Burkitt Lymphoma Genome Sequencing Project: Introduction

Bruno M. Grande, Daniela S. Gerhard, BLGSP Consortium, Marco A. Marra, Ryan D. Morin, Louis M. Staudt
What is Burkitt lymphoma (BL)?

Aggressive B-cell non-Hodgkin lymphoma

Most common in children located in malaria-endemic regions

Three clinical variants:

1) Endemic BL
2) Sporadic BL
3) Immunodeficiency-related BL

Image source Ribeiro and Sandlund, 2008
Challenges in treating endemic BL

Late stage at presentation

Limitations in the ability to support intensive chemotherapeutic regimens

More relevant to sporadic BL:

• Less effective in adult and elderly patients

• Treatment-resistant disease

Image source  Ribeiro and Sandlund, 2008
Mutational landscape in sporadic BL

Translocation in MYC

Image source Schmitz et al., 2015
Recent genomic studies on endemic BL

In 2015: 20 samples with RNA-seq (Abate et al.)
In 2017: 28 samples with RNA-seq (Kaymaz et al.)

Limitations of these studies:

• Difficulty detecting true somatic variants
• Inability to compare gene expression with other RNA-seq dataset (due to batch effects)
• Limited sample sizes
Introducing the Burkitt Lymphoma Genome Sequencing Project (BLGSP)

An integrative molecular characterization of a large comprehensive BL cohort including an unprecedented representation of endemic cases

We aim to sequence 160 BL tumor-normal pairs

- 50% will be endemic (mostly paediatric)
- 38% will be sporadic (paediatric and adult)
- 12% will be from HIV+ patients
Multi-dimensional data

Whole genome sequencing (WGS)
  - 80X for tumours and 40X for normals

Ribo-depleted RNA sequencing (RNA-seq)
  - On average, 200 million reads per sample

miRNA sequencing (miRNA-seq)

Patient outcome and other clinical metadata
BLGSP Consortium: Over a dozen institutions

- Uganda Cancer Institute
  Kampala, Uganda
- EMBLEM
  Gulu, Uganda
- Foundation for Burkitt Lymphoma Research
  Geneva, Switzerland
- Lyon University Hospital
  Lyon, France
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Memorial Sloan Kettering Cancer Center
  New York, NY, USA
- National Cancer Institute
  Washington, DC, USA
- Children’s Oncology Group
  Columbus, OH, USA
- Massachusetts General Hospital
  Boston, MA, USA
- BC Cancer Agency
  Vancouver, Canada
- Simon Fraser University
  Burnaby, Canada
- Fred Hutchinson Cancer Research Center
  Seattle, WA, USA
- University of Nebraska Medical Center
  Omaha, NE, USA
- St. Jude Research Hospital
  Memphis, TN, USA
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Foundation for Burkitt Lymphoma Research
  Geneva, Switzerland
- Lyon University Hospital
  Lyon, France
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Memorial Sloan Kettering Cancer Center
  New York, NY, USA
- National Cancer Institute
  Washington, DC, USA
- Children’s Oncology Group
  Columbus, OH, USA
- Massachusetts General Hospital
  Boston, MA, USA
- BC Cancer Agency
  Vancouver, Canada
- Simon Fraser University
  Burnaby, Canada
- Fred Hutchinson Cancer Research Center
  Seattle, WA, USA
- University of Nebraska Medical Center
  Omaha, NE, USA
- St. Jude Research Hospital
  Memphis, TN, USA
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Foundation for Burkitt Lymphoma Research
  Geneva, Switzerland
- Lyon University Hospital
  Lyon, France
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Memorial Sloan Kettering Cancer Center
  New York, NY, USA
- National Cancer Institute
  Washington, DC, USA
- Children’s Oncology Group
  Columbus, OH, USA
- Massachusetts General Hospital
  Boston, MA, USA
- BC Cancer Agency
  Vancouver, Canada
- Simon Fraser University
  Burnaby, Canada
- Fred Hutchinson Cancer Research Center
  Seattle, WA, USA
- University of Nebraska Medical Center
  Omaha, NE, USA
- St. Jude Research Hospital
  Memphis, TN, USA
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Foundation for Burkitt Lymphoma Research
  Geneva, Switzerland
- Lyon University Hospital
  Lyon, France
- Nationwide Children’s Hospital
  Columbus, OH, USA
- Memorial Sloan Kettering Cancer Center
  New York, NY, USA
- National Cancer Institute
  Washington, DC, USA
- Children’s Oncology Group
  Columbus, OH, USA
- Massachusetts General Hospital
  Boston, MA, USA
- BC Cancer Agency
  Vancouver, Canada
- Simon Fraser University
  Burnaby, Canada
- Fred Hutchinson Cancer Research Center
  Seattle, WA, USA
- University of Nebraska Medical Center
  Omaha, NE, USA
- St. Jude Research Hospital
  Memphis, TN, USA
- Nationwide Children’s Hospital
  Columbus, OH, USA
### BLGSP discovery cohort (so far)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Discovery (n = 109)</th>
<th>ICGC MALY * (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pediatric</td>
<td>101</td>
<td>17</td>
</tr>
<tr>
<td>Adult</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Not submitted yet</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td><strong>Clinical Variant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endemic</td>
<td>77</td>
<td>0</td>
</tr>
<tr>
<td>Sporadic</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>HIV-associated</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Unknown</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Not submitted yet</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

* Some analyses are supplemented by 17 ICGC sporadic BLs
Importance of sequencing germline DNA, especially for African cases

Identifying somatic variants in RNA-seq data requires the removal of germline variation and RNA editing events.

Removing germline variation is especially difficult with African cases:

• Current knowledge of germline variation (dbSNP) is biased towards non-African populations.

• The African population harbours the highest genetic diversity in the world.
Case in point: Higher false positive rate for nonsynonymous mutations in endemic cases

Which dbSNP database is used to filter RNA-seq variants?

![Graph showing comparison between Common SNPs and All SNPs for endemic and sporadic cases.](image)

- Common SNPs: 1.5x increase in false positives
- All SNPs: 1.9x increase in false positives

**Clinical variant**

- Endemic
- Sporadic
Significantly mutated genes (SMGs) in BL

<table>
<thead>
<tr>
<th>Cohort-wide</th>
<th>Endemic BL Cases</th>
<th>Sporadic BL Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDX3X</td>
<td>52%</td>
<td>17%</td>
</tr>
<tr>
<td>ARID1A</td>
<td>45%</td>
<td>17%</td>
</tr>
<tr>
<td>FOXO1</td>
<td>35%</td>
<td>13%</td>
</tr>
<tr>
<td>FBXO11</td>
<td>20%</td>
<td>33%</td>
</tr>
<tr>
<td>TP53</td>
<td>20%</td>
<td>50%</td>
</tr>
<tr>
<td>CCND3</td>
<td>14%</td>
<td>47%</td>
</tr>
<tr>
<td>SMARCA4</td>
<td>11%</td>
<td>43%</td>
</tr>
<tr>
<td>ID3</td>
<td>32%</td>
<td>50%</td>
</tr>
<tr>
<td>GNA13</td>
<td>20%</td>
<td>10%</td>
</tr>
<tr>
<td>TFAP4</td>
<td>14%</td>
<td>13%</td>
</tr>
<tr>
<td>PCBP1</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>SIN3A</td>
<td>12%</td>
<td>3%</td>
</tr>
<tr>
<td>GNAI2</td>
<td>11%</td>
<td>7%</td>
</tr>
<tr>
<td>HIST1H1E</td>
<td>10%</td>
<td>7%</td>
</tr>
<tr>
<td>RHOA</td>
<td>11%</td>
<td>7%</td>
</tr>
<tr>
<td>KMT2D</td>
<td>7%</td>
<td>17%</td>
</tr>
<tr>
<td>P2RY8</td>
<td>6%</td>
<td>17%</td>
</tr>
<tr>
<td>BCL7A</td>
<td>6%</td>
<td>0%</td>
</tr>
<tr>
<td>CHD8</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td>TCF3</td>
<td>4%</td>
<td>13%</td>
</tr>
<tr>
<td>USP7</td>
<td>4%</td>
<td>13%</td>
</tr>
<tr>
<td>RFX7</td>
<td>4%</td>
<td>7%</td>
</tr>
</tbody>
</table>

Legend:
- % Mutation rates
- Truncating mutation
- Inframe mutation
- Missense mutation
- Genes not previously linked to BL
Endemic or EBV-positive cases have a higher mutation burden

![Mutation load comparison graph]

- **Endemic Clinical variant**
- **Sporadic variant**
- **Positive EBV infection status**
- **Negative EBV infection status**
Endemic or EBV-positive cases have a higher nonsynonymous mutation burden
EBV-positive tumours harbour fewer nonsynonymous mutation in BL genes
A single high-confidence EBV integration event has been identified so far. It affects the GAS6 gene, but no striking expression pattern
Differential mutation rates for several genes

<table>
<thead>
<tr>
<th>Cohort-wide</th>
<th>Endemic BL Cases</th>
<th>Sporadic BL Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>42% DDX3X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37% ARID1A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29% FOXO1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19% FBXO11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29% TP53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24% CCND3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21% SMARCA4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38% ID3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17% GNA13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14% TFAP4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13% PCBP1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12% SIN3A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10% GNAI2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10% HIST1H1E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10% RHOA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9% KMT2D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9% P2RY8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8% BCL7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8% CHD8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8% TCF3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7% USP7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5% RFX7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend: % Mutation rates ■ Truncating mutation ■ Inframe mutation ■ Missense mutation ■ Significant difference between endemic and sporadic BL
Most BL tumours resemble cells in the germinal center dark zone.
There’s a need to reconcile seemingly contradictory results

BL tumours are more similar to **dark zone** cells

BL tumours are more similar to **light zone** cells

**Source**: Victora et al., 2012
Summary

We identified high-confidence significantly mutated genes

- Including novel genes not previously linked to BL
- Some genes show differential mutation rates

EBV-positive tumours show attenuated selection for driver mutations in BL genes

EBV integration events are rare and likely passenger events

BL tumors resemble cells in the germinal center dark zone
Acknowledgements

British Columbia Cancer Agency
Vancouver, Canada
Andy Mungall
Karen Novik
Marco A. Marra
Yussanne Ma

Foundation for Burkitt Lymphoma Research
Geneva, Switzerland
Jean Paul Martin
John D. Irvin
Marie-Reine Martin

George Washington University
Washington, DC
Fabio Leal
Jeffrey Bethony

Infectious Disease Research Institute
Seattle, WA
Corey Casper

Leidos Biomedical Research
Frederick, MD
Maureen Dyer

Massachusetts General Hospital
Boston, MA
Jeremy S. Abramson
Nancy Lee Harris

Memorial Sloan Kettering Cancer Center
New York, NY
Ariela Noy

National Cancer Institute
Bethesda, MD
Daniela S. Gerhard
Elaine S. Jaffe
Louis M. Staudt
Nicholas B. Griner
Patee Gesuwan
Roland Schmitz
Sam M. Mbulaiyte
Tanja M. Davidsen
Thomas Gross
Wyndham Wilson
Yiwen He

Nationwide Children’s Hospital
Columbus, OH
Hilary Allen
Jay Bowen
Julie M. Gastier-Foster

Simon Fraser University
Burnaby, Canada
Bruno M. Grande
Ryan D. Morin

St. Jude Children’s Hospital
Memphis, TN
Charles G. Mullighan
John Kim Choi
John T. Sandlund
Thomas Alexander

Uganda Cancer Institute
Kampala, Uganda
Abraham Omoding
Constance Namirembe
Jackson Orem

University of Nebraska Medical Center
Omaha, NE
Timothy C. Greiner

And the patients and their families
Merci pour votre attention!

Thank you for your attention!