High-grade serous ovarian adenocarcinoma transcriptome sequencing

Andrew J. Mungall, Ph.D.
British Columbia Cancer Agency Genome Sciences Centre
amungall@bcgsc.ca
High-grade serous ovarian cancer

- Most deaths from advanced-stage, high-grade serous ovarian carcinoma
  - 489 tumours: mRNA, miRNA, DNA copy number and methylation.
  - 316 cases: exome T/N sequencing.
  - Simple mutational spectrum, *TP53* in 96% tumours
  - High frequency of somatic copy number aberrations

- **Aim of this study**
  - Transcriptome (mRNA & miRNA) sequencing: subtypes, structural variants and alternatively spliced transcripts.
Transcriptome datasets

- 490 tumour samples (15 TSSs)
- 420 RNA-seq libraries sequenced
  - 420 submitted to CGHub and DCC
  - 300 expression datasets passed QC* submitted to DCC
- 485 miRNA-seq
  - All passed QC# & data submitted to CGHub and DCC
- Analyses:
  - Unsupervised NMF consensus clustering
  - miRNA anti-correlations with mRNA isoform expression
  - Trans-ABySS & UC-fusion-finder identification of gene fusions

*: >5Gb total; >21,000 genes; <20% rRNA; <20% mitochondrial; >0.6 5’/3’ ratio etc.
#: >750,000M miRNA aligned reads
Sequence-based mRNA expression profiling suggests two additional groups.

Microarray-based

TCGA

489 tumours

1,000 genes

Gene expression

Low

High

Tumour/gene groups

Differentiated

Immunoreactive

Mesenchymal

Proliferative

Sequence-based

300 tumours

TCGA Research Network (2011) Figure 2a

The Cancer Genome Atlas
miRNA expression profiling identifies at least 3 clusters

TCGA Research Network (2011) Figure S6.5
Interplay between miR-29a and DNMT3A transcript isoforms

Only DNMT3A mRNA isoforms harboring the miR-29a binding site have negatively correlated expression profiles with miR-29a.
Gene fusions

- No tumour total RNA available for verification, so orthogonal analysis methods were used
- 1,538 calls overlap between the two methods
- 64 recurrent (≥2) gene fusions

Trans-ABySS
n=420

UC-fusion-finder
n=394

2,758 1,538 [64] 649
Recurrent gene fusions

In-frame

Out-of-frame

Fusion event in Mitelman
One, or both, genes in Mitelman
Genes not reported in Mitelman
In-frame *MECOM* fusion events

- *MDS1* and *EVI1* complex locus (*MECOM*) was focally amplified in >20% OV tumours (TCGA Res Network 2011)
- *MECOM* is a target of the therapeutic compounds aurintricarboxilic acid, arsenic trioxide
- We identify *MECOM* in-frame fusions with several different gene partners in at least 14 (3%) OV cases

<table>
<thead>
<tr>
<th>Fusion</th>
<th>Recurrence</th>
<th>Event type</th>
<th>Event breakpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MECOM-LRRC31</em></td>
<td>6</td>
<td>duplication</td>
<td>exon1,5’UTR (5) exon1,exon3 (1)</td>
</tr>
<tr>
<td><em>MECOM-LRRC34</em></td>
<td>4</td>
<td>duplication</td>
<td>exon1,exon2</td>
</tr>
<tr>
<td><em>MECOM-CLDN1</em></td>
<td>2</td>
<td>duplication</td>
<td>exon1,exon2</td>
</tr>
<tr>
<td><em>MECOM-LMAN2L</em></td>
<td>1</td>
<td>translocation</td>
<td>exon3,exon1</td>
</tr>
<tr>
<td><em>MECOM-SLC7A14</em></td>
<td>1</td>
<td>duplication</td>
<td>exon1,exon3</td>
</tr>
</tbody>
</table>
MECOM-LRRC31 recurrent, in-frame fusion

MDS1 and EVI1 complex locus (MECOM)

Leucine-rich repeat-containing 31 (LRRC31)

MECOM-LRRC31 recurrent, in-frame fusion

MAPK9, SMAD3 and SUV39H1 interaction domain
Known cancer-related pathways are significantly enriched with fusion genes

- 2,415 unique genes in 1,538 fusions
- 105 genes in COSMIC: causally implicated with cancer (p=1.8e-12)

<table>
<thead>
<tr>
<th>Pathway database</th>
<th>Pathway</th>
<th>Adj. p-val (2415 genes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEGG; IPA</td>
<td>Pathways in cancer</td>
<td>3.0E-04; 5.1E-03</td>
</tr>
<tr>
<td>KEGG; IPA</td>
<td>Tight junction signaling</td>
<td>4.3E-02; 2.0E-02</td>
</tr>
<tr>
<td>KEGG; IPA</td>
<td>Cell cycle</td>
<td>4.4E-02; 3.3E-02</td>
</tr>
<tr>
<td>KEGG</td>
<td>WNT signaling</td>
<td>1.5E-02</td>
</tr>
<tr>
<td>KEGG</td>
<td>Ubiquitin-mediated proteolysis</td>
<td>2.6E-04</td>
</tr>
<tr>
<td>KEGG</td>
<td>ERBB signaling</td>
<td>2.4E-03</td>
</tr>
<tr>
<td>IPA</td>
<td>PI3K/AKT signaling</td>
<td>1.3E-02</td>
</tr>
<tr>
<td>IPA</td>
<td>TGF-b signaling</td>
<td>1.4E-02</td>
</tr>
<tr>
<td>IPA</td>
<td>Role of BRCA1 in DNA damage response</td>
<td>1.4E-02</td>
</tr>
</tbody>
</table>
Summary

- Generated mRNA-seq and miRNA-seq for 420 and 485 of the TCGA high-grade serous ovarian adenocarcinoma cohort
- Unsupervised clustering of mRNA/miRNA expression profiles identifies additional sample groups
- An exploration of putative miRNA and mRNA interactions identifies significant expression anti-correlations including miR-29a with specific isoforms of \textit{DNMT3A}
- In contrast to other cancers, such as AML, duplication is the primary rearrangement leading to gene fusions
- \textit{MECOM} fusions are the most recurrent in-frame events
Future work

- Recurrent PTDs and ITDs (Lucas Swanson – poster #106 ‘Barnacle’)
- Rearrangements e.g. *MECOM*
- Differential expression and discriminatory gene analysis for unsupervised clusters and for gene rearrangements
- Further integrated analyses with our TCGA collaborators
Acknowledgements

Reanne Bowlby
Andy Chu
Hye-Jung Chun
Gordon Robertson
Karen Mungall
Readman Chiu
Kirstine Hamilton
Emilia Lim
Richard Corbett
Payal Sipahimalani
Dominik Stoll

Ka Ming Nip
Jenny Qian
Nina Thiessen
An He
Martin Krzywinski
Jacquie Schein
Richard Varhol
Angela Tam
Yongjun Zhao
Richard Moore
Inanc Birol

Poster #146

BCGSC Production Teams
Marco Marra
Steve Jones
Kevin White
Chai Bandlamudi
TCGA Research Network

National Cancer Institute Award Number U24CA143866
Comparison of RNA and miRNA cluster membership

292 shared IDs

RNA-seq

300 IDs

miRNA-seq

462 IDs

The Cancer Genome Atlas
Compare miRNA-seq 3-groups to microarray 3-groups
Compare miRNA-seq 6-groups to microarray 3-groups
Compare miRNA-seq 6 to 3-groups

miRNA-seq 6 groups

miRNA-seq 3 groups

462 samples
**Putative miRNA:mRNA Isoform-specific Interactions**

miRNA:Gene pairs where only mRNA transcript isoforms with miRNA binding sites have negatively correlated miRNA and mRNA expression profiles.

These interactions suggest an interplay between alternative isoform expression (AIE) & miRNA-mediated repression (MMR).

### Top 19 miRNA:Gene Pairs That Display AIE-MMR Interplay

<table>
<thead>
<tr>
<th>miRNA Name</th>
<th>MIMAT ID</th>
<th>Gene Name</th>
<th>With MBS (Max Q-val)</th>
<th>Without MBS (Max Q-val)</th>
<th>With MBS Mean Rho</th>
<th>Without MBS Mean Rho</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-mir-129</td>
<td>MIMAT0000242</td>
<td>ATF5</td>
<td>3.22E-05</td>
<td>0.077171081</td>
<td>-0.310180981</td>
<td>0.143693947</td>
</tr>
<tr>
<td>hsa-mir-509-3</td>
<td>MIMAT0004975</td>
<td>CHN1</td>
<td>1.31E-13</td>
<td>0.934810527</td>
<td>-0.452791643</td>
<td>0.01070408</td>
</tr>
<tr>
<td>hsa-let-7b</td>
<td>MIMAT000063</td>
<td>EGFLAM</td>
<td>1.14E-05</td>
<td>0.999946714</td>
<td>-0.305326604</td>
<td>0.040647623</td>
</tr>
<tr>
<td>hsa-let-7i</td>
<td>MIMAT0000415</td>
<td>GABPB1</td>
<td>1.89E-05</td>
<td>0.865677253</td>
<td>-0.30646911</td>
<td>0.011480833</td>
</tr>
<tr>
<td>hsa-mir-186</td>
<td>MIMAT0000456</td>
<td>GABPB1</td>
<td>9.75E-06</td>
<td>0.705176441</td>
<td>-0.335491053</td>
<td>0.089241001</td>
</tr>
<tr>
<td>hsa-mir-877</td>
<td>MIMAT0004950</td>
<td>PPP2R2A</td>
<td>2.71E-05</td>
<td>0.000496779</td>
<td>-0.31261063</td>
<td>0.270502113</td>
</tr>
<tr>
<td>hsa-mir-509</td>
<td>MIMAT0002881</td>
<td>RCAN2</td>
<td>1.42E-06</td>
<td>0.963815897</td>
<td>-0.31998578</td>
<td>0.000418671</td>
</tr>
<tr>
<td>hsa-mir-331</td>
<td>MIMAT0000760</td>
<td>SEC14L1</td>
<td>1.30E-06</td>
<td>0.957171386</td>
<td>-0.33796075</td>
<td>0.13258906</td>
</tr>
<tr>
<td>hsa-mir-1180</td>
<td>MIMAT0005825</td>
<td>SET</td>
<td>7.09E-06</td>
<td>0.950153414</td>
<td>-0.301748309</td>
<td>0.016205006</td>
</tr>
<tr>
<td>hsa-mir-29b</td>
<td>MIMAT000100</td>
<td>TIPRL</td>
<td>4.73E-06</td>
<td>0.486361605</td>
<td>-0.31272921</td>
<td>0.062702709</td>
</tr>
<tr>
<td>hsa-mir-9</td>
<td>MIMAT0000441</td>
<td>CHN1</td>
<td>9.70E-07</td>
<td>0.117736333</td>
<td>-0.348815743</td>
<td>0.193468941</td>
</tr>
<tr>
<td>hsa-mir-514</td>
<td>MIMAT0002883</td>
<td>CHN1</td>
<td>1.39E-12</td>
<td>0.981482288</td>
<td>-0.437334338</td>
<td>0.019167667</td>
</tr>
<tr>
<td>hsa-mir-509-3</td>
<td>MIMAT0004975</td>
<td>CHN1</td>
<td>1.31E-13</td>
<td>0.934810527</td>
<td>-0.452791643</td>
<td>0.01070408</td>
</tr>
<tr>
<td>hsa-let-7i</td>
<td>MIMAT0000415</td>
<td>GABPB1</td>
<td>1.89E-05</td>
<td>0.865677253</td>
<td>-0.30646911</td>
<td>0.011480833</td>
</tr>
<tr>
<td>hsa-mir-186</td>
<td>MIMAT0000456</td>
<td>GABPB1</td>
<td>9.75E-06</td>
<td>0.705176441</td>
<td>-0.335491053</td>
<td>0.089241001</td>
</tr>
<tr>
<td>hsa-mir-1228</td>
<td>MIMAT0005583</td>
<td>MCM7</td>
<td>4.74E-05</td>
<td>0.915797117</td>
<td>-0.30606887</td>
<td>0.014436598</td>
</tr>
<tr>
<td>hsa-mir-509</td>
<td>MIMAT0002881</td>
<td>RCAN2</td>
<td>1.42E-06</td>
<td>0.963815897</td>
<td>-0.31998578</td>
<td>0.000418671</td>
</tr>
<tr>
<td>hsa-mir-331</td>
<td>MIMAT0000760</td>
<td>SEC14L1</td>
<td>1.30E-06</td>
<td>0.957171386</td>
<td>-0.33796075</td>
<td>0.13258906</td>
</tr>
<tr>
<td>hsa-mir-29b</td>
<td>MIMAT000100</td>
<td>TIPRL</td>
<td>4.73E-06</td>
<td>0.486361605</td>
<td>-0.31272921</td>
<td>0.062702709</td>
</tr>
</tbody>
</table>
Only ATF5 mRNA isoforms harboring the miR-129 binding site have negatively correlated expression profiles with miR-129.
Expressed mutations and RNA-editing


- Focus on gene mutations listed in marker paper: *TP53* (96%), *NF1*, *BRCA1* & *BRCA2*, *RB1*, *CDK12*
- Look for ITDs, PTDs and SNVs
- Can we find evidence for TP53 mutations in the 4% of patients (~16?) missing from the marker paper?
Orthogonal gene fusion detection

- We observe 60 recurrent (≥2 libraries), high-confidence events when overlapping Trans-ABYSS and FusionFinder results.
- These include:
  - **NCOR2-UBC** (11);
  - **XPR1-ACBD6** (9);
  - **GTF2I-GTF2IRD1** (8);
  - **CCDC6-ANK3** (7);
  - **GATAD2B-CRTC2** (6);
  - **TFG-GPR128** (6);
  - **COL14A1-DEPDC6** (5);
  - **MECOM-LRRC31** (5)

Trans-ABYSS and Fusion Finder results:

- 3007 events in total
- Sense and anti-sense fusions reported.
- Includes 2758 gene fusions
- Includes 249 Large-scale rearrangements (one or no annotated genes)

- All but 15 are in sense orientation.
- 2178 events in Chai’s excel file
- Of the 1127:
  - 487 are short contigs (e.g. <150bp) or have insufficient read evidence for Trans-ABYSS calls.
  - 360 have poor contig-to-genome alignments
  - 280 putative novel FF events
In-frame **MECOM-LRRC31** fusion found in 5/420 libraries

A14132_k44_1851400 180-3057bp 100.0% chr3:169556616-169587660 (-)
A14132_k44_1851400 1-179bp 100.0% chr3:169381124-169381302(-)
## Summary of all MECOM events

<table>
<thead>
<tr>
<th>Fusion</th>
<th>Recurrence</th>
<th>In frame</th>
<th>Event type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MECOM_LRRC31</td>
<td>6</td>
<td>yes</td>
<td>duplication</td>
<td>exon1,5utr (5) (A08215,A12152,A14126,A14132,A14134) exon1,exon3 (1) A14327</td>
</tr>
<tr>
<td>MECOM_LRRC34</td>
<td>4</td>
<td>yes</td>
<td>duplication</td>
<td>exon1,exon2 (3) (A14376,A14193,A14280) exon1,exon3 (1) (A14180)</td>
</tr>
<tr>
<td>MECOM_CLDN1</td>
<td>2</td>
<td>yes</td>
<td>duplication</td>
<td>exon1,exon3 (A14279)</td>
</tr>
<tr>
<td>MECOM_LMAN2L</td>
<td>1</td>
<td>yes</td>
<td>translocation</td>
<td>exon3,exon1 (A08240)</td>
</tr>
<tr>
<td>MECOM_SLC7A14</td>
<td>1</td>
<td>yes</td>
<td>duplication</td>
<td>exon1,exon3 (A14279)</td>
</tr>
<tr>
<td>MECOM_NA</td>
<td>1</td>
<td></td>
<td>deletion</td>
<td>exon1,NA (A12038)</td>
</tr>
<tr>
<td>MECOM_NA</td>
<td>1</td>
<td></td>
<td>duplication</td>
<td>exon1,NA (A12132)</td>
</tr>
<tr>
<td>MECOM_NA</td>
<td>1</td>
<td></td>
<td>duplication</td>
<td>exon1,NA (A12095)</td>
</tr>
<tr>
<td>MECOM_SKIL (AS)</td>
<td>2</td>
<td></td>
<td>duplication</td>
<td>exon1,intron3 (A12132) intron2,exon3 (A14308)</td>
</tr>
<tr>
<td>MECOM_NA</td>
<td>2</td>
<td></td>
<td>inversion</td>
<td>exon1,NA (A08245,A08257)</td>
</tr>
<tr>
<td>SEC62_MECOM</td>
<td>1</td>
<td></td>
<td>inversion</td>
<td>exon7,intron2 (A14184)</td>
</tr>
<tr>
<td>MECOM_NA</td>
<td>1</td>
<td></td>
<td>translocation</td>
<td>exon1,NA (A12110)</td>
</tr>
</tbody>
</table>
“HGS-OvCa demonstrates a remarkable degree of genomic disarray”

- Point 1
- Point 2

Partial and tandem duplications

- Table of most recurrent ITDs/PTDs to come from Karen
- E.g. MSLN PTD in 26/420 libraries
- ARID1A ~20/420
miRNA saturation in ovarian cancer

![miRNA Saturation in OV](chart.png)

- X-axis: # reads aligned to miRNAs (Millions)
- Y-axis: # miRNA species
- Legend:
  - 1x coverage
  - 10x coverage