Identifying hub genes among different breast cancer subtypes and racial diversity through integrated bioinformatics analysis (WGCNA)

PS4-05-17

Daniel Lee-Kay-Pen, MD[1]; Yomali Ferreyra, MSc[2]; Zaida Morante, MD[3–4]; Henry L. Gómez, MD[4–5]
[1]Clínica Padre Luis Tezza, Department of Medical Oncology, Lima, Peru, [2]Health Innovation Laboratory, Institute of Tropical Medicine "Alexander von Humboldt", Universidad Peruana Cayetano Heredia, Lima, Peru,
[3]Instituto Nacional de Enfermedades Neoplásicas, Lima, Peru, [4]Oncosalud–AUNA, Lima, Peru[5]Universidad Peruana Cayetano Heredia, Lima, Peru

SAN ANTONIO BREAST CANCER SYMPOSIUM®

Introduction

Focusing only on driver genes may ignore important interactions and pathways in breast cancer across subtypes and racial groups. Understanding their dysfunctional network or hub genes is crucial, as these are related to mRNA expression patterns. Our goal is to identify key hub genes related to race and subtype.

Methods and Materials

RNA-seq counts and clinical data from TCGA (GSE62944) were filtered to include only invasive breast cancer cases with complete information on race (White, Black, Asian), intrinsic subtype (Luminal A/B, HER2-enriched, Basal-like), and survival status. Differential expression analysis was performed using DESeq2 to compare tumor samples from deceased versus living patients within each race×subtype group. Genes with $|\log_2 \text{fold change}|>1$ and FDR<0.05 were considered differentially expressed (DEGs). All DEGs were used to apply the WGCNA algorithm in three stages: (1) construction of weighted gene correlation networks, (2) identification of coexpression modules, and (3) correlation of module eigengenes with clinical traits (race, subtype, and their interaction). Hub genes were defined as those with high module membership (MM>0.8) and gene significance (GS>0.2). We prioritized hub genes that were prevalent across multiple traits.

Results

Figure 1 shows a total of 3548 unique DEGs that were identified across intrinsic subtypes and race. In **Figure 2**, a total of 7 gene coexpression modules were identified

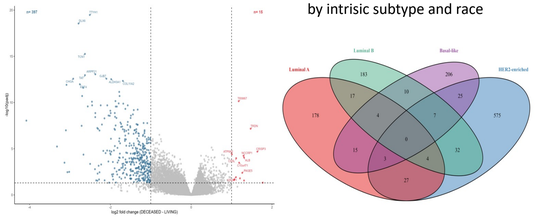
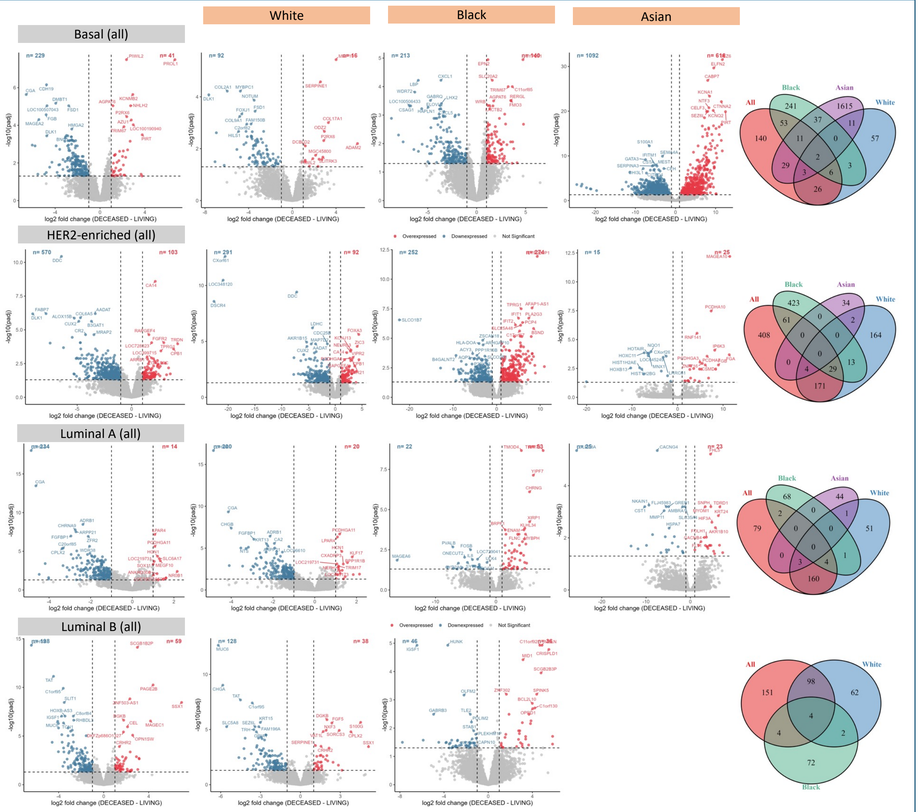


Figure 1: Volcanos plots stratified by intrisic subtype and race

The turquoise module was the largest (709 genes), followed by blue (538 genes) and brown (407 genes). However, the last one exhibited the strongest inverse correlation with basal-like tumors across racial groups, including both White (R=-0.66, p<0.001) and Black (R=-0.47, p<0.001) patients.

Results



In contrast, it showed a positive association with luminal A tumors, particularly among White patients (R=0.57, p<0.001). **Table 1** shows several hub genes from brown module that demonstrated consistent associations across intrinsic subtypes and racial groups. Notably, C6orf97 and ESR1 showed high MM and GS in more than 50% of the evaluated traits, including basal-like, HER2-enriched, luminal A/B subtypes, and race-specific combinations (Basal-Black, Luminal-White)

Results

Figure 2: Module – trait relationships

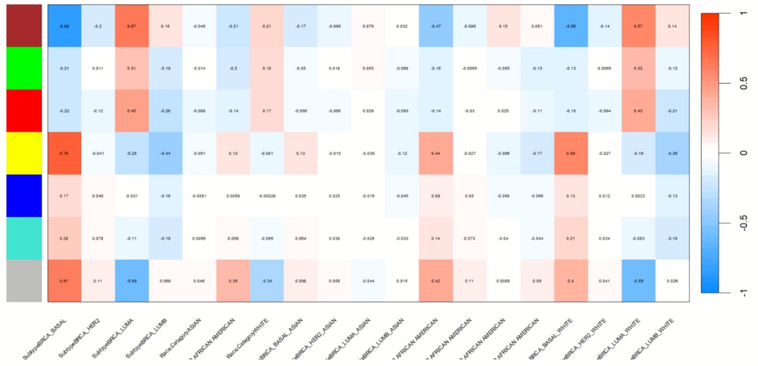


Table 1: Prevalence of hub genes obtained from brown module

GENE	n TRAITS	TRAITS: INTRINSIC SUBTYPES BASAL	HER 2	LUM A	LUM B	RACE BLACK	WHITE	RACE AND INTRINSIC SUBTYPES BLACK BASAL	WHITE BASAL	WHITE HER 2	WHITE LUM A	WHITE LUM B	PREVALENCE
C6orf97	10	X	X	X	X	X		X	X	X	X	X	0.53
ESR1	10	X	X	X	X		X	X	X	X	X	X	0.53
TBC1D9	8	X		X	X	X	X	X	X		X		0.42
AGR3	7	X	X	X	X			X	X		X		0.37
DNALI1	7	X	X	X	X			X	X		X		0.37
GATA3	7	X		X	X			X	X		X	X	0.37
THSD4	7	X	X	X		X		X	X		X		0.37
LOC402160	5	X		X				X	X		X		0.26
PRR15	5	X		X				X	X		X		0.26
SLC44A4	5	X		X				X	X		X		0.26

Discussion

Our findings reveal that breast cancer gene networks vary across subtypes and racial groups. The strong correlation of the brown module with luminal and basal-like tumors suggests subtype-specific regulation. Hub genes such as ESR1 and GATA3 reaffirm their central role in luminal differentiation, while C6orf97, TBC1D9, and AGR3 may reflect additional regulatory mechanisms influenced by race. These results highlight the value of network-based analyses like WGCNA to uncover shared and race-specific molecular drivers that extend beyond traditional single-gene approaches.

Conclusions

Using WGCNA, we identified 10 hub genes consistent across intrinsic subtypes, but not when considering race alone. Basal and Luminal A hub genes were unaffected by race, while those related to HER2 and Luminal B showed minor correlations when stratified by race.

Contact

Daniel Lee-Kay Pen
Clinica Padre Luis Tezza
Email: tpklde@gmail.com
Phone: (+51) 997566993

References

P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," BMC Bioinformatics, vol. 9, no. 1, pp. 559–559, Dec. 2008, doi: https://doi.org/10.1186/1471-2105-9-559.
Qiu J, Du Z, Wang Y, Zhou Y, Zhang Y, Xie Y, Lv Q. Weighted gene co-expression network analysis reveals modules and hub genes associated with the development of breast cancer. Medicine (Baltimore). 2019 Feb;98(6):e14345. doi: 10.1097/MD.0000000000014345. PMID: 30732163; PMCID: PMC6380702.

This presentation is the intellectual property of the author/presenter. Contact them at tpklde@gmail.com for permission to reprint and/or distribute.

© Poster Template by Genigraphics® 1.800.790.4001 www.genigraphics.com