RNA-based gene fusion detection: NGS applications and sample workflow improvements

Joshua Stahl
Chief Scientific Officer and General Manager
ArcherDX
Cancer is a disease of the genome

Genomic rearrangements
SNVs/indels
Copy number variations
RNA abundance
Fusions

- cDNA or DNA fragments
- End repair, d/A-tailing
- Adapter ligation
- Barcode
- P5 Primer
- GSP1
- GSP2

SNVs

- Relative Expression

CNVs

- ALK (CN=1.02)
- ALKRA (CN=2.49)
- CCND1 (CN=1.12)
- CDKN2A (CN=0.1)
- MYC (CN=0.22)
- NF1 (CN=0.89)
- RET (CN=0.97)
- MET (CN=24.59)

RNA abundance

For research use only. Not for use in diagnostic procedures.
Fusions

SNVs

cDNA or DNA fragments

End repair, d/A-tailing

Adapter ligation

GSP1

GSP2

CNVs

RNA abundance
Cancer is a disease of the genome

Genomic rearrangements
Sensitivity: False negatives - AMP vs. opposing primers

100% sensitivity

90% sensitivity

Opposing PCR methods
ALK fusion

- Tissue: lung
- Tumor: 80%
- Fusion detected: SQSTM1 → ALK
- Start sites: 100
- Not covered in competitor panels
ALK fusion detected by FISH

5' ALK

3' ALK
ALK fusion with Crizotinib resistance

Assay Result
- Strong Evidence Fusions: 3
  - EML4 → ALK
  - EML4 → ALK
  - EML4 → ALK
- Weak Evidence Fusions: 4
- Novel Isoforms: 16
- Variants Found: 1
  - ALK:p.C1156Y

QC Result
- FUSION QC: PASS
- VARIATION QC: PASS

Fusion EML4 → ALK

GSP2
- ALK_chr2_29446347_28_p_GSP2
  - Filters: ✓
  - Reads (#/%): 184 / 57.0
  - Start Sites (#/%): 29 / 52.7

Visualize | Translation | Quiver | Blast

For research use only. Not for use in diagnostic procedures.
Cancer is a disease of the genome

Genomic rearrangements/isoforms

SNVs/indels
Exon skipping/deletion detection

Wild-type MET

Exon 13  Exon 14  Exon 15

For research use only. Not for use in diagnostic procedures.
Exon skipping/deletion detection

MET variant

Exon 13  Exon 14  Exon 15

GSP2  GSP2  GSP2

Splice site mutation
MET exon 14 splice site mutation
MET exon 14 skipping

**Oncogenic Isoform MET**

<table>
<thead>
<tr>
<th>Filter</th>
<th>MET_chr7_116414939_29_n_GSP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reads</td>
<td>39 / 86.7</td>
</tr>
<tr>
<td>Start Sites</td>
<td>12 / 70.6</td>
</tr>
</tbody>
</table>

**Sample Name**

MRB544_S1_L001_R1_001

**Assay Result**

- Strong Evidence Fusions
- No Strong Evidence Fusions Detected
- Weak Evidence Fusions
- Oncogenic Isoforms
- MET
- Novel Isoforms

**FUSION QC:** PASS

---

For research use only. Not for use in diagnostic procedures.
Cancer is a disease of the genome.
Relative abundance detection

1. De-duplicate reads with MBC
2. Map reads to genome
3. Count MBCs associated with target regions and housekeeping controls
4. The counts are clustered into two groups
5. Calculate a relative abundance as a ratio of RNA specific reads

Moderate Relative Abundance

<table>
<thead>
<tr>
<th>MBC depth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

Internal controls  Target

High Relative Abundance

<table>
<thead>
<tr>
<th>MBC depth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

Internal controls  Target
Expression markers within panels help identify tissue of origin.
Expression markers within panels help identify tissue of origin.

Principal components extracted from per-exon RNA expression in normal lung, thyroid and adrenal inputs.
Cancer is a disease of the genome

SNVs/indels

RNA abundance
RNA + DNA variant calling

Wild-type DNA

Expressed wild-type RNA

Wild-type protein

DNA mutation

Expressed wild-type RNA

Mutant protein

Expressed mutant RNA

Mutant protein
RNA + DNA variant calling

Wild-type DNA

Expressed wild-type RNA

Wild-type protein

DNA mutation

Expressed wild-type RNA

Mutant protein

Expressed mutant RNA
RNA ideal for sensitive detection

D. Lung cancer case 46-3765

Source: PubMed
SNV detection in both RNA and DNA

NSCLC FFPE sample – EGFR L858R (AF = 22%)

Evidence of allelic imbalance – despite 22% AF, mutant allele is primarily expressed
Cancer is a disease of the genome
Orthogonal validation of CNV calls

Copy number variation

Normal sample

CNV sample

RNA abundance
Gene amplifications are confirmed by mRNA over-expression

<table>
<thead>
<tr>
<th>Driver mutation</th>
<th>Normal lung</th>
<th>NSCLC FFPE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EGFR L858R</td>
<td>CCND1 4.2X</td>
</tr>
<tr>
<td></td>
<td>22.5% AF</td>
<td>EGFR 15.2X</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EGFR 14.2X</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EGFR 3.5x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MET 24.5X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Copy number</th>
<th>CCND1</th>
<th>EGFR</th>
<th>MET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.6X</td>
<td>1.3X</td>
<td>1.1X</td>
</tr>
<tr>
<td></td>
<td>4.2X</td>
<td>1.1X</td>
<td>1.4X</td>
</tr>
<tr>
<td></td>
<td>1.9X</td>
<td>15.2X</td>
<td>0.8X</td>
</tr>
<tr>
<td></td>
<td>1.1X</td>
<td>14.2X</td>
<td>1.8X</td>
</tr>
<tr>
<td></td>
<td>1.3X</td>
<td>3.5X</td>
<td>0.9X</td>
</tr>
<tr>
<td></td>
<td>1.0X</td>
<td>1.0X</td>
<td>24.5X</td>
</tr>
</tbody>
</table>
Understanding your sample requires matching sequencer reads to input molecules.
Understanding your sample requires matching sequencer reads to input molecules.
Understanding your sample requires matching sequencer reads to input molecules

For research use only. Not for use in diagnostic procedures.
Understanding your sample requires matching sequencer reads to input molecules.
Archer PreSeq™ RNA QC Assay

- cDNA from universally expressed RNA
- PreSeq Assay

<table>
<thead>
<tr>
<th></th>
<th>PreSeq</th>
<th>UV Spectrometry, Fluorescent Dyes</th>
<th>Capillary Electrophoresis</th>
<th>Gel Electrophoresis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Length</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crosslinking</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

For research use only. Not for use in diagnostic procedures.
PreSeq predicts fragmentation and input
RNA integrity vs. age

- Empirically determined PreSeq cut off of Ct=29
- 100% sensitivity at Ct=29
- $p < 0.0001$ for linear trend
Increased input rescues library quality

- High-, medium- and low-quality FFPE samples
PreSeq is predictive of sensitivity and specificity
What about extraction?

- **Agencourt® FormaPure® Kit** (cat # A33341)
- **QIAGEN® AllPrep® DNA/RNA FFPE Kit** (cat # 80234)
- **Covaris® truXTRAC™ FFPE RNA Kit** (cat # 520161)
- **Ambion® RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE** (cat # AM1975)
- **Zymo Research Pinpoint™ Slide RNA Isolation System II** (cat # R1007)
- **Promega ReliaPrep™ FFPE Total RNA Miniprep System** (cat # Z1001)
- **Ambion® TR Izol® Reagent** (cat # 15596-026)

The extraction kits above are for research use only and not for use in diagnostic procedures.
What about extraction? – optimization

- Crosslinking reversal time
- Crosslinking temperature
- Crosslinking pH
- Divalent cation concentration
- Inclusion of a commercially available RNA storage reagent
Agencourt FormaPure - Additives

![Graph showing the relative amount of amplifiable VCP RNA across different conditions.](image-url)
Agencourt FormaPure – Crosslink reversal time and temperature

- 55°C/15h: 1.03
- 65°C/1h: 1.35
- 65°C/15h: 1.70
- 80°C/1h: 1.80
Covaris – Fusion calling and extraction

![Bar chart showing unique start sites for different conditions and samples](chart.png)

<table>
<thead>
<tr>
<th></th>
<th>EML4:ALK</th>
<th>CCDC6:RET</th>
<th>SLC34A2:ROS</th>
</tr>
</thead>
<tbody>
<tr>
<td>truXTRACT 80°C 15 min</td>
<td>3 of 3</td>
<td>2 of 3</td>
<td>3 of 3</td>
</tr>
<tr>
<td>truXTRACT 80°C 1 hour</td>
<td>3 of 3</td>
<td>3 of 3</td>
<td>3 of 3</td>
</tr>
<tr>
<td>truXTRACT 55°C O.N.</td>
<td>1 of 2</td>
<td>2 of 2</td>
<td>2 of 2</td>
</tr>
</tbody>
</table>
What about extraction? - Conclusion

The diagram shows a comparison of unique start sites for different DNA extraction methods under two conditions:

1. **55°C/16 hours crosslink reversal**
2. **80°C/1 hour crosslink reversal**

Methods compared include:
- Agencourt FormaPure
- Ambion RecoverAll
- Zymo Research Pinpoint
- Promega ReliaPrep
- QIAGEN AllPrep
- Covaris truXTRAC
- Universal RNA

The data is represented with error bars indicating variability. The methods are categorized by DNA and RNA content.
What about extraction? - Conclusion

• **Agencourt** FormaPure Kit - proteinase K digestion at 55°C for 1 hour followed by crosslink reversal at 80°C for 1 hour
• **Covaris** truXTRAC FFPE RNA Kit - crosslink reversal at 80°C for 1 hour
• **QIAGEN** AllPrep DNA/RNA FFPE Kit - crosslink reversal at 80°C for 1 hour; mineral oil is interchangeable with the QIAGEN® Deparaffinization Solution for sample deparaffinization.
• **Promega** ReliaPrep FFPE Total RNA Miniprep System - crosslink reversal at **80°C for 1 hour**
• **Zymo Research** Pinpoint Slide RNA Isolation System II - proteinase K digestion at 55°C for 1 hour followed by a crosslink reversal at 80°C for 1 hour; elute using 20μL water
• **Ambion** RecoverAll – **not recommended**
What about extraction? - Conclusion

- **Agencourt** FormaPure Kit - proteinase K digestion at 55°C for 1 hour followed by crosslink reversal at 80°C for 1 hour
- **Covaris** truXTRAC FFPE RNA Kit - crosslink reversal at 80°C for 1 hour
- **QIAGEN** AllPrep DNA/RNA FFPE Kit - crosslink reversal at 80°C for 1 hour; mineral oil is interchangeable with the QIAGEN® Deparaffinization Solution for sample deparaffinization.
- **Promega** ReliaPrep FFPE Total RNA Miniprep System - crosslink reversal at **80°C for 1 hour**
- **Zymo Research** Pinpoint Slide RNA Isolation System II - proteinase K digestion at 55°C for 1 hour followed by a crosslink reversal at 80°C for 1 hour; elute using 20μL water
- **Ambion** RecoverAll – **not recommended**
What about the Maxwell® RSC FFPE Kit?

For research use only. Not for use in diagnostic procedures.
Maxwell Instruments
Easy-to-Use Particle Mover Automation

- **Mixing**
  - Add Sample
  - LEV Plunger

- **Capture**
  - 1, 2, 3, 4, 5, 6, 7, 8

- **Binding**
  - 1, 2, 3, 4, 5, 6, 7, 8

- **Washing & Elution**
  - Elution Tube (30 - 100µl)

Easily process up to 16 samples simultaneously.
## Experimental design

<table>
<thead>
<tr>
<th>Library #</th>
<th>Tube Label</th>
<th>Extraction Type</th>
<th>Sample Type</th>
<th>RNA PreSeq CT value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PG1207</td>
<td>MF1</td>
<td>Maxwell</td>
<td>FFPE</td>
<td>25.5</td>
</tr>
<tr>
<td>PG1208</td>
<td>MF2</td>
<td>Maxwell</td>
<td>FFPE</td>
<td>26.8</td>
</tr>
<tr>
<td>PG1209</td>
<td>MN1</td>
<td>Maxwell</td>
<td>negative</td>
<td>26.0</td>
</tr>
<tr>
<td>PG1210</td>
<td>MN2</td>
<td>Maxwell</td>
<td>negative</td>
<td>24.7</td>
</tr>
<tr>
<td>PG1211</td>
<td>MN3</td>
<td>Maxwell</td>
<td>negative</td>
<td>25.1</td>
</tr>
<tr>
<td>PG1212</td>
<td>MN4</td>
<td>Maxwell</td>
<td>negative</td>
<td>25.2</td>
</tr>
<tr>
<td>PG1213</td>
<td>MP1</td>
<td>Maxwell</td>
<td>positive (SureShot)</td>
<td>25.4</td>
</tr>
<tr>
<td>PG1214</td>
<td>MP2</td>
<td>Maxwell</td>
<td>positive (SureShot)</td>
<td>24.9</td>
</tr>
<tr>
<td>PG1215</td>
<td>MP3</td>
<td>Maxwell</td>
<td>positive (SureShot)</td>
<td>25.6</td>
</tr>
<tr>
<td>PG1216</td>
<td>MP4</td>
<td>Maxwell</td>
<td>positive (SureShot)</td>
<td>25.6</td>
</tr>
<tr>
<td>PG1217</td>
<td>RF1</td>
<td>ReliaPrep</td>
<td>FFPE</td>
<td>27.2</td>
</tr>
<tr>
<td>PG1218</td>
<td>RF2</td>
<td>ReliaPrep</td>
<td>FFPE</td>
<td>26.5</td>
</tr>
<tr>
<td>PG1219</td>
<td>RN1</td>
<td>ReliaPrep</td>
<td>negative</td>
<td>25.6</td>
</tr>
<tr>
<td>PG1220</td>
<td>RN2</td>
<td>ReliaPrep</td>
<td>negative</td>
<td>26.6</td>
</tr>
<tr>
<td>PG1221</td>
<td>RN3</td>
<td>ReliaPrep</td>
<td>negative</td>
<td>26.4</td>
</tr>
<tr>
<td>PG1222</td>
<td>*RN5</td>
<td>ReliaPrep</td>
<td>negative</td>
<td>26.4</td>
</tr>
<tr>
<td>PG1223</td>
<td>RP1</td>
<td>ReliaPrep</td>
<td>positive (SureShot)</td>
<td>25.9</td>
</tr>
<tr>
<td>PG1224</td>
<td>RP2</td>
<td>ReliaPrep</td>
<td>positive (SureShot)</td>
<td>25.7</td>
</tr>
<tr>
<td>PG1225</td>
<td>RP3</td>
<td>ReliaPrep</td>
<td>positive (SureShot)</td>
<td>26.1</td>
</tr>
<tr>
<td>PG1226</td>
<td>*RP5</td>
<td>ReliaPrep</td>
<td>positive (SureShot)</td>
<td>25.5</td>
</tr>
</tbody>
</table>
Relative amplifiable VCP RNA (PreSeq)

- ReliaPrep positive: 1
- ReliaPrep negative: 0.76
- ReliaPrep FFPE: 0.50
- Maxwell positive: 1.40
- Maxwell negative: 1.50
- Maxwell FFPE: 0.83

For research use only. Not for use in diagnostic procedures.
Relative amplifiable VCP RNA (PreSeq)

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Relative Amount of Amplifiable VCP RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReliaPrep positive</td>
<td>1</td>
</tr>
<tr>
<td>ReliaPrep negative</td>
<td>0.76</td>
</tr>
<tr>
<td>ReliaPrep FFPE 1</td>
<td>0.37</td>
</tr>
<tr>
<td>ReliaPrep FFPE 2</td>
<td>0.62</td>
</tr>
<tr>
<td>Maxwell positive</td>
<td>1.40</td>
</tr>
<tr>
<td>Maxwell negative</td>
<td>1.50</td>
</tr>
<tr>
<td>Maxwell FFPE 1</td>
<td>1.19</td>
</tr>
<tr>
<td>Maxwell FFPE 2</td>
<td>0.47</td>
</tr>
</tbody>
</table>
Unique start sites

- **ReliaPrep positive**
  - Ambig: 234 ± 142
  - DNA: 1462 ± 234
  - RNA: 976 ± 142

- **ReliaPrep negative**
  - Ambig: 267 ± 142
  - DNA: 1389 ± 267
  - RNA: 922 ± 142

- **ReliaPrep FFPE 1**
  - Ambig: 620 ± 314
  - DNA: 1104 ± 620
  - RNA: 484 ± 314

- **ReliaPrep FFPE 2**
  - Ambig: 214 ± 110
  - DNA: 778 ± 214
  - RNA: 564 ± 110

- **Maxwell positive**
  - Ambig: 412 ± 234
  - DNA: 2376 ± 412
  - RNA: 1000 ± 234

- **Maxwell negative**
  - Ambig: 380 ± 196
  - DNA: 1968 ± 380
  - RNA: 588 ± 196

- **Maxwell FFPE 1**
  - Ambig: 727 ± 372
  - DNA: 1548 ± 727
  - RNA: 899 ± 372

- **Maxwell FFPE 2**
  - Ambig: 271 ± 169
  - DNA: 899 ± 271
  - RNA: 620 ± 169
Conclusion

- Archer FusionPlex assays offer distinct advantages over FISH and traditional target enrichment methods
- Archer Analysis software empowers users with multiple methods to support genetic abnormalities
- RNA quality on prospective FFPE sample higher than archived samples
- Extraction is an often overlooked but vital piece of any molecular assay
- Maxwell RSC instrument combined with Maxwell RSC RNA FFPE kit is the best-performing extraction technology on the market for FusionPlex assays
Acknowledgements

Archer Mol Bio Group
Jason Myers
Brian Kudlow
Brady Culver
Josh Haimes
Laura Johnson
Ryan Walters
Katie Moore
Paula Roberts
Namitha Manoj
Ian Hoskins
Marc Bessette
Megan Wemmer
Ian McKittrick
Aaron Garnett
James Covino
Eric Davis
Elina Baravik
Jens Eberlein
Helen Wang
External Advisors
Long Le
John Iafrate
Robert Daber

Archer Bioinformatics
Doug Wendel
Abel Licon
Jeremy Widmann
Aaron Berlin
Eric Reckase
Mike Banos
Jody Stephens
Mike Montgomery
Ben Van Deusen
Johnny Flowers
David Finke
Karamjeet Khalsa
Chris Davies
Vince Reuter
John Callaway

Collaborators
Cheng-Han Lee – University of Alberta
Valentina Nardi - MGH
Milhan Telatar – City of Hope
Danielle Murphy – Ignyta
Robert Shoemaker – Ignyta
Jason Christiansen - Ignyta

Archer Marketing
Jason Amsbaugh
Darius Fugere
Tyler Newberry
Rich Kulesus
Jared Snider