The potential role of chromosome 3 copy number variation in the ESCC carcinogenesis

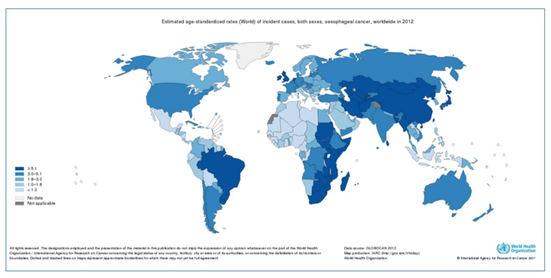
Mariana Severo Ramundo[1], Sheila Coelho Soares Lima[1], Paulo Thiago de Souza Santos[2], Luis Felipe Ribeiro Pinto[1,3]

Authors affiliation: [1]Molecular Carcinogenesis Program – National Cancer Institute, Rio de Janeiro, Brazil. [2]FIOCRUZ, Rio de Janeiro

[3]Universidade Estadual do Rio de Janeiro

INTRODUCTION

- Copy number variations (CNV) at chromosome 3 are genetic alterations commonly observed in many human cancers, including esophageal cancer;
- Esophageal cancer is the eighth most frequent and the sixth cancer with the highest mortality rates in the world, and the most frequent histological type in Brazil is the Esophageal Squamous Cell Carcinoma (ESCC);
- The main risk factors associated with ESCC are tobacco and alcohol consumption;
- The knowledge about the role of chromosome 3 copy number variation in the ESCC carcinogenesis is still limited;



OBJECTIVE

This study aims to verify the incidence of CNV at chromosome 3 in ESCC and evaluate the role of tobacco as a cause of these alterations.

METHODS

- Analysis of CNV data, using bioinformatic tools (cBioPortal and R software), for 1030 genes localized at chromosome 3, of 74 ESCC samples, 511 Lung Squamous Cell Carcinoma (LSCC) samples, 220 Colorectal Cancer (CRC) samples and 273 Glioblastoma samples, obtained from TCGA;
- To evaluate whether these changes are also found in brazilian ESCC patients, we analyzed a comparative genomic hybridization (CGH) previously performed, using 24 samples;
- Some genes localized at 3q, that were differentially expressed in the RNAseq, previously performed by the Molecular Carcinogenesis Program, were selected to perform univariate analysis.
- Chi-square test, Fisher exact test and Univariate analysis using TCGA data were performed in the GraphPad Prism 5.02.

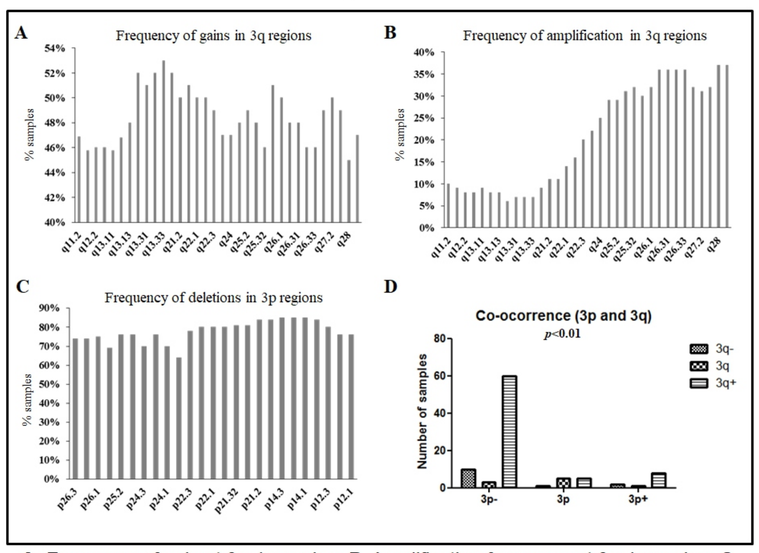
RESULTS

Frequency of copy number gain and deletions at Chromosome 3 in ESCC samples from patients included in the TCGA

We observed high frequencies of copy number gain (CNG) in 3q (89%) and deletions in 3p (85%).

When the data were stratified into gain or amplification and deletions, it was possible to observe that:

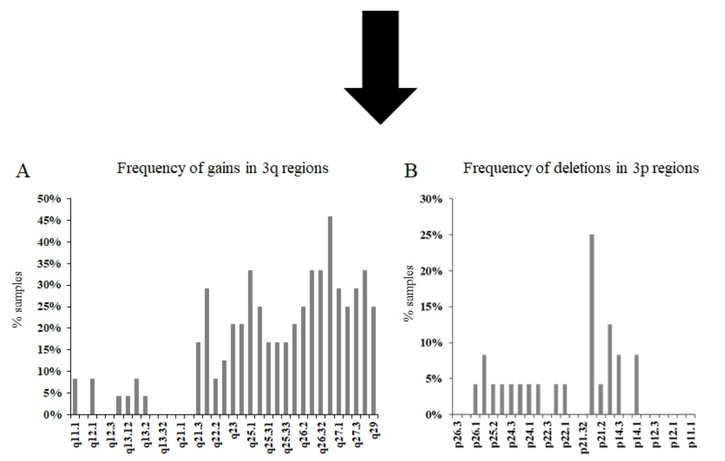
- The gain rates were around 44%, being this gain distributed among all regions of 3q, remaining between 44-53% throughout the long arm (q11. 2-q29) (Figure A);
- The amplification is present in 45% of the ESCC samples, however, the regions presented considerable differences in frequency, being more observed in regions 3q25.31 to 3q29, between 30% and 35% of the samples (Figure B);
- The results obtained for the reduction in the number of copies in the short arm of chromosome 3 show that this event occurs in 63 of the 74 samples (85%), maintaining a frequency of 70-85% in all regions (Figure 4C; 3-p12.1);
- Further, 80% of the patients showed both events occurring simultaneously (isochromosome 3), being this co-occurrence statistically significant according to the chi-square test ($p<0.01$; Figure D).



A. Frequency of gain at 3q, by region. **B.** Amplification frequency at 3q, by region. **C.** Frequency of deletions in 3p, by region. **D.** Co-occurrence of amplification in 3q and deletion in 3p.

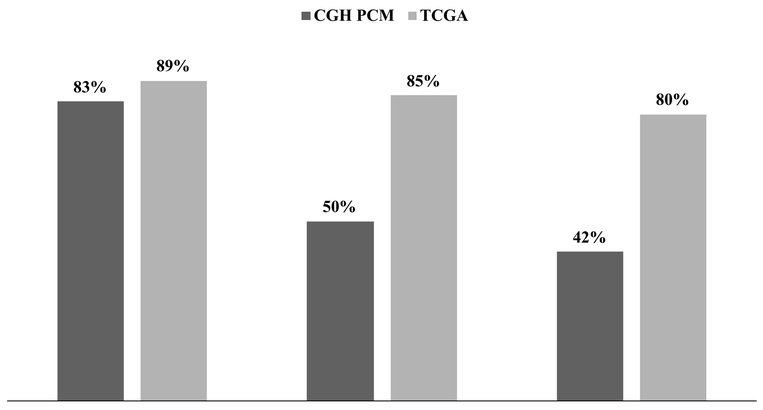
Copy Number Variation at chromosome 3 in ESCC samples from INCA patients

- We observed that 83% of samples presented CNG at 3q, being 3q25.1-q29 the most affected region (Figure A), as observed in the TCGA data.
- The deletion in the short arm of chromosome 3 was observed in 50% of the samples, most frequently in 3p21.31 (Figure B).
- Isochromosome 3 could be observed in 42% of the samples, however, it was not possible to detect the co-occurrence of the two events in this set.



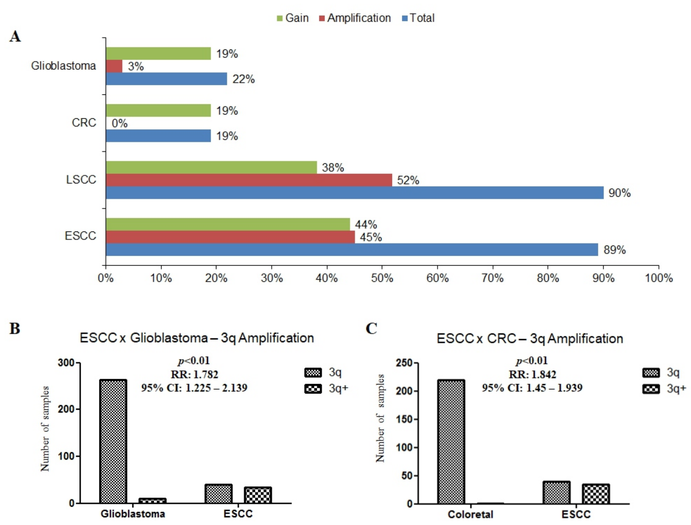
A. Frequency of copy number gain at 3q, by region. **B.** Frequency of copy number gain at 3p, by region.

Comparison between the results obtained from the analysis of the TCGA and CGH results.



The tobacco smoke contribution to the CNG observed at 3q

- We observed that 90% of the ESCC and LSCC samples (tumors clearly associated with tobacco smoking) exhibited CNG at 3q, being high level amplification presented in 50% of the samples. In contrast, no significant amplification rates were found in GBM and CRC (tumors that smoking is not a risk factor) – Figure A.
- Fisher's exact test showed that the difference in amplification rates betweenESCC and non-smoking tumors was statistically significant and the relative risk of ESCC patients to present amplification is 78,2% (Figure B) and 84.2% (Figure C), higher than patients with GBM and CRC, respectively.



A. Copy number variation in tumors with different levels of association with tobacco smoking. Comparison of the presence of 3q amplification between ESCC and glioblastoma (**B**) and ESCC and CRC (**C**).

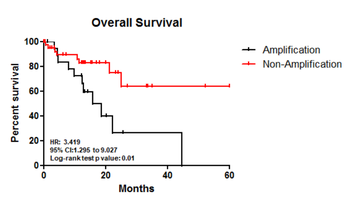
The impact of 3q amplification on the prognosis of ESCC patients

- We selected one representative gene of each region with the highest amplification rates (q25.1-q29) in ESCC for the evaluation of the impact of amplification on patient survival.
- Univariate analysis, using CNV and Clinical Data from TCGA, for the select genes showed that the amplification of all regions tested, except for the q25.32 and q25.33 regions, has a significant impact on overall survival of these patients.

Gene	Region	HR	95% CI	pvalue
PTX3	q25.1	2.383	1.025-5,53	0.04*
MLF1	q25.2	2.383	1.025-5,53	0.04*
MIR15B	q25.31	3.102	1.217 - 7,911	0.01*
ZBBX	q25.32	2.448	0.98-6,07	0.053
EIF5A2	q25.33	2.097	0.89-4,92	0.08
ECT2	q26.1	2.603	1.077-6,291	0.03*
PIK3CA	q26.2	2.432	1.062-5,567	0.03*
PEX5L	q26.31	2.432	1.062-5,567	0.03*
MIR1224	q26.33	2.879	1.212-6,840	0.01*
ETV5	q27.1	3.172	1.349-7,462	0.008*
RFC4	q27.2	3.172	1.349-7,462	0.008*
ATP13A3	q28	2.928	1.276-6,719	0.01*
TFRC	q29	2.789	1.220-6,379	0.01*

HR: Harzard Ratio
CI: Confidence Interval
* p<0.05

Concomitant amplification of the 13 genes analyzed also has a significant impact on prognosis.



CONCLUSION

Our data demonstrated frequent 3p deletion and 3q amplification in ESCC, possibly associated with tobacco consumption and poor survival of ESCC patients. Furthermore, *in vitro* and *in vivo* analysis are needed to validate these data.

REFERENCES

Dai F et al., 2017. The global expression profiling in esophageal squamous cell carcinoma. Genomics. 2017 Jul;109(3-4):241-250. doi: 10.1016/j.ygeno.2017.04.005. Epub 2017 Apr 22. PubMed PMID: 28442363.

INCA. Estimativa 2017: Incidência de Câncer no Brasil. Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. 124p.

Fields, A.P., Justilien, V., Murray, N.R. The chromosome 3q26 OncCassette: A multigenic driver of human cancer. Adv Biol Regul. 60:47-63, 2016.

International Agency for Research on Cancer. (2012) Globocan 2012. [Online]. Available: "http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx"http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspxts new potential targets for further investigation. Support: CAPES, CNPq, FAPERJ, Ministério da Saúde.

Funding source: Ministério da Saúde, FAPERJ, Swiss Brigde, CNPq

Projeto Gráfico: Área de Edição e Produção de Materiais Técnico-Científicos / INCA