

THE RS6942067 GG GENOTYPE IS ASSOCIATED WITH A WORSE OVERALL SURVIVAL IN HEAD AND NECK SQUAMOUS CELL CARCINOMA AND WITH N+ PATIENTS EITHER UNDER 42 YEARS OLD OR NON-SMOKING HPV NEGATIVE

Guillaume B. Cardin¹; Monique Bernard¹; Jessica Bourbonnais¹; Houda Bahig^{1,2}; Felix-Phuc Nguyen-Tan²; Edith Filion²; Denis Soulieres²; Danielle Charpentier²; Olguta Gologan²; Tareck Ayad^{1,2}; Louis Guertin²; Eric Bissada²; Jean-Claude Tabet²; Francis Rodier¹; Apostolos Christopoulos^{1,2}

¹Centre de Recherche du Centre Hospitalier de l'Université de Montréal; ²Centre hospitalier de l'Université de Montréal

Please send questions and requests to guillaume.cardin.chum@sssss.gouv.qc.ca and a.christopoulos@umontreal.ca

Abstract

Background: We previously showed that the DCBLD1 promoter region harbors a single nucleotide polymorphism (rs6942067) for which the GG genotype increases the risk of young and non-smoking HPV negative head and neck squamous cell carcinoma (HNSCC). A similar risk has been shown for female non-smoking lung cancer. DCBLD1 gene expression has been shown to have a prognostic value in HNSCC and HNSCC patients with the rs6942067 GG genotype (GG patients) have higher DCBLD1 gene expression then other patients. High expression of DCBLD1 has also been associated with a worse overall survival in non-small cell lung carcinoma and invasive breast carcinoma.

Objective: To evaluate rs6942067 GG associations with overall survival and clinical parameters in HNSCC.

Methods: Retrospective study of 100 HNSCC patients of the Montreal University Hospital Center (CHUM) and 155 HNSCC patients of the Cancer Genome Atlas (TCGA). Genotype of rs6942067 was evaluated for every participants and rs6942067 GG patients were compared to other patients for overall survival and clinical characteristics.

Results: 37 participants have rs6942067 GG genotype and 218 have AA or AG phenotype. GG patients have a worse overall survival than AA or AG patients ($p = 0.038$). GG patients have significantly ($p = 0.01$) higher occurrence of N+ tumors (67.6%) than AA or AG patients (45.8%). N+ patients ≤ 42 years old or with no classical HNSCC risk factors (never-smoking and HPV negative) have an higher concurrency of rs6942067 GG (34.4%) than N0 patients ≤ 42 years old or with no classical HNSCC risk factors (11.4%, $p = 0.02$), other N+ patients (15.6%, $p = 0.02$), other N0 patients (8.7%, $p = 0.0005$), and from the healthy population (11.7%, $p < 0.0001$).

Conclusion: A single nucleotide polymorphism (rs6942067) is causing an upregulation of DCBLD1 gene expression is associated with a worse overall survival and N+ tumors in HNSCC. It is also associated with young (≤ 42 years old) and with non-smoking HPV negative HNSCC.

Introduction

While most head and neck squamous cell carcinoma (HNSCC) patients have either a history of tobacco use or human papillomavirus (HPV) infection, a subset of young and/or no classical HNSCC risk factors patients exists for which hereditary genetic factor are yet to be identified.¹ The DCBLD1 gene was first reported in a genome-wide association study that identified a susceptibility locus in the DCBLD1 promoter in women with never-smoking lung cancer.² In HNSCC, the rs6942067 GG genotype (SNP in the DCBLD1 promoter) was associated with a higher expression of DCBLD1 and with a higher risk to develop HNSCC for patients with no classical HNSCC risk factors (HPV negative non-smokers) and for patients under 40 years old.³ High DCBLD1 gene expression was associated with a worse overall survival and upregulation of the integrin signaling pathway, which is important notably for cell migration, in HNSCC, non-small cell lung carcinoma and invasive breast carcinoma.^{3,4} This study aims to better define the impact of rs6942067 GG genotype in HNSCC in a larger cohort.

Methods

A total of 100 HNSCC patients were recruited at the Centre Hospitalier de l'Université de Montréal (CHUM) between 2010 and 2019. Sequencing of rs6942067 (chr6:117464533, A>G, GRCh38.p12) was done by Genome Quebec using a “TGACTGTGGTGTTTGAATGGA” primer on germline gDNA obtained from blood PBMCs. This study also included the 155 previously genotyped participants of the Cancer Genome Atlas (TCGA) HNSCC cohort (TCGA, Provisional).^{3,5} As a reference cohort for rs6942067 status, we used the 15654

of the Genome Aggregation Database (gnomAD) from the Genotype-Tissue Expression (GTEx) project.⁶ Lifelong non-smokers and previous smokers for over 15 years with less than 10 pack-year were classified as non-smokers and others as smokers. The age cutoff at which rs6942067 GG genotype is associated with young HNSCC was evaluated by calculating the threshold maximizing the prediction value on a ROC curve associating age to the rs6942067 GG genotype. This study was approved by our Institutional Ethics Review Committee (#19.323). Statistical analysis are Pearson's chi-square test and log-rank test. They were performed using JMP 12.0.1 statistical software (SAS Institute Inc.).

Results

HNSCC patients of the CHUM and TCGA cohorts were combined and stratified in a rs6942067 AA or AG ($n = 218$) and a rs6942067 GG ($n=37$) group and differences in demographic and clinical characteristics were evaluated (Table). The only significant association was with nodal status, rs6942067 GG patients having a significantly higher occurrence of N+ tumors than AA or AG patients. Young patients (≤ 42 years old) and patients with no classical HNSCC risk factors were pooled together (young and/or no risk factor) under

Table. Demographic and clinical characteristics of HNSCC patients stratified by rs6942067 genotype

Characteristics	rs6942067 (AA/AG)	rs6942067 (GG)	P
Total	218	37	
Cohort, n (%)			0.58
CHUM	87 (40)	13 (35)	
TCGA	131 (60)	24 (65)	
Age group, n (%)			0.14
≤ 42 years old	23 (11)	7 (19)	
>42 years old	195 (89)	30 (81)	
Sex, n (%)			0.35
Female	57 (26)	7 (19)	
Male	161 (74)	30 (81)	
Anatomical subsite, n (%)			0.36
Oral tongue	90 (41)	13 (35)	
Other oral cavity	57 (26)	11 (30)	
Oropharynx	39 (18)	4 (11)	
Larynx and hypopharynx	32 (15)	9 (24)	
Smoking Status, n (%)			0.74
Current / previous ≥ 10 pack-year	159 (73)	26 (70)	
Never / previous < 10 pack-year	59 (27)	11 (30)	
HPV, n (%)			0.59
Positive	50 (23)	7 (19)	
Negative	168 (77)	30 (81)	
Tumor size, n (%)			0.25
T1-2	97 (44)	13 (35)	
T3+	117 (54)	24 (65)	
Tx, unknown	4 (2)	0 (0)	
Nodal status, n (%)			0.01
N0	115 (53)	12 (32)	
N+	97 (44)	25 (68)	
Nx, unknown	6 (3)	0 (0)	

the assumption that HPV and tobacco should not had time to induce cancer for young patients. Occurrence of rs6942067 GG genotype was significantly higher in young and/or no risk factor HNSCC with positive nodal status than in other HNSCC groups and than the healthy reference population (Figure 1). There was no significant difference between the other HNSCC groups and the healthy reference population. The impact of rs6942067 genotype on patients' overall survival was also evaluated showing that the rs6942067 GG is associated with a worse HNSCC overall survival (Figure 2).

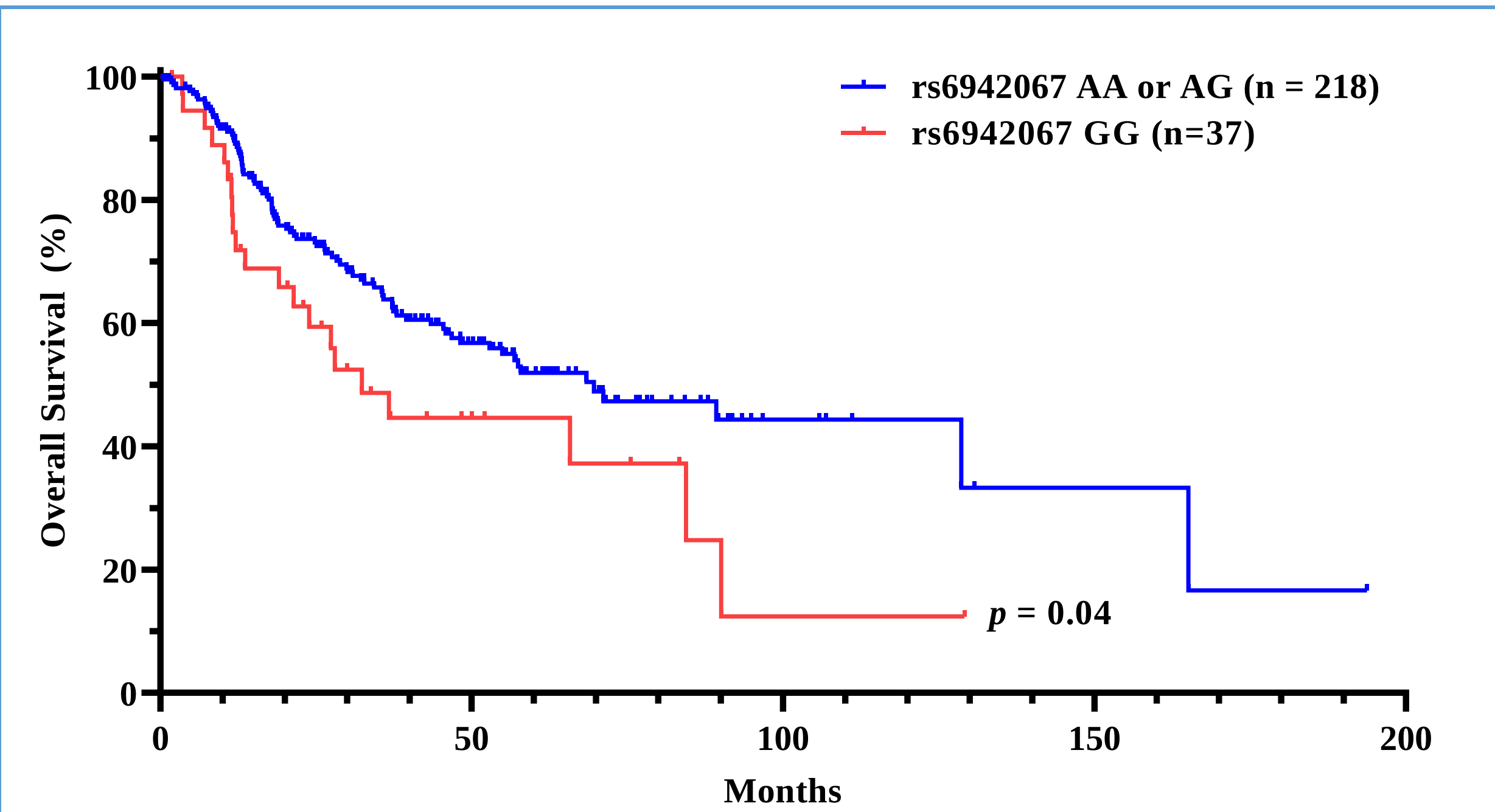


Figure 1. HNSCC overall survival comparison for the combined CHUM and TCGA cohorts using the Kaplan-Meier curve estimation for rs6942067 AA and AG (blue) versus rs6942067 GG (red). P value is calculated using a log-rank test.

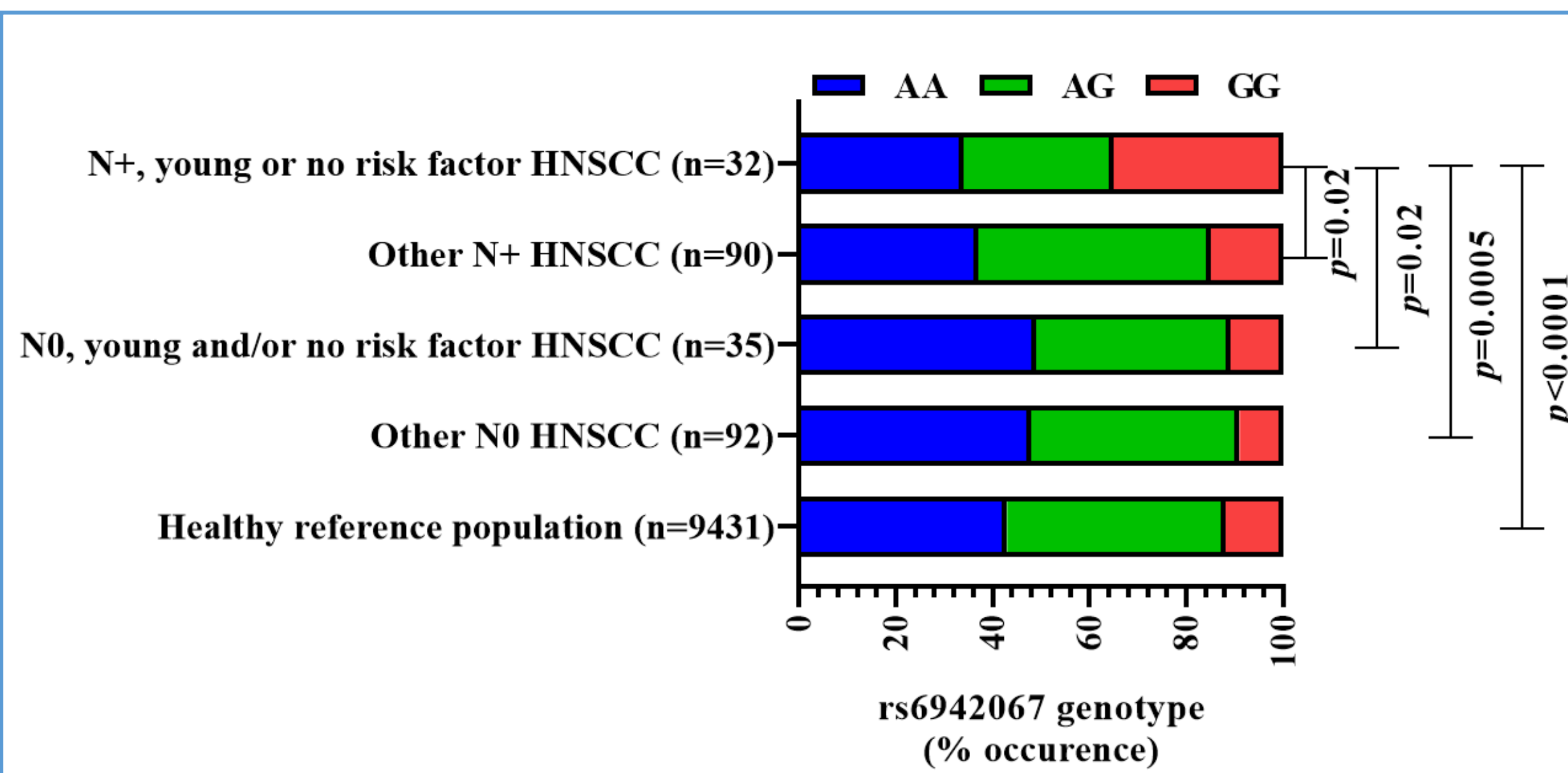


Figure 2. Comparison of the rs6942067 genotype status in the reference healthy cohort (European gnomAD), with HNSCC patients stratified by N0 or N+ nodal status and by over 42 years old and risk factor or under 43 years old and no known risk factor. Risk factor considered were smoking (≥ 10 pack-year) and HPV. Relative risk and p values are for a Pearson's chi-square test and a recessive model of penetrance.

Discussion

HNSCC patients usually present classic risk factors which are tobacco and alcohol use and/or HPV infection. While increased occurrence of HPV has been implicated in young patients developing HNSCC, there has also been a parallel increasing incidence of young patients with HNSCC with no known risk factors.¹ In this study, the association between the rs6942067 GG genotype and overall survival was expected because of the strong association with that genotype and high DCBLD1 expression and the association of high DCBLD1 expression and worse HNSCC overall survival, but the previous study with the 155 TCGA participants alone had failed to have the statistical power to show this association.³ The association between rs6942067 and N+ tumors concurs with previous results showing DCBLD1 association with the integrin signaling pathway^{3,4} and the hypothesis that it increases tumor cell migration. Interestingly, the higher occurrence of rs6942067 GG genotype observed in young or no risk factor HNSCC is specific to the more aggressive N+ tumors. Although rs6942067 GG is associated with young or no risk factor HNSCC, it is not a high risk mutation, such as for instance the BRCA genes for breast and ovarian cancer. There will need to have a large scale study with multiple thousands of genotyped young and no risk factors HNSCC to correctly evaluate which germline SNPs are involved. We believe the accumulation of concordant results^{2,3,4} is now large enough to justify the following two conclusions:

- 1) High DCBLD1 expression is associated to worse overall survival and integrin signaling pathway upregulation in three different cancers including HNSCC.
- 2) There is a SNP region in the DCBLD1 promoter which appears to slightly increase the risk of no risk factor HNSCC and lung cancer.

Acknowledgments

This research was funded by the Dr Azar-Angelil Research Chair in Head and Neck Oncology. We thank all the clinicians and staff involved in the CHUM head and neck cancer biobank since 2010. The results published here are in part based upon data generated by The Cancer Genome Atlas (TCGA), the Broad Institute and the Genotype-Tissue Expression (GTEx) Project. The TCGA is managed by the NCI and NHGRI. Information about TCGA can be found at <http://cancergenome.nih.gov>. The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

References

1. Pickering CR, *et al.* Squamous cell carcinoma of the oral tongue in young non-smokers is genomically similar to tumors in older smokers. Clinical Cancer Research, 2014. doi: 10.1158/1078-0432.CCR-14-0565
2. Lan Q, *et al.* Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nature Genetics, 2012. doi: 10.1038/ng.2456
3. Cardin GB, *et al.* Single nucleotide polymorphism rs6942067 is a risk factor in young and in non-smoking patients with HPV negative head and neck squamous cell carcinoma. Cancers, 2019. doi: 10.3390/cancers12010055
4. Cardin GB, *et al.* DCBLD1 is associated with the integrin signaling pathway and has prognostic value in non-small cell lung and invasive breast carcinoma. Scientific Reports, 2021. doi: 10.1038/s41598-021-92090-6
5. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature, 2015. doi: 10.1038/nature14129
6. Karczewski KJ, *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 2020. doi: 10.1038/s41586-020-2308-7