Biological Drivers of Breast Invasive Carcinoma in Black/African-American Patients

G. Acquaah-Mensah, Ph.D.

Presented to *Toxicologists of African Origin* SIG, Society of Toxicology

October 29th, 2021
Network Inference Algorithms Elucidate Nrf2 Regulation of Mouse Lung Oxidative Stress

Ronald C. Taylor1,2, George Acquah-Mensah3, MuditaaSingh1, Deepita Malhotra1, Shyam Biswal2

1 Computational Biology and Bioinformatics Group, Pacific Northwest National Laboratory, U.S. Department of Energy, Richland, Washington, United States of America; 2 Department of Pharmacological Sciences, Massachusetts College of Pharmacy and Health Sciences, Worcester, Massachusetts, United States of America; 3 Department of Environmental Health Sciences, Bloomberg School of Pub

Suppressed Expression Involved in Senesce Pulmonary Disease

George K. Acquah-Mensah1,9, Deep Shyam Biswal2

1 Department of Pharmacological Sciences, Massachusetts College of Environmental Health Sciences, Bloomberg School of Public Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America; 2 University of North Carolina at Chapel Hill, Raleigh-Durham, North Carolina, United States of America; 3 Computational Biology and Bioinformatics Group Pacific Northwest National Laboratory, Richland, Washington, United States of America

NRF2 Activation Promotes Aggressive Lung Cancer and Associates with Poor Clinical Outcomes

Anju Singh1, Analeena Daemen2, Dorothy Nickels2, Sang-Min Jeon3,4, Oded Foreman5, Kuladheep Sudini6, Florian Gnad2, Stepanha Lopite1, Naina Gour1, Wayne Mittner1, Samir Chatterjee1, Eun-Ji Choi1, Buvana Ravishankar1, Amy Rappaport1, Namrata Patil1, Mark McClendon1, Lesia Johnson1, George Acquah-Mensah6, Edward Gabrielson1, Shyam Biswal1, and George Hatzivassiliou2,5

herapeutic target in ncer

Iey1, Qiong Kang1, George Acquah-Mensah2,5, Naincio L. Wonders1, Vassilidou Papadimitrakopoulos4, a Carretero5, Kwok-Kin Wong1, John D. Hale1, 83

Health, 3 Department of Pathology, and Sidney Kimmel Health Sciences, Hopkins University, Baltimore, Maryland, USA; 4 Department of Biochemistry and Molecular Cell Biology, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA; 5 Department of Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA; 6 Department of Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA

Research article

Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis

Anju Singh1, Christine Happel1, Soumen K. Manna1, George Acquah-Mensah2, Julian Carretero3, Sarvesh Kumar1, Poonam Nasipuri1, Kristopher Krausz1, Nobunao Wakabayashi4, Ruby Dew1, Leoel G. Boros5, Frank J. Gonzalez6, Edward Gabrielson7, Kwok K. Wong8,9, Geoffrey Gimmi8,9, and Shyam Biswal1

1 Laboratory of Metabolism, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, USA; 2 Department of Pharmacological Sciences, Massachusetts College of Pharmacy and Health Sciences, Worcester, Massachusetts, USA; 3 Department of Physiology, University of Valencia, Valencia, Spain; 4 Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; 5 Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, Maryland, USA; 6 Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA; 7 Department of Genetics, Sidney Kimmel Comprehensive Cancer Center, School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA; 8 Department of Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA

Machine learning approaches to decipher hormone and HER2 receptor status phenotypes in breast cancer

Emmanuel S. Adabor and George K. Acquah-Mensah

Corresponding author: Emmanuel S. Adabor, African Institute for Mathematical Sciences, 4 Meito Road, Munsieberg, Cope Town 7945, South Africa. Tel.: +27819466598; E-mail: heid79@gmail.com
NATIONAL CANCER INSTITUTE
THE CANCER GENOME ATLAS

TCGA BY THE NUMBERS

TCGA produced over
2.5
PETabytes
of data

To put this into perspective, 1 petabyte of data is equal to
212,000
DVDs

TCGA data describes
33
DIFFERENT TUMOR TYPES
...including
10
RARE CANCERS

...based on paired tumor and normal tissue sets collected from
11,000
PATIENTS

...using
7
DIFFERENT DATA TYPES

TCGA RESULTS & FINDINGS

MOLECULAR BASIS OF CANCER
Improved our understanding of the genomic underpinnings of cancer

TUMOR SUBTYPES
Revolutionized how cancer is classified

TCGA revolutionized how cancer is classified by identifying tumor subtypes with distinct sets of genomic alterations.*

THERAPEUTIC TARGETS
Identified genomic characteristics of tumors that can be targeted with currently available therapies or used to help with drug development

TCGA’s identification of targetable genomic alterations in lung squamous cell carcinoma led to NCI’s Lung-MAP Trial, which will treat patients based on the specific genomic changes in their tumor.

THE TEAM

20
COLLABORATING INSTITUTIONS
across the United States and Canada

WHAT’S NEXT?

The Genomic Data Commons (GDC) houses TCGA and other NCI-generated data sets for scientists to access from anywhere. The GDC also has many expanded capabilities that will allow researchers to answer more clinically relevant questions with increased ease.

*TCGA’s analysis of stomach cancer revealed that it is not a single disease, but a disease composed of four subtypes, including a new subtype characterized by infection with Epstein-Barr virus.

www.cancer.gov/ccg
Survival in BrCA, Race and Stage: All

$p = 0.85$
Survival in BrCA: Stage 3 and 4, Over 60 Years

$p = 0.081$
## TCGA July 2016: Patients

### Number of patients **up to age 50 at diagnosis** in dataset at different stages of BrCA

<table>
<thead>
<tr>
<th>Stage</th>
<th>Black or African-American</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/ IA/ IB</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td>II/ IIA/ IIB</td>
<td>39</td>
<td>121</td>
</tr>
<tr>
<td>III/ IIIA/ IIIB/ IIIC</td>
<td>14</td>
<td>62</td>
</tr>
<tr>
<td>IV</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

### Number of patients **of all ages at diagnosis** in dataset at different stages of BrCA

<table>
<thead>
<tr>
<th>Stage</th>
<th>Black or African-American</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/ IA/ IB</td>
<td>32</td>
<td>142</td>
</tr>
<tr>
<td>II/ IIA/ IIB</td>
<td>106</td>
<td>415</td>
</tr>
<tr>
<td>III/ IIIA/ IIIB/ IIIC</td>
<td>35</td>
<td>176</td>
</tr>
<tr>
<td>IV</td>
<td>5</td>
<td>11</td>
</tr>
</tbody>
</table>

Number of patients with “number of days to death” annotation even fewer.
Survival in BrCA: Stage 2

All Ages

50 Years Old or Younger

$\text{Survival probability}$

$p = 0.93$

$p < 0.0001$
Survival in BrCA

Stage 3 (no 3a) and 4, Over 60 Years

Stage 3, 3b, 3c, and 4
TERT Gene Expression in Samples from Caucasian and B/AA Patients

RB Expression in Samples from White and B/AA Patients

hsa-let-7a Expression in Samples from Caucasian and B/AA Patients

IMPACT OF RELATIVE CHANGES IN miR AND GENE EXPRESSION IN BLACKS

1. E2F1
2. RB1
3. hsa-let-7a
4. cMYC
5. TERT

- E2F targets and multiple cell cycle regulators and etc...
- Entrance into S phase of cell cycle

*In renal cell carcinoma cells Liu Y et al., 2012
*In primary fibroblasts Bueno MJ et al., 2010; absence of miRs leads to significant increase in S-phase entrance and DNA damage

Red -> White (n=849); Blue -> B/AA (n=188)
Immortalization Gene Signature Expression in Primary Tumor Cultures

Oncogene (2007) 26, 6269–6279; doi:10.1038/sj.ong.1210452
IMMORTALIZATION SIGNATURE GENES

WHITE PATIENTS (n=925)

BLACK OR AFRICAN-AMERICAN PATIENTS (n=189)
E. coli TRANSCRIPTIONAL REGULATORY NETWORK INFERRED USING CONTEXT LIKELIHOOD OF RELATEDNESS

Faith JJ et al., 2007
Mutual Information

• For two discrete random variables $X$ and $Y$, mutual information:

$$I(X; Y) = \sum_{i,j} P(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

For genes, $X$ and $Y$ represent a transcription factor and its potential target gene, and $x_i$ and $y_i$ represent particular expression levels.

• For continuous variables integrals are used in place of summation.
<table>
<thead>
<tr>
<th>p-value</th>
<th>Set</th>
<th>Act</th>
<th>Exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000118</td>
<td>HMGA1</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>0.000176</td>
<td>TERT</td>
<td>643</td>
<td></td>
</tr>
<tr>
<td>0.000234</td>
<td>ATF4</td>
<td>521</td>
<td></td>
</tr>
<tr>
<td>0.000575</td>
<td>TCF3</td>
<td>991</td>
<td></td>
</tr>
<tr>
<td>0.000878</td>
<td>HES4</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>0.00107</td>
<td>HSF1</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>0.00121</td>
<td>PBP1</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>0.00154</td>
<td>FOXE3</td>
<td>3492</td>
<td></td>
</tr>
<tr>
<td>0.00159</td>
<td>DRAP1</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>0.00186</td>
<td>TAF6</td>
<td>669</td>
<td></td>
</tr>
<tr>
<td>0.00203</td>
<td>IRF3</td>
<td>231</td>
<td></td>
</tr>
<tr>
<td>0.00189</td>
<td>ZNF396</td>
<td>1775</td>
<td></td>
</tr>
<tr>
<td>0.00111</td>
<td>CREB1</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>0.00105</td>
<td>DENND4A</td>
<td>269</td>
<td></td>
</tr>
<tr>
<td>0.000957</td>
<td>SP1</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>0.000578</td>
<td>BDP1</td>
<td>290</td>
<td></td>
</tr>
<tr>
<td>0.000538</td>
<td>THRB</td>
<td>843</td>
<td></td>
</tr>
<tr>
<td>0.000268</td>
<td>ARID4A</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>0.000175</td>
<td>RB1</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>0.000175</td>
<td>AR</td>
<td>888</td>
<td></td>
</tr>
</tbody>
</table>
ARACNe-derived, 2 races only, all ages; green nodes have suppressed expression in stage iii in Black/African Americans, white nodes have elevated expression
ARACNe-derived, 2 races only, all ages; green nodes have suppressed expression in stage iii in Black/African Americans, white nodes have elevated expression.
AHR and its Dimerization Partner ARNT Significantly Suppressed among Blacks
Non-response to Tamoxifen, Anastrozole, Lestrozole, Fulvestrant, Exemestane, Goserelin, or Leuprolide
Differentially Mutated at Stage II

BAA, Up To Age 50 (n = 29) v/s White, Up To Age 50 (n = 115)

HUWE1

PIK3CA

log10 (Odds Ratio)

BAA, Up To Age 50  White, Up To Age 50

TCGA 2015
British Journal of Cancer (2006) 94(11), 1555–1558
Genes Most Mutated between White and B/AA Breast Cancer Patients
Difference in Survival between HUWE1-mutated and PIK3CA-mutated BRCA Patients

HUWE1-mutated: n=5; PIK3CA-mutated: n=22

p = 0.52
## BRCA Molecular Subtypes

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Expression of steroid receptors, HER2, and cytokeratins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminal A</td>
<td>ER+ and/or PgR+, HER2–, CK5/6–</td>
</tr>
<tr>
<td>Luminal B</td>
<td>ER+ and/or PgR+, HER2+, CK5/6–</td>
</tr>
<tr>
<td>Basal-like</td>
<td>ER–, PgR–, HER2–, CK5/6+</td>
</tr>
<tr>
<td>HER2+</td>
<td>ER–, PgR–, HER2+, CK5/6–</td>
</tr>
<tr>
<td>Normal breast-like</td>
<td>Cancers not classified in mentioned subtypes</td>
</tr>
</tbody>
</table>

*Contemp Oncol (Pozn). 2016; 20(6): 436–443*
### PAM50 Molecular Subtype Predictions for HUWE1 and PIK3CA Mutants

<table>
<thead>
<tr>
<th></th>
<th>Basal-like</th>
<th>HER2-enriched</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with Either Mutation (n=61)</td>
<td>7 (11.5%)</td>
<td>6 (9.8%)</td>
<td>23 (37.7%)</td>
<td>21 (34.4%)</td>
<td>4 (6.6%)</td>
</tr>
<tr>
<td>All Patients with HUWE1 Mutation (n=10)</td>
<td>6 (60%)</td>
<td>0 (0%)</td>
<td>3 (30%)</td>
<td>1 (10%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>B/AA Patients with HUWE1 Mutation (n=5)</td>
<td>2 (40%)</td>
<td>0 (0%)</td>
<td>1 (20%)</td>
<td>1 (20%)</td>
<td>1 (20%)</td>
</tr>
<tr>
<td>All Patients with PIK3CA Mutation (n=51)</td>
<td>3 (5.9%)</td>
<td>5 (9.8%)</td>
<td>14 (27.5%)</td>
<td>23 (45.1%)</td>
<td>6 (11.8%)</td>
</tr>
<tr>
<td>White Patients with PIK3CA Mutation (n=38)</td>
<td>3 (7.9%)</td>
<td>5 (13.2%)</td>
<td>10 (26.3%)</td>
<td>15 (39.5%)</td>
<td>5 (13.2%)</td>
</tr>
</tbody>
</table>

Note: stage II patients aged 50 or younger at diagnosis
Differential Expression of Ubiquitination, SUMOylation, and TP53 Regulation Genes in HUWE1/BAA Patients vs. PIK3CA/White Patients

age 50 or younger at diagnosis
<table>
<thead>
<tr>
<th></th>
<th>Basal-like</th>
<th>HER2-Enriched</th>
<th>Luminal B</th>
<th>Luminal A</th>
<th>Normal-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALL SUBJECTS (n=1213)</strong></td>
<td>16.4% (n=199)</td>
<td>6.3% (n=77)</td>
<td>49.8% (n=604)</td>
<td>18.1% (n=220)</td>
<td>9.3% (n=113)</td>
</tr>
<tr>
<td><strong>WHITE SUBJECTS (n=864)</strong></td>
<td>13.8% (n=119)</td>
<td>5.1% (n=44)</td>
<td>48.1% (n=416)</td>
<td>20.9% (n=181)</td>
<td>12% (n=104)</td>
</tr>
<tr>
<td><strong>BLACK/AFRICAN-AMERICAN SUBJECTS (n=189)</strong></td>
<td>34.9% (n=66)</td>
<td>9.5% (n=18)</td>
<td>42.9% (n=81)</td>
<td>10% (n=19)</td>
<td>2.6% (n=5)</td>
</tr>
</tbody>
</table>

BREAST CANCER SUBTYPE PREDICTIONS USING THE PAM50 ALGORITHM -TCGA 2016
# Triple-Negative Breast Cancer (TNBC) Patients by Race - TCGA

<table>
<thead>
<tr>
<th></th>
<th>ALL (n=358)</th>
<th>B/AA (n=75)</th>
<th>WHITE (n=235)</th>
<th>ASIAN (n=18)</th>
<th>OTHER (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TNBC</strong></td>
<td>126 (35%)</td>
<td>42 (56%)</td>
<td>69 (29%)</td>
<td>8 (44%)</td>
<td>7 (23%)</td>
</tr>
<tr>
<td><strong>non-TNBC</strong></td>
<td>232 (64%)</td>
<td>33 (44%)</td>
<td>166 (71%)</td>
<td>10 (56%)</td>
<td>23 (77%)</td>
</tr>
</tbody>
</table>
Word Plot: Mutations found in non-TNBC samples (TCGA, 2016)
Driving the Differences in Gene Expression between Races in Stage III

Note: TRN based on samples from 2 races only, all ages; DE between 2 races, stage iii only, all ages
The Ubiquitin Proteasome System

B. Suresh et al., 2016; http://dx.doi.org/10.1155/2016/6705927
Selected SUMOylation/Ubiquitination Genes in TNBC - TCGA All Ages

non-TNBC n=256

TNBC n=127

SUMO1
UBE2L3
HMGA1
UBE2G
UBE2M
RNF4
PIAS1
PIAS2
UBE2S
UBE2C
SUMOylation of Intracellular Receptor Targets - TCGA All Ages

Non-TNBC n=256

TNBC n=127

Color Key and Histogram

Row Z-Score

0

ESR1
RORA
NR3C2
NR5A2
THR9
NR4A2
AR
THRA
PPARG
NR3C1
Leading Edge Genes from PuriNet in TNBC

HDAC1, PAK3, NUP153, SUPT3H, DAXX, SUMO1, NOLC1, LCK, RFC3, HMGB1, SOCS1, NUDT1, RPA2, NFKB1B, ANAPC1, NOC2L, ZFP42, PCNA, CCNB1, NUP50, MYC, HMMR, CEBPZ, COPS3, RAD51, CDK1, CENPE, AURKA, MYBBP1A, RPA1, CKS2, MCM4, PSIP1, BUB1B, PFBN1, MCM2, GMNN, PLSCR1, DDX11, MCM10, CHEK2, MCM6, MYBL2, TRIP13, RFC2, MCM3, MSH6, DCC1, C51B, CCNA2, PRNP, CDC7, EXO1, CDT1, BAK1, BLM, CDC45, BIRC5, CCNB2, CHEK1, RFC4, HDAC2, TPX2, MSH2, POU5F1, MCM5, HMGA1, RAD54L, AURKB, and CCNE1
Selected “STEMness” Genes in TNBC
Summary

• B/AA more likely to present with the more aggressive basal-like or HER2-enriched molecular subtype than White patients
• Mutation patterns differ: PIK3CA, HUWE1
• Gene expression patterns indicate enhanced entry into S-phase of Cell Cycle
• OncomiR miR-221 and migration- and invasion-promoting miR-135b expressed at higher levels in BAA
• Master regulators include: TERT, HMGA1, HES4, PQBP1, RB1, and AR
• TERT knockdown results in apoptosis and decreased cell viability
• TERT-linked immortalization gene signature enriched in B/AA
• Biological differences exist between BAA non-responders to endocrine therapy and White non-responders.