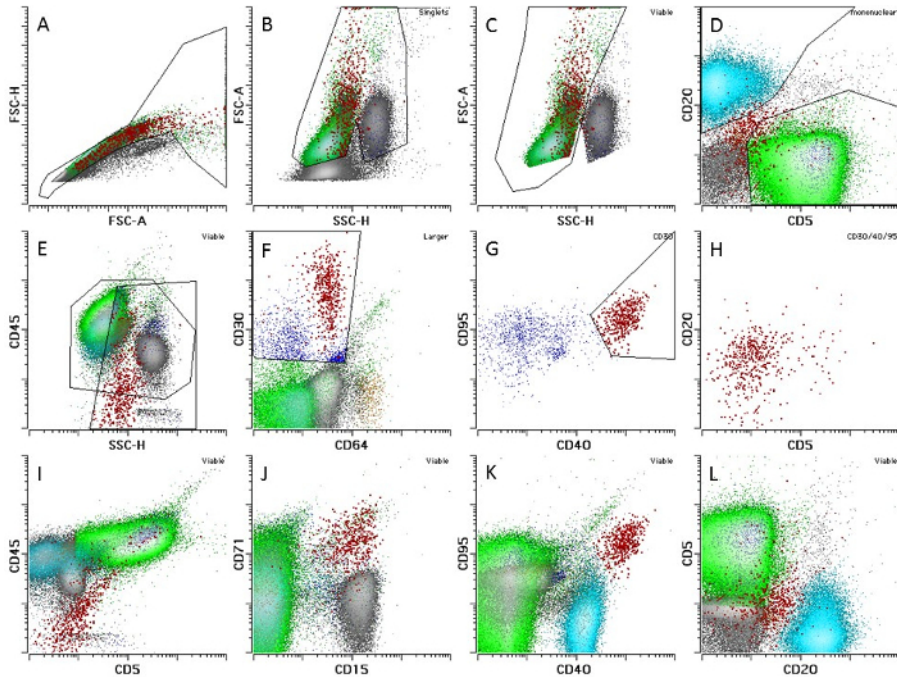


Figure 1: Multiparameter Gating for Flow-sorting of Viable HRS, B, and T cells. Population names are shown on plots and indicate the parent population of any gate. (A) shows all cells: On the FSC-H versus FSC-A histogram, draw a SINGLETs gate that includes populations with proportional FSC-H and FSC-A increases. Exclude doublets that have high FSC-A and lower FSC-H. Hodgkin cells are very large; expand the gate to include high FSC-H and FSC-A events. (B) Exclude debris, dead cells, and unlysed RBCs that normally have lower FSC-A and slight increases in SSC-H without cutting out viable populations (VIABLE gate). (C) Gate MONONUCLEAR cells on FSC-A vs. SSC-H to exclude granulocytes with high SSC-H; note that the gate is only used for the identification of T and B cells. (D) Gate T CELLS positively identified by CD5 (green) without CD20 and B CELLS (blue), which have CD20 expression without CD5. (E) Draw a LARGER gate on the CD45/SSH plot, including 10 - 20% of the lymphoid cells and all granulocytic cells. (F) Gate CD30-positive events with dim to negative CD64/FITC autofluorescence. (G) Among CD30 positive events, gate CD40- and CD95-positive events (CD30/40/95 HRS gate, red). (H-L) Observe the confirmatory immunophenotypic features on HRS cells; HRS cells generally do not express CD20 (plots H and L). The expression of CD5 is correlated with CD45 due to T-cell rosetting (I); most cases will show expression of CD15 and will be bright for CD71 (J). CD40 and CD95 expression is generally above the level of B and T cells, respectively (K). This research was originally published in *Blood*⁷. [Please click here to view a larger version of this figure.](#)

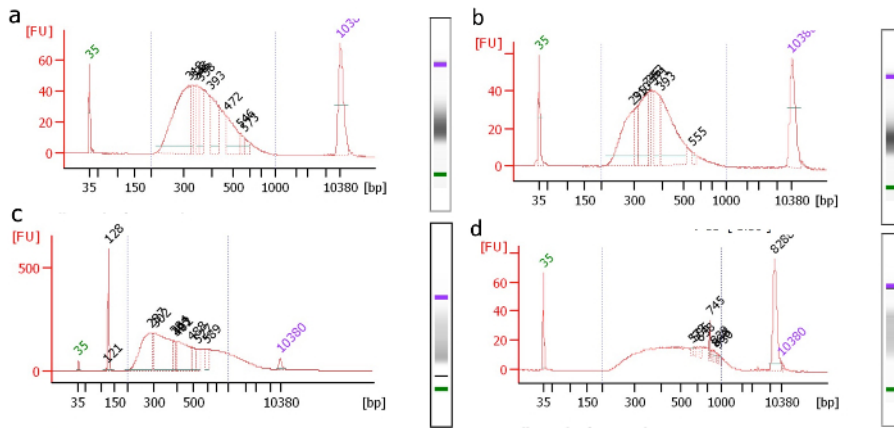


Figure 2: Size Distribution Plots of Samples after Adapter Ligation and Amplification. Ideally, libraries with visible artifacts should be repeated and not sequenced. (a) Well-formed library; no visible artifacts. (b) Asymmetry in the curve reveals a problem. (c) Significant adapter dimer (peak at 128) and a high molecular weight artifact (>400 bp). The dimer may be cleaned with an additional 0.8x bead cleanup. (d) Significant high molecular weight artifact. [Please click here to view a larger version of this figure.](#)

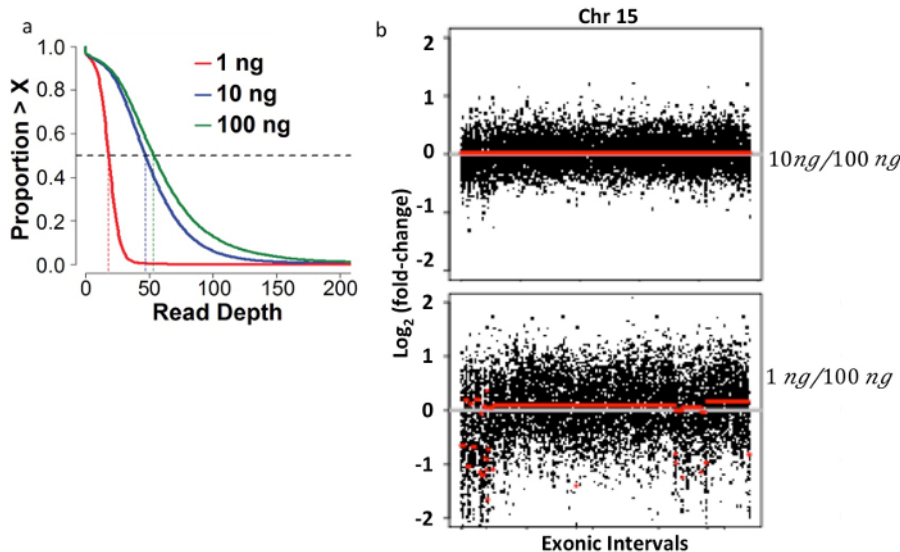


Figure 3: Bioinformatic Quality Control Measurements for Higher-quality Libraries made from 10 ng of Input DNA Compared to Lower-quality Libraries made from 1 ng of Input DNA. The magnitude and distribution of the unique read depth of coverage expectations. (a) A median depth of coverage of at least 50 is expected when using 10-100 ng of input DNA. (b) Each of 2 panels depicts copy number variation analysis results comparing data between 2 sequenced libraries. Exonic probe segments (x-axis) versus copy number change on a \log_2 scale (y-axis) are plotted for a single representative chromosome (chr 6). Comparing data from a 10 ng low-input library from intratumoral T-cell DNA to a 100 ng normal-input library from intratumoral T-cell DNA from the same case showed no significant false-positive results; that is, low-input and normal-input libraries are copy-neutral (top). Numerous false-positive segmental copy number alterations were called when data from a 1 ng low-input library from intratumoral T-cell DNA was compared to a 100 ng normal-input library from intratumoral T-cell DNA from the same case (bottom). This research was originally published in *Blood*⁷. [Please click here to view a larger version of this figure.](#)

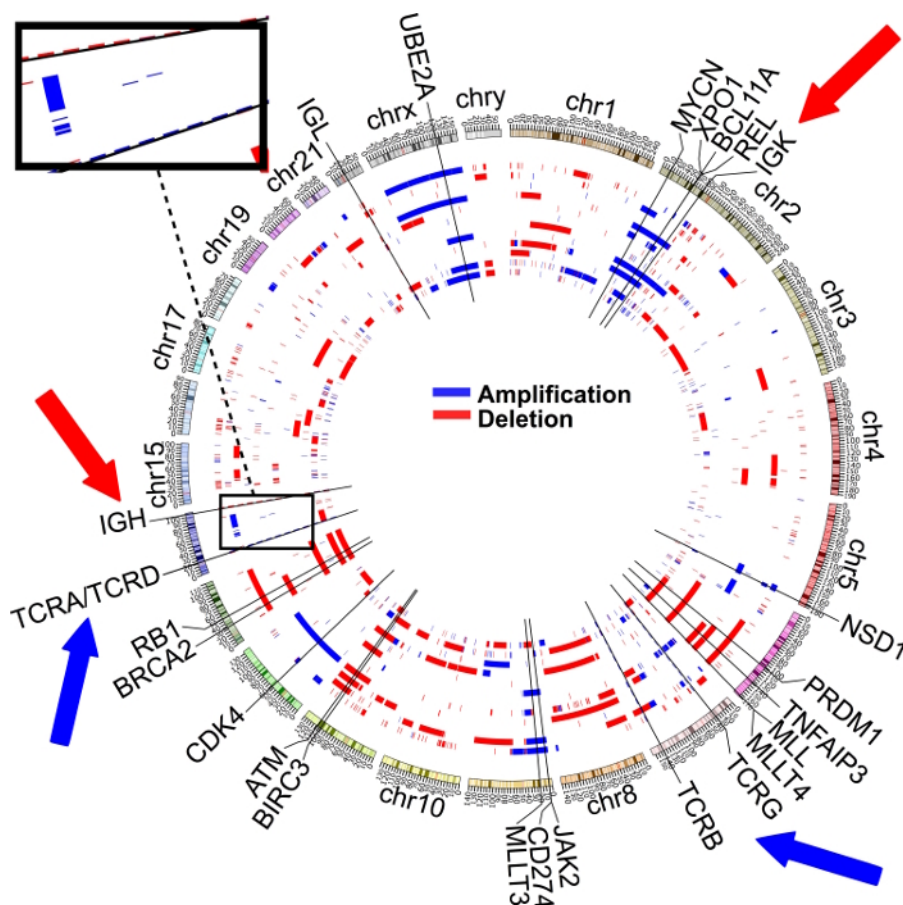


Figure 4: Recurrent Copy Number Gains and Losses, with Tumor Cells (Primary HRS and Cell Lines) Compared Against Healthy Intratumor T Cells from Primary Cases. If T cells are used as the somatic control against B cell-derived HRS and cell line populations, recurrent gains can be seen in T-cell receptors alpha/delta and beta. Likewise, recurrent losses are detectable in immunoglobulin heavy and kappa chains. This research was originally published in *Blood*⁷. [Please click here to view a larger version of this figure.](#)

DNA input mass (n)	1-10 ng	25 ng	100 ng
Adapter : insert molar ratio (r)	65:1	25:1	15:1

Table 1: DNA Input Mass and the Corresponding Adapter:Insert Molar Ratio. The values in this table can be used as a guide to calculate the amount of adapter oligo to add during the adapter ligation step of library construction.

Discussion

Future applications or directions after mastering this technique

This work allows for exome sequencing from samples containing at least 10 ng of DNA. In the clinical context, this limit excludes most fine-needle aspiration samples due to insufficient material, but it includes adequate core biopsies and excisional biopsy samples. This will enable the acquisition of data from a larger set of possible samples.

Critical steps within the protocol

Proper freezing and dissociation techniques are critical to the success of the experiment. The entire sample should be dissociated, including fibrotic sections, as much as possible. Utilizing a 130 μm nozzle for cell sorts appears to be critical to maximize the yield of the large HRS cells and DNA. While rigorous controlled experiments with multiple nozzle sizes were not performed, significant increases in cell yields were observed using the larger nozzle compared to the 100-μm, and virtually no cells were detected on the slide post 85-μm nozzle sort (unpublished observation).

Modifications and troubleshooting

When working with less than 100 ng masses of input DNA for library construction, care should be taken to reduce all unnecessary processing and cleanup steps, especially prior to PCR amplification, since every lost unique molecule results in a reduction of the complexity of the final library. The protocol deliberately recommends eluting extracts in 50 μL because this value is compatible with both sonication and end-repair using the equipment recommended. In this way, it is not necessary to reduce the elution volume using a centrifugal evaporator or a column for the sonication step or to reduce the sonication volume for the end-repair step. Additional modifications to this protocol are possible; for example,

new library construction kits may improve adapter ligation efficiency. One could troubleshoot/optimize by comparing adapter-mediated library amplification curves using real-time amplification on libraries made with different kits from the same quantity and quality of input DNA.

Limitations of the technique

At least 10 ng of purified HRS DNA (approximately 1,000 sorted cells after losses in the sorting and extraction steps) appears to be the minimal amount for high-quality results. In pilot experiments, DNA from a single sample of purified T cells was first used to make libraries spanning a range of input masses. 100 ng (the standard recommended amount) of input DNA was compared to libraries generated using this low-input protocol for 10 ng and 1 ng of input DNA. The library generated from 10 ng of input DNA had no significant differences from the 100-ng input library in magnitude and distribution of coverage across the target, or in PCR duplication fraction. However, the library generated using 1 ng of input DNA resulted in a significantly higher PCR duplicate fraction and a corresponding decrease (approximately 3-fold) in mean coverage across the target. Moreover, the 1 ng library also exhibited deviation from uniformity in coverage, which created false-positive copy number alterations with respect to the 100 ng library, unlike the 10 ng input library, which was copy-neutral in the same assessment. For this reason, exert caution using less than 10 ng of input DNA using this method.

Significance of the technique with respect to existing/alternative methods

Following the protocol will allow full exome sequencing of HRS cells from primary Hodgkin lymphoma samples. CHL lymph node samples that contain at least 2×10^7 cells are suitable for this procedure. This was not previously available to researchers, who relied on techniques such as laser capture microdissection and whole genome amplification. It is predicted that genomic approaches to primary CHL cases will continue to be valuable for further understanding CHL pathogenesis. Data also points to significant genome-level heterogeneity within this disease entity, with potentially at least two defined subsets that loosely follow morphological stratification between nodular sclerosis and mixed cellularity CHL. Finally, it is believed that genome-level data will lead to the development of targeted therapy for difficult-to-treat relapsed/refractory CHL cases.

Disclosures

The authors have nothing to disclose.

Acknowledgements

The development of this project method was funded by the Department of Pathology and Laboratory Medicine of Weill Cornell Medical College. We acknowledge the Tri-Institutional Training Program in Computational Biology and Medicine for partial funding. We would like to thank the scientists who shared their time and knowledge with us, especially Maryke Appel; Dan Burgess; Iwanka Kozarewa; Chad Locklear; and everyone from the Weill Cornell Medical College Genomics Core Facility, including Jenny Zhang, Xiaobo (Shawn) Liang, Dong Xu, Wei Zhang, Huimin Shang, Tatiana Batson, and Tuo Zhang.

References

- Abrams, J. *et al.* National Cancer Institute's Precision Medicine Initiatives for the new National Clinical Trials Network. *American Society of Clinical Oncology educational book / ASCO. American Society of Clinical Oncology. Meeting.* 71-76 (2014).
- Gagan, J., & Van Allen, E. M. Next-generation sequencing to guide cancer therapy. *Genome Medicine.* 7, 80 (2015).
- Matsuki, E., & Younes, A. Lymphomagenesis in Hodgkin lymphoma. *Seminars in Cancer Biology.* 34, 14-21 (2015).
- Fromm, J. R., Kussick, S. J., & Wood, B. L. Identification and purification of classical Hodgkin cells from lymph nodes by flow cytometry and flow cytometric cell sorting. *Am J Clin Pathol.* 126, 764-780 (2006).
- Fromm, J. R., Thomas, A., & Wood, B. L. Flow cytometry can diagnose classical Hodgkin lymphoma in lymph nodes with high sensitivity and specificity. *Am J Clin Pathol.* 131, 322-332 (2009).
- Roshal, M., Wood, B. L., & Fromm, J. R. Flow cytometric detection of the classical Hodgkin lymphoma: clinical and research applications. *Advances in Hematology.* 2011, 387034 (2011).
- Reichel, J. *et al.* Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. *Blood.* 125, 1061-1072 (2015).
- Kozarewa, I. *et al.* A Modified Method for Whole Exome Resequencing from Minimal Amounts of Starting DNA. *PLoS ONE.* 7, e32617 (2012).
- Brunnicardi, F. C. *Schwartz's Principles of Surgery.* 10th edn, McGraw-Hill Education / Medical, (2014).
- Kantor AB, R. M. *FACS analysis of leukocytes.* Vol. Chapter 49 Blackwell Scientific, (1996).
- Sanders, M. E. *et al.* Molecular pathways of adhesion in spontaneous rosetting of T-lymphocytes to the Hodgkin's cell line L428. *Cancer Res.* 48, 37-40 (1988).
- Biosciences, *FACSDiva Software v6.0.* (2007).
- Biosciences, *FACSARIA II User's Guide. Part No. 644832, Revision A, March 2009.* (2009).
- Life Technologies, *Qubit 3.0 Fluorometer.* Catalog Number Q33216. (2016).
- Roche Nimblegen. *SeqCap EZ Library SR User's Guide ver. 4.1.* (2013).
- Illumina, *HiSeq High-Throughput Sequencing System.* <http://www.illumina.com/systems/hiseq_2500_1500.html> (2016).
- Li, H., & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 25, 1754-1760 [pii] (2009).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25, 2078-2079 (2009).
- The Broad Institute, *Picard Tools.*, <<http://broadinstitute.github.io/picard/>> (2016).
- Saunders, C. T. *et al.* Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics.* 28, 1811-1817 (2012).
- Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 6, 80-92 (2012).
- Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 29, 15-21 [pii] (2013).

23. Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*. **14**, 178-192 (2013).
24. DNA copy number data analysis. *v.R package version 1.34.0*. (2017).