Heritability, polygenic risk scoring, and TWAS analyses for esophageal cancer

Oleg Borisov1*, Julia Schröder2, Michael Knapp3, Puya Gharahkhani4, Peter Krawitz1, Anne C. Böhmer2, Carlo Maj1, Johannes Schumacher2,3

Corresponding author
olegbor@uni-bonn.de

Introduction
Barrett’s esophagus (BE) is a premalignant precursor of Esophageal adenocarcinoma (EA) which is one of the leading causes of cancer deaths. Both conditions are complex traits and we aimed at dissecting their polygenic structure in association with tissue specific gene expression regulation.

Sample

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barrett’s esophagus (BE)</td>
<td>16.176</td>
</tr>
<tr>
<td>Esophageal adenocarcinoma (EA)</td>
<td>4,112</td>
</tr>
<tr>
<td>Controls</td>
<td>17,159</td>
</tr>
</tbody>
</table>

Meta-analysis of 4 studies (Gharahkhani et al., Lancet Oncology 2016)

Specific gene expression regulation.

1 – Heritability (h²) and genetic correlation
Method – LD Score Regression (LDSC)

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Barrett’s esophagus</th>
<th>Esophageal adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>h² (SE)</td>
<td>0.20 (0.03)</td>
<td>0.13 (0.02)</td>
</tr>
</tbody>
</table>

Liability scale conversion was performed using the lifetime risk estimates 1.6% for BE, 0.25% for EA (Ek et al, 2013)

Genetic correlation between BE and EA

r²=0.85 (0.06), (p=4E-41)

2 – Polygenic risk scoring (PRS)

Formula of PRS

\[ PRS_i = \sum_{j=1}^{n} \ln(OR)_{ij} \times c_i \]

PRS_i = polygenic risk score of sample i

n = selected variants

\( \ln(OR) \) = weight (base sample)

\( c_i \) = risk allele counts (target sample)

2 independent data sets:

<table>
<thead>
<tr>
<th>Base data</th>
<th>Target data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2406 BE</td>
<td>1037 BE</td>
</tr>
<tr>
<td>1508 EA</td>
<td>1609 EA</td>
</tr>
<tr>
<td>4,112 BE</td>
<td>3,537 controls</td>
</tr>
<tr>
<td>2406 BE</td>
<td>1508 EA</td>
</tr>
<tr>
<td>4,112 BE</td>
<td>2,357 controls</td>
</tr>
</tbody>
</table>

Comparison        Summary statistics   Model p-value   R²
BE vs controls    BE vs controls    3.00E-05        0.56%
EA vs controls    EA vs controls    3.00E-07        0.67%
BE-EA vs controls BE-EA vs controls 5.00E-13        1.09%
BE vs controls    MDD vs controls   0.09            0.15%

3 – Functional Mapping and Annotation of GWAS (FUMA)

Significance threshold after multiple correction

Gastroesophageal Junction tissue

Summary

1) Relatively high SNP-h² and genetic correlation for BE and EA.
2) Strong polygenic effect associated with case-control status.
3) Genetic signal can be driven by the tissue-specific gene-expression regulation.
4) Potential candidate genes in relevant tissues identified by means of TWAS.
5) Tissue-specific GE-PRS prioritize esophagus mucosa tissue

4 – Transcriptome-wide association study (TWAS)

5 – Gene-expression polygenic risk scores (GE-PRS)