Openventio PUBLISHERS

ISSN 2377-1518

CANCER STUDIES AND MOLECULAR MEDICINE

Open Journal

| December 2017 | Volume 3 | Issue 2 |

www.openventio.org

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518

Open Journal

CONTENTS

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518 Open Journal http://dx.doi.org/10.17140/CSMMOJ-3-e005

Editorial

Corresponding author
Alexandros Charalabopoulos, MD
Consultant Oesophago-Gastric Surgeon, Department of Upper Gastrointestinal Surgery
Regional Oesophagogastric Cancer Centre, Broomfield Hospital
Mid Essex Hospital Services NHS Trust, Chelmsford, England, UK
E-mail: acharalabopoulos@yahoo.com

Volume 3 : Issue 2
Article Ref. #: 1000CSMMOJ3e005

Article History
Received: July 25th, 2017
Accepted: July 27th, 2017
Published: July 31st, 2017

Citation
Charalabopoulos A, Liakakos T. Dysadherin: A novel oncogenic molecular biomarker in oesophageal cancer. *Cancer Stud Mol Med Open J*. 2017; 3(2): e5-e6. doi: 10.17140/CSMMOJ-3-e005

Dysadherin: A Novel Oncogenic Molecular Biomarker in Oesophageal Cancer

Alexandros Charalabopoulos, MD, MSc, PhD, MRCS, FRCS[1*]; Theodoros Liakakos, MD, PhD[2]

[1]Department of Upper Gastrointestinal Surgery, Regional Oesophagogastric Cancer Centre, Broomfield Hospital, Mid Essex Hospital Services NHS Trust, Chelmsford, England, UK

[2]Upper Gastrointestinal Unit, 1st Department of Surgery, Laiko Hospital, National and Kapodistrian University of Athens, Athens, Greece

Dysadherin or FXYD5 is a transmembrane glycoprotein, identified for the first time in 2002 by a team of Japanese researchers.[1] Although, it is generally accepted that most of its action is derived from the increased maximal velocity (V_{max}) of the Na^+-K^+-ATPase that promotes, there are several actions of dysadherin that cannot be attributed to its interaction with the Na^+-K^+-ATPase.[2] Dysadherin is a well-described cancer-associated protein and a potential oncogenic molecular target with promising future significances. It has been distinctly shown, in both the cases of *in vivo* and *in vitro* studies, that when expressed in malignant tumours, it signifies augmented tumorigenesis with enhanced metastatic potential and haematogenous spread and is thus linked to poor prognosis.[3,4] On the contrary, studies based on human cancer cell lines have demonstrated that introduction of small interfering RNA (siRNA) against dysadherin leads to downregulation and subsequent membranous underexpression of dysadherin, instigating accordingly decreased collective and individual cell motility with subsequent inhibition of tumour aggressiveness and metastatic potential.[5] Its main interaction in carcinogenesis takes place with E-cadherin (epithelial cadherin); the most important adhesion molecules of the cadherin family. E-cadherin is involved in cancer development in more than 90% of human cancers, as most are carcinomas of epithelial origin. E-cadherin mediates homophilic cell-to-cell adhesion and promotes adherence between neighboring epithelial cells; its decreased cellular expression is hence associated with enhanced disconnection and dispersal of epithelial cells, promoting cell detachment from the primary lesion and consequently exhibiting increased metastatic potential.

Dysadherin and its closely related E-cadherin, has been studied in various cancers and its correlation with cancer aggressiveness and poor prognosis has been collectively and undoubtedly reported.[6-8] While the role of dysadherin in gastric cancer has been studied in detail, in oesophageal cancer its role remains mainly unexplored.[7,9,10] In a clinical study from Shimada et at[11] expression of dysadherin was studied in 117 patients with oesophageal squamous cell carcinoma of all stages. The correlation between immunohistochemically evaluated dysadherin expression with patient clinicopathological data was studied. Oesophageal tumours with expression of dysadherin, displayed worse prognosis in comparison to tumours negative for dysadherin. When dysadherin positivity was combined with absence of E-cadherin expression, a statistically significant detrimental effect was exhibited in that patient group, revealing the worst prognosis of all patient groups. In the group of oesophageal tumours expressing both dysadherin and E-cadherin, prognosis was slightly better, but still poorer compared to the group of tumours that displayed no dysadherin expression. In accordance to what has been shown in relevant studies exploring the role of dysadherin in extraoesophageal human cancers, multivariate analysis revealed that in squamous cell carcinomas of the oesophagus, although expression of E-cadherin is not an independent prognostic factor, expression of dysadherin is ($p<0.05$).[6,11] At present and in contrary to E-cadherin, there is no research investigating the association of dysadherin with oesophageal adenocarcinoma.[12] This is particularly important in the West, where adenocarcinoma subtype comprises more than 80% of oesophageal cancers.

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-e005

Further research is needed in order to explore the role of this promising oncogenic molecular biomarker in oesophageal carcinomas. Dysadherin plays a pivotal role in the carcinogenic process and can therefore become a target of novel oncological agents in an attempt to arrest disease progression and metastasis with obvious implications in prognosis and long-term survival.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Ino Y, Gotoh M, Sakamoto M, Tsukagoshi K, Hirohashi S. Dysadherin, a cancer-associated cell membrane glycoprotein, down-regulates E-cadherin and promotes metastasis. *Proc Natl Acad Sci U S A*. 2002; 99(1): 365-370. doi: 10.1073/pnas.012425299

2. Lubarski I, Karlish SJD, Garty H. Structural and functional interactions between FXYD5 and the Na+-K+-ATPase. *Am J Physiol Renal Physiol*. 2007; 293: F1818-F1826. doi: 10.1152/ajprenal.00367.2007

3. Lee YK, Lee SY, Park JR, et al. Dysadherin expression promotes the motility and survival of human breast cancer cells by AKT activation. *Cancer Sci*. 2012; 103: 1280-1289. doi: 10.1111/j.1349-7006.2012.02302.x

4. Nam JS, Hirohashi S, Wakefield LM. Dysadherin: A new player in cancer progression. *Cancer Lett*. 2007; 255(2): 161-169. doi: 10.1016/j.canlet.2007.02.018

5. Shimamura T, Yasuda J, Ino Y, et al. Dysadherin expression facilitates cell motility and metastatic potential of human pancreatic cancer cells. *Cancer Res*. 2004; 64: 6989-6995. doi: 10.1158/0008-5472.CAN-04-1166

6. Batistatou A, Charalabopoulos AK, Scopa CD, et al. Expression patterns of dysadherin and E-cadherin in lymph node metastases of colorectal carcinoma. *Virchows Arch*. 2006; 448(6): 763-767. doi: 10.1007/s00428-006-0183-8

7. Shimada Y, Yamasaki S, Hashimoto Y, et al. Clinical significance of dysadherin expression in gastric cancer patients. *Clin Cancer Res*. 2004; 10(8): 2818-2823. doi: 10.1158/1078-0432.CCR-0633-03

8. Raman P, Purwin T, Pestell R, Tozeren A. FXYD5 is a marker for poor prognosis and a potential driver for metastasis in ovarian carcinomas. *Cancer Inform*. 2015; 14: 113-119. doi: 10.4137/CIN.S30565

9. Maehata Y, Hirahashi M, Aishima S, et al. Significance of dysadherin and E-cadherin expression in differentiated-type gastric carcinoma with submucosal invasion. *Hum Pathol*. 2011; 42(4): 558-567. doi: 10.1016/j.humpath.2010.08.016

10. Yanagihara K, Tanaka H, Takigahira M, et al. Establishment of two cell lines from human gastric scirrhous carcinoma that possess the potential to metastasize spontaneously in nude mice. *Cancer Sci*. 2004; 95(7): 575-582. doi: 10.1111/j.1349-7006.2004.tb02489.x

11. Shimada Y, Hashimoto Y, Kan T, et al. Prognostic significance of dysadherin expression in esophageal squamous cell carcinoma. *Oncology*. 2004; 67(1): 73-80. doi: 10.1159/000080289

12. Charalabopoulos A, Golias C. E-cadherin expression in Barrett's esophagus and esophageal carcinoma. *Esophagus*. 2014; 11: 153-161. doi: 10.1007/s10388-014-0424-x

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518 Open Journal http://dx.doi.org/10.17140/CSMMOJ-3-117

Research

*Corresponding author
Joana Espiga de Macedo, MD
Department of Medical Oncology
Centro Hospitalar de Entre o Douro e Vouga, Rua Dr. Cândido de Pinho
4520-211 Santa Maria Da Feira
Portugal
Tel. +351-93-6050138
Fax: +351-25-6373867
E-mail: joanamacedo@hotmail.com

Volume 3 : Issue 2
Article Ref. #: 1000CSMMOJ3117

Article History
Received: May 4th, 2017
Accepted: May 16th, 2017
Published: May 16th, 2017

Citation
de Macedo JE. Neurofibromatosis type I and multiple gastrointestinal stromal tumors: A unique identity. *Cancer Stud Mol Med Open J*. 2017; 3(2): 19-20. doi: 10.17140/CSMMOJ-3-117

Neurofibromatosis Type I and Multiple Gastrointestinal Stromal Tumors: A Unique Identity

Joana Espiga de Macedo, MD*

Consultant of Medical Oncologist, Department of Medical Oncology, Centro Hospitalar de Entre, Douro e Vouga, Portugal

INTRODUCTION

Von Recklinghausen's disease (VRD) is a genetic disorder characterized by the growth of multiple noncancerous tumors of nerves and skin (neurofibromas) and areas of hypo or hyperpigmentation of the skin. It is also known as *Neurofibromatosis type I (NF1)* and represents one type of VRD. It is an autosomal dominant disorder with a rate of occurrence of 1 in 3000 in the general population.[1] The cause of *NF1* is a mutation in the *NF1* gene, located at the chromosome 17q11.2, which encodes the tumor suppressor gene, neurofibromin. Loss of neurofibromin function results in activity of Ras oncogene and consequently tumor formation.

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal tumors of the gastrointestinal tract. The prevalence of GISTs in *NF1* varies from 5 to 30%. They are more common in the small intestine (90%) and only 5% occur in the stomach, and they don´t metastasize immediately.[2] The majority are c-kit and PDGFRA wild-type, which means that are not related to oncogenic mutations. In this paper, we report a case of a man with a late diagnosis of neurofibromatosis and multiple GISTs.

CASE REPORT

We report a 75-years single-old man with the diagnosis of VRD - Neurofibromatosis Type I since 2000. He had multiple cutaneous neurofibromas and pigmented macules *(café-au-lait* spots) over the entire body surface. In his past history he was also diagnosed with epilepsy and occasional seizures and benign hypertrophy of the prostate gland. His usual medication was tegretol 400 mg, kepra 1000 mg. He had learning disabilities. He had no history of hypertension, diabetes mellitus, cardiovascular or pulmonary disease or previous gastrointestinal pathology. He had no other relevant past medical history. No known family history of diagnosis of Von Recklinghausen disease. He lives in a social center and always accompanied by a social assistant.

This patient reported to the emergency room in 2011 with abdominal pain and an abdominal computed tomography (CT) scan was performed and solid mass with 57×44 mm in the left hypochondrium was diagnosed. He was submitted to surgery with no complications. Histological analysis revealed a gastrointestinal stromal tumor with 7 cm long, tumor capsule was intact, 3 mitosis/50 HPF and was classified with an intermediate risk, according to the National Institutes of Health (NIH) classification, with 24% risk of recurrence.[3] The patient stayed in surveillance.

Three years later in 2014, a second tumor was diagnosed in the duodenum, jejunum as well as in the mesentery. Histology revealed once again a GISTs. The duodenum fragment revealed 2 nodules with 7 and 15 mm long, 1 mitosis/50 HPF and free margins. Immunohistochemical staining revealed: positive for CD117 (oncogene c-kit), negative for CD34, protein S100 and SMA. It was classified as pT1, very low risk. The jejunum nodule was 8 mm long; no mitosis were identified, free margins. Immunostaining was positive for CD117 and CD34,

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-117

negative for protein S100 and SMA. It was classified as pT1 with very low risk. The mesentery nodule was 9 mm longest axis, spindle cell proliferation like the ones above and with the same immunostaining characteristics. In the three samples no c-kit mutations were identified, because of which resistance to imatinib has been described and the patient stayed in surveillance once again.

A year later, an abdominal CT scan showed a solid mass with 21 mm was identified in the lateral wall of the first portion of the duodenum. The mass shows exocentric growth from the duodenum wall. Cleavage was maintained with other proximal structures. He was once again submitted to surgery with no complications. Histology revealed the following characteristics: the Duodenum nodule was 20×18×11 mm, spindle cell proliferation, mitotic index was inferior to 5 mitosis/50 HPF; proliferative index (Ki67) inferior to 1%. Immunostaining was positive for CD117 (c-kit) and CD34, poorly positive for desmin, negative for protein S100, keratin 7 and SMA. It was classified as pT1, low risk. The 3 jejunum nodules each with 5 mm long showed smooth muscle, spindle cell proliferation. No mitotic figures were identified. Immunostaining was positive for CD117 and no mutations of c-kit were identified once again. The patient until today is in surveillance, with no more recurrences diagnosed up to last image evaluation.

DISCUSSION

GISTs are the most common mesenchymal tumor in the digestive tract, which originate from the cells of Cajal.[4] The most common cases are the sporadic GISTs which harbor c-kit gene mutations in 75%. Another group of patients without c-kit mutations may harbor mutations of the *PDGFRA* gene in 10%.[5]

GISTs associated to *NF1* differ in many ways, and only 5-25% arise in *NF1*. They are typically localized to the small intestine, show multiplicity, strongly positive for KIT by Immunohistochemistry (IHC), but have neither mutations of c-kit or *PDGFRA* gene. These GISTs are commonly accidental findings during other abdominal surgeries. According to the NIH classification, they are frequently low grade, small size and with few or mitotically inactive. The presence of multiple small tumors is not associated with progressive disease. Most patients in the follow-up have a good prognosis.

Consequently, *NF1* cases don´t have any oncogene mutations which could be targeted by a "shut down" molecule as is imatinib mesylate, a tyrosine kinase inhibitor which acts in sporadic cases with oncogene mutations. Therefore, point mutations of c-kit (exon 9,11,13 and 17) and *PDGFRA* (exon 12 or 18) gene may play a small role in the tumorigenesis pathway of *NF1* associated to GIST.[5] In other words, the pathogenetic mechanisms arise via a distinct way.

We report a case of a late known diagnosis of *NF1* associated with multiple GISTs submitted to three surgeries. He was never submitted to imatinib treatment, as previously explained, and with a good prognosis and with good quality of life (QoL).

This is a unique identity of Gist's that clinicians have to be aware of. It appears within the Von Recklinghausen disease, as Neurofibromatosis Type I, and are distinct from the usual GISTs, thus require different care approaches. This case report has some interesting particularities which in this case, spares the patient from molecular target therapy and its toxicities, and on the other hand, spares the country's health system from a heavy economic burden.

CONSENT

The author has taken oral consent from the patient.

REFERENCES

1. Yantiss RK, Rosenberg AE, Sarran L, Besmer P, Antonescu CR. Multiple gastrointestinal stromal tumors in type I neurofibromatosis: A pathologic and molecular study. *Mod Pathol*. 2005; 18(4): 475-484. doi: 10.1038/modpathol.3800334

2. Tomatsu M, Isogaki J, Watanabe T, et al. Multiple gastric gastrointestinal stromal tumors in a patient with neurofibromatosis type 1. *Case Reports in Surgery*. 2016; 2016: 1515202. doi: 10.1155/2016/1515202

3. Joensuu H, Hohenberger P, Corless CL. Gastrointestinal stromal tumour. *The Lancet*; 2013; 382(9896): 973-983, doi: 10.1016/S0140-6736(13)60106-3

4. Hirota S, Isozaki K, Moriyama Y, et al. Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. *Science*. 1998; 279(5350): 577-580. doi: 10.1126/science.279.5350.577

5. Corless CL. Gastrointestinal stromal tumors: What do we know now? *Mod Pathol*. 2014; 27: S1-S16: doi: 10.1038/modpathol.2013.173

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518 Open Journal http://dx.doi.org/10.17140/CSMMOJ-3-118

Letter to the Editor

Corresponding author
Vasudeva G. Kamath, MSc, PhD
Assistant Professor
Department of Biochemistry and Medical Genetics
Touro College of Osteopathic Medicine
60 Prospect Avenue
Middletown, NY 10940, USA
Tel. 845-648-1250
E-mail: vasudeva.kamath@touro.edu

Volume 3 : Issue 2
Article Ref. #: 1000CSMMOJ3118

Article History
Received: May 19th, 2017
Accepted: May 25th, 2017
Published: May 25th, 2017

Citation
Kamath VG. Is it time to start using mitochondrial DNA copy number as an indicator of health and diseases? *Cancer Stud Mol Med Open J.* 2017; 3(2): 21-22. doi: 10.17140/CSMMOJ-3-118

Is it Time to Start Using Mitochondrial DNA Copy Number as an Indicator of Health and Diseases?

Vasudeva G. Kamath, MSc, PhD*

Department of Biochemistry and Medical Genetics, Touro College of Osteopathic Medicine, Middletown, NY 10940, USA

Clinical biochemistry and pathology have contributed too many assays for diagnosis and prognosis of human health and diseases. Bedside biochemistry has revolutionized modern medicine and the invention of new generation biosensors have educated the patients like never before. Although, there are many gadgets which are available for monitoring the blood glucose in the patients with metabolic syndromes, hardly any assays can accurately predict and aid in early detection of metabolic syndrome. By the time a person is diagnosed and treated for metabolic syndrome, molecular and pathological damages further push the patient towards the metabolic syndrome. Early detection and early treatments are necessary to further revolutionize the field of Medicine.

Metabolic syndrome is not a completely solved riddle, there are several factors which contribute towards the genesis and progression of the disease. Energy metabolism is in the center of metabolic syndrome, and mitochondria are one such organelle where majority of the energy metabolism takes place. Given that it is at the center of energy metabolism, most of the proteins required for this process are supplied by the nuclear gene products. Additionally, the mitochondria with its genome is responsible for 13 of the key electron transport complex proteins. Mutations in both nuclear and mitochondrial genes result in Mitochondrial Encephalomyopathy Lactic Acidosis and Stroke-like Episodes (MELAS), Myoclonic Epilepsy and Ragged-Red Fiber Disease (MERF), Mitochondrial Neurogastrointestinal Encephalopathy (MNGIE) and Leigh syndrome.

The mitochondrial genome is a simple closed circular molecule that contains 16,569 DNA base pairs compared to the complex nuclear genome. But on the other hand, the nuclear genome has specified amount i.e., two copies at the least of every gene and the quality of the genome is well maintained by "well-oiled" DNA repair machinery. In mitochondrial genome contains numerous copies of the genes. Apart from this the number of mitochondria in each tissue also varies. This imbalance in quantity of proteins generated from mitochondria and nucleus for electron transport complex is a testimony of their importance in energy metabolism. Mitochondrial DNA (mtDNA) quantity and the quality is an important aspect of the energy metabolism. The factors attributing towards determining the number of mitochondria and mitochondrial DNA copy number is yet unclear. Nonetheless, recent investigations suggest a negative correlation between mitochondrial DNA copy number and disease process and progression. A recent article Reznik et al[1] clearly demonstrated the correlation between mtDNA copy number to the expression of mitochondrially-localized metabolic pathways, suggesting that mtDNA copy number variation reflects gross changes in mitochondrial metabolic activity. As we discussed the significance of the mitochondrial DNA copy number, it is also important to note the quality of the DNA, it is observed that over a period of time the bad copies of the mitochondrial DNA accumulate within tissue mitochondria as the mitochondrial DNA repair mechanism is not as effective as the nuclear DNA repair mechanisms and the chances of mutation at the end of every replication cycle is higher. With this delicate balance of the quality and quantity of the mtDNA determining the efficacy of the energy metabolism, there is a high possibility that these changes can be used as indicators of the metabolic problems to occur in near future. It is up to

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-118

the scientific communities to decide whether it is time for us to seriously investigate the numerical relation of mitochondrial DNA copy number as a health indicator. Appropriate detection of mtDNA mutations and quantification of mtDNA can be a very important means in the resources of the modern day physicians.

REFERENCE

1. Reznik E, Miller ML, Şenbabaoğlu Y, et al. Mitochondrial DNA copy number variation across human cancers. *eLife*. 2016; 5: e10769. doi: 10.7554/eLife.10769

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518 Open Journal http://dx.doi.org/10.17140/CSMMOJ-3-119

Mini Review

Corresponding author
Joana Espiga de Macedo, MD
Department of Medical Oncology
Centro Hospitalar de Entre o Douro e Vouga, Rua Dr. Cândido de Pinho
4520-211 Santa Maria Da Feira
Portugal
Tel. +351-93-6050138
Fax: +351-25-6373867
E-mail: joanamacedo@hotmail.com

Volume 3 : Issue 2
Article Ref. #: 1000CSMMOJ3119

Article History
Received: May 31st, 2017
Accepted: June 13th, 2017
Published: June 21st, 2017

Citation
de Macedo JE. How target therapy can induce cardiotoxicity: The onco-cardiologist point of view. *Cancer Stud Mol Med Open J*. 2017; 3(2): 23-26. doi: 10.17140/CSMMOJ-3-119

How Target Therapy can Induce Cardiotoxicity: The Onco-Cardiologist Point of View

Joana Espiga de Macedo, MD*

Consultant of Medical Oncologist, Department of Medical Oncology, Centro Hospitalar de Entre, Douro e Vouga, Portugal

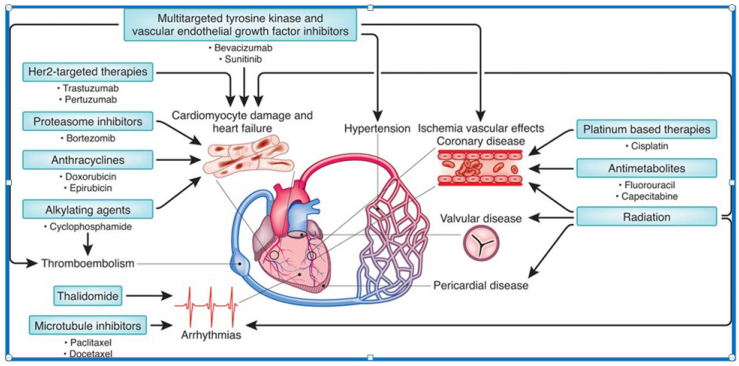
INTRODUCTION

The life expectancy of an oncological patient has been increased significantly decades due to the evolution of cancer therapies. The aim of this paper is to review how these new target molecules, which have a direct impact on the cancer cell as opposed to the classical chemotherapy, known to all of us for a long time. How do they act specifically? What are their cardiotoxic effects at a short and long time period? How can they be monitored? How can they be prevented and treated? And above all, how can we prevent an oncological patient to become a cardiovascular one later on, due to induced cardiotoxicity caused by oncological therapies. We have to treat our patients as a whole and not only, target therapy to one of the pieces of the puzzle.

DEFINITION

Cardiotoxicity can be defined as all damage caused directly or indirectly to the cardiomyocyte. This is a broad term which affects the cardiovascular system in many ways. The classical definition refers a reduction in left ventricular ejection fraction and the development of heart failure.[1] However, they can be divided in ten categories: myocardial dysfunction, heart failure, coronary artery disease, arrhythmias, arterial hypertension, thromboembolic disease, peripheral vascular disease, stroke, pulmonary hypertension and pericardial complications (Figure 1).

Figure 1: Cardiovascular Effects of Target Therapies.[1]



First of all, when a patient presents himself with a certain lifestyle risk factors (smoking, alcohol, obesity), demographic and other cardiovascular risk factors (age, arterial hypertension or diabetes mellitus) then it is needed to be checked whether the factors are known or unknown, treated and controlled or not. Depending on past history, he may also have cardiovascular disease such as heart failure or myocardial infarct. Or he might even, have already have had, a primary cancer diagnose and submitted to chemotherapy or radiotherapy previously.

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-119

As mentioned above, the three most important cardiotoxicities are myocardial dysfunction, heart failure and coronary artery disease. The mechanisms of action of these target therapies and how they act and their statistical incidence can be analysed in Tables 1 and 2. However, as already mentioned, in Table 3, we can evaluate the population baseline cardiovascular risk factors, which may be increased by specific target therapies as will be explained further on.

HUMAN EPIDERMAL RECEPTOR 2 (HER2)

Human epidermal receptor 2 (HER2), is a member of the ErbB protein family, with an extracellular domain and an intracellular domain. It is overexpressed for example, in patients with breast cancer. It is an extremely important control centre of unregulated cell growth and proliferation. Blocking this protein by target therapies, leads to inhibitionor downregulation of the

Table 1: Heart Failure and Myocardial Dysfunction.[3]

Agent	Incidence (%)	Agent	Incidence (%)
Doxorubicin	2-48	Pertuzumab	<1.5
Epirubicin	1-3.3	Lapatinib	<1
Lipsomal A	2	Sunitinib	2.7-19
Cyclophosphamide	7-28	Sorafenib	4-8
Ifosfamide	17	Pazopanib	7-11
Docetaxel	2-13	Imatinib	<3
Paclitaxel	<1	Everolimus	<1
Bevacizumab	1.6-4	Temsirolimus	<1
Trastuzumab	1.7-20		

Table 2: Coronary Artery Disease.[3]

Therapies	Possible mechanisms	Incidence
Fluoropyrimidines (5-FU, capecitabine)	Endothelial injury and vasospasm	Myocardial ischemia: 18% Silent myocardial ischemia: 50%
Platinum salts	Procoagulant status; arterial thrombosis	Thrombosis: 2% In long-term survivors (8% 20 y testicular cancer)
VEGF inhibitors	Procoagulant status; arterial thrombosis; endothelial injury	AT with bevacizumab 3.8%
Radiotherapy	Thrombosis; endothelial injury, plaque rupture 30 y	In long-term survivors (13% 30 y Hodgkin lymphoma)

Table 3: Factors Associated with Risk of Cardiotoxicity Following Therapeutics with Anti-HER2 and VEGF Inhibitors.[3]

Agent	Risk factors
Anti-HER2 compounds	
Antibodies - Trastuzumab - Pertuzumab - T-DM1 Tyrosine Kinase inhibitors - Lapatinib	• Previous or concomitant anthracycline treatment (short time between anthracycline and anti-HER2 treatment) • Age (>65 years) • High BMI > kg/mg^2 • Previous LV dysfunction • Arterial hypertension • Previous radiation therapy
VEGF inhibitors	
Antibodies - Bevacizumab - Ramucirumab	• Pre-existing HF, significant CAD or left side VHD (e.g. mitral regurgitation), chronic ischaemic cardiomyopathy • Previous anthracycline
Tyrosine Kinase inhibitors - Sunitinib - Pazopanib - Axitinib - Neratinib - Afatinib - Sorafenib - Dasatinib	• Arterial hypertension • Pre-existing cardic disease

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-119

phosphoinositide 3-kinase (P13-Kinase/AKT) and the mitogen-activated protein kinase (MAPK/ERK1/2) pathways (Figure 2).[2]

This protein can be inhibited in two places: in the extracellular domain by trastuzumab, a monoclonal antibody, or in the intracellular domain by lapatinib, an inhibitor of the receptor of the tyrosine kinase domain. This leads to downregulation of this pathway. Consequently, leading to impaired myocardial response to stress, leading to hypertension.

By blocking the extracellular domain this inhibitor starts competing with neuregulin-1 (NRG-1) which is released by the endothelial cells. It has an autocrine function, stimulating angiogenesis and vascular homeostasis, and also a paracrine effect, acting on the myocyte leading to growth, survival and maintenance of the sarcomere organisation by stimulating the erbB2 pathway. Together these effects protect the cardiac structure and function under normal conditions. It has a cardio-protective effect. However if this pathway is blocked, it induces apoptosis leading to heart failure. On the other hand, if tyrosine signalling is maintained, with constant stimulation by neuregulin-1, preservation of the myocyte haemostasis is achieved.[2]

VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF) AND PLATELET DERIVED GROWTH FACTOR (PDGF)

Inhibition of the VEGF may occur in different places along the signalling pathway. Directly, on the ligand for example as dose bevacizumaba monoclonal antibody. It blocks its connection to the VEGF receptor (VEGFR). Secondly, by blocking the receptor directly as does Sunitinib. It doesn't allow an adaptative modulation of vascular structure and function, leading to hypertension and left ventricle disfunction. Sunitinib blocks the AMP-activated protein kinase (AMPK) pathway disrupting the regulation of the endothelial metabolism (Figure 3).

PDGF can also be blocked in the intracellular domain of the PDGF receptor (PDGFR) by Sorafenib, a tyrosine kinase inhibitor, which causes disruption of the cardiomyocyte haemostasis leading to hypertension.

Figure 2: HER2/ERBB2 – Target Therapies.[1]

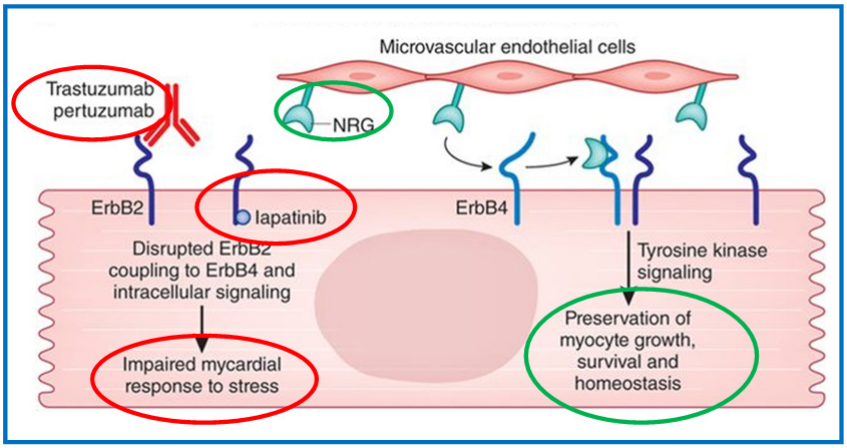
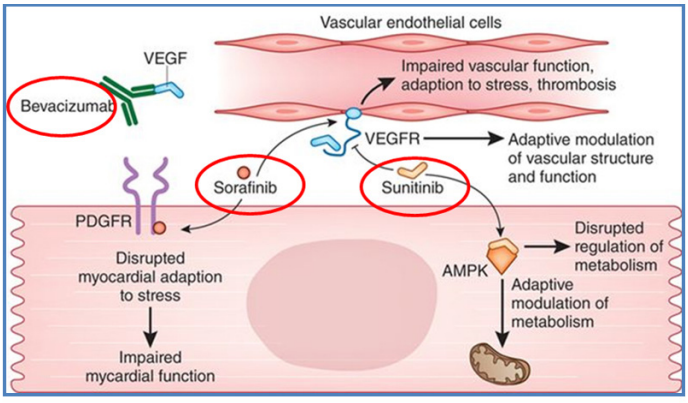


Figure 3: VEGF – Target Therapies.[1]



ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-119

HOW CAN THESE CARDIOVASCULAR TOXICITIES BE MONITORED, PREVENTED AND TREATED?

First of all, a rigorous screening can be done to exclude patients at unacceptable high cardiovascular risk. This is achieved with a complete patient's history, physical examination, alerted to signs and symptoms of heart failure. Evaluate cardiac co-morbidities and if it was well medicated and controlled. Also patients should be questioned regarding previous cancer history and therapies performed, if any.

Secondly, it is necessary to adapt each treatment to the patient we have in front, and induce less toxicity as possible in all areas. Regular screening can be done by echocardiography with Doppler imaging to evaluate the left ventricle ejection fraction. It is still the standard of care, which best characterizes the cardiac morphology and function, detects cardiac systolic and diastolic dysfunction, although it's operator dependent and time consuming. On the other hand it doesn't expose the patient to radiation as multi-gated acquisition (MUGA). Cardiac magnetic resonance imaging (MRI) it's not widely available, causes claustrophobia and is expensive comparing theechocardiography with Doppler imaging.[3]

Cardiac biomarkers, such as brain natriuretic peptide (BNP) and troponin I (TNT) have been wildly used in many trials, but have never been validated as true biomarkers.[3] However, it has been demonstrated that TNT early increase and detection, is predictive of early cardiac dysfunction in oncological patients. Consequently, early treatment with inhibition of the angiotensin converting enzyme, may avoid late dysfunction of the left ventricle. When early cardiomyopathy cell death is detected both present elevated parameters.[4,5]

Treatment has long been more beneficial and effective, when it is started as early as possible with cardioprotective drugs. The most common used are β-blockers and inhibition of the angiotensin converting enzyme pathway (ACE). Also correction and control of other co-morbidities are also extremely important to prevent further damage to the cardiomyocyte. Adopt healthy lifestyle is essential in all medical areas, such as: a healthy diet, not smoking controlling body weight and above all regular aerobic exercise.

FUTURE ORIENTATIONS

Clinicians have to be more aware of heart disease in oncological patients and survivors. The aim of this paper was to widen the view of how these new targeting molecules may act on the cardiovascular system, what's the best way to diagnose and prevent and also how to treat. More than ever, efforts between cardiologists who best know how to deal with cardiovascular disease, in its differentclinical ways, the oncologist who best deals with the targeting therapies and its consequences and also, the family physician who has the best knowledge of the patient's clinical history and co-morbidities. It depends on the fusion of these essential practitioners that they can best treat the patient as whole, with the best quality of life (QoL). Nevertheless, larger prospective studies are needed to validate appropriate methods of monetarization, prevention and treatment of cardiovascular effects of cancer target therapies.

REFERENCES

1. Lenneman CG, Sawyer DB. Cardio-oncology an update on cardiotoxicity of cancer-related treatment. *Circ Res*. 2016; 118: 1008-1020. doi: 10.1161/CIRCRESAHA.115.303633

2. Hahn VS, Lenihan L, Ky B. Cancer therapy–induced cardiotoxicity: Basic mechanisms and potential cardioprotective therapies virginia shalkey. *J Am Heart Assoc*. 2014; 3(2): e000665. doi: 10.1161/JAHA.113.000665

3. Zamorano JL, Lancellotti P, Muñoz DR, et al. 2016 ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines The Task Force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC). *Eur Heart J*. 2016; 37: 2768-2801. doi: 10.1093/eurheartj/ehw211

4. Cardinale D, Colombo A, Lamantia G, et al. Anthracycline-induced cardiomyopathy: Clinical relevance and response to pharmacologic therapy. *J Am Coll Cardiol*. 2010; 55(3): 213-220. doi: 10.1016/j.jacc.2009.03.095

5. Cardinale D, Colombo A, Sandri MT, et al. Prevention of high-dose chemotherapy–induced cardiotoxicity in high-risk patients by angiotensin-converting enzyme inhibition. *Circulation*. 2006; 114: 2474-2481. doi: 10.1161/CIRCULATIONAHA.106.635144

CANCER STUDIES AND MOLECULAR MEDICINE

Openventio PUBLISHERS

ISSN 2377-1518 Open Journal http://dx.doi.org/10.17140/CSMMOJ-3-120

Research

*Corresponding author
Chi Wang, PhD
Markey Cancer Center
800 Rose St., CC433
Lexington, KY 40536, USA
Tel. 859-323-2045
E-mail: chi.wang@uky.edu

Volume 3 : Issue 2
Article Ref. #: 1000CSMMOJ3120

Article History
Received: September 2nd, 2017
Accepted: September 19th, 2017
Published: September 19th, 2017

Citation
Huang Z, Chen L, Wang C. Classifying lung adenocarcinoma and squamous cell carcinoma using RNA-seq data. *Cancer Stud Mol Med Open J*. 2017; 3(2): 27-31. doi: 10.17140/CSMMOJ-3-120

Classifying Lung Adenocarcinoma and Squamous Cell Carcinoma using RNA-Seq Data

Zhengyan Huang, MPA[1]; Li Chen, PhD[1,2]; Chi Wang, PhD[1,2*]

[1]Department of Biostatistics, University of Kentucky, Lexington, KY 40536, USA
[2]Markey Cancer Center, University of Kentucky, Lexington, KY 40536, USA

ABSTRACT

Background: Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are two primary subtypes of non-small cell lung carcinoma (NSCLC). Currently, the most widely used method to discriminate between LUAD and LUSC is hematoxylin-eosin (HE) staining. However, this method is not always able to precisely diagnose LUAD or LUSC. More accurate diagnostic approaches are highly desired.
Methods: We propose to use gene expression profile to discriminate a patient's NSCLC subtype. We leveraged RNA-Seq data from The Cancer Genome Atlas (TCGA) and randomly split the data into training and testing subsets. To construct classifiers based on the training data, we considered three methods: logistic regression on principal components (PCR), logistic regression with LASSO shrinkage (LASSO), and kth nearest neighbors (KNN). Performances of the classifiers were evaluated and compared based on the testing data.
Results: All gene expression-based classifiers show high accuracy in discriminating between LUSC and LUAD. The classifier obtained by LASSO has the smallest overall misclassification rate of 3.42% (95% CI: 3.25%-3.60%) when using 0.5 as the cutoff value for the predicted probability of belonging to a subtype, followed by classifiers obtained by PCR (4.36%, 95% CI: 4.23%-4.49%) and KNN (8.70%, 95% CI: 8.57%-8.83%). The LASSO classifier also has the highest average area under the receiver operating characteristic curve (AUC) value of 0.993, compared to PCR (0.987) and KNN (0.965).
Conclusions: Our results suggest that mRNA expressions are highly informative for classifying NSCLC subtypes and may potentially be used to assist clinical diagnosis.

KEY WORDS: LUAD; LUSC; Principal Components; LASSO; Kth Nearest Neighbors.

ABBREVIATIONS: NSCLC: Non-Small Cell Lung Carcinoma; LUAD: Lung adenocarcinoma; LUSC: Lung squamous cell carcinoma; RNA-Seq: RNA sequencing; PCR: Principal Components Regression; LASSO: Logistic regression with lasso shrinkage; KNN: Kth nearest neighbors; ROC: Receiver Operating Characteristic; AUC: Area under the ROC curve.

INTRODUCTION

Lung cancer has the second highest estimated new case rates and the highest estimated death rates for both males and females. According to American Cancer Society (ACS), the estimated number of new cases in 2015 are 115,610 men and 105,590 women which account for about 14% and 13% of all new cancer cases for males and females respectively. The estimated deaths are 86,380 men and 71,660 women which account for about 28% and 26% of all deaths associated with cancer for males and females respectively.[1] Lung cancer can be classified as small cell lung carcinoma and non-small cell lung carcinoma (NSCLC). NSCLC weighs more than 80% of all lung cancer. NSCLC can be sub-classified into Lung squamous cell carcinoma (LUSC), Lung adenocarcinoma (LUAD), and large cell carcinoma.[2] Approximately 20% of all lung cancers are LUSC; it has the worst prognosis. About 40% of all NSCLC cancers are LUAD.[2]

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-120

Despite the differences in prognosis, subtypes of NSCLC have been treated by similar strategies. The treatment effects vary in LUSC and LUAD patients.[3] With the rapid development of the targeted therapies for NSCLC, more efficient treatments are available for both NSCLC subtypes. However, to choose treatments, especially targeted therapies and combination of interventions, we need more accurate sub-typing between them.[3]

Currently, the most widely used method to distinguish between LUAD and LUSC is hematoxylin-eosin (HE) staining of the tumor tissue sections observed under a light microscope. However, due to the drawbacks of HE stainning, such as unclear structures in tumors and small biopsies with a limited number of tumor cells, it is difficult to make a precise diagnosis between LUAD and LUSC.[4] Since the molecular profiling is different between LUAD and LUSC, Immunohistochemical (IHC) staining can help to diagnose between LUAD and LUSC. However, it needs knowledge of reliable IHC markers.[4] Yu et al designed a fully automated informatics pipeline to extract quantitative image features and build classifiers to distinguish survival outcomes for lung cancer. They applied the classifiers to distinguish between LUAD and LUSC and obtained 0.75 as the highest area under the curve (AUC) value.[5]

High-throughput data obtained from microarrays and RNA sequencing can be used to identify appropriate biomarkers for IHC staining.[4] However, limited biomarkers are applied to IHC staining. By using all available high-throughput data, we may obtain better diagnostic outcomes. In our study, we applied three methods directly to publicly available RNA-Seq data released by the cancer genome atlas (TCGA) from National Cancer Institute (NCI).

METHODS

Data Source

Normalized level 3 RNA sequencing data of tumor samples were obtained from the R package RTCGA RNA seq.[6] The data include 576 LUAD and 552 LUSC cases. The outcome to predict is LUAD *versus* LUSC, and the predictors are 20,259 gene expressions quantified using RNA-Seq by expectation maximization (RSEM) values.[7,8] The original data includes 20531 genes; however, 272 have all zero values and thus were excluded from the analyses. In addition, we applied log transformation to achieve an approximately normal distribution for the data.

Overview of Classifier Construction and Evaluation

We applied the following three methods to predict lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC): principal components regression (PCR), logistic regression with LASSO shrinkage (LASSO), and kth nearest neighbors (KNN). We randomly split the data into training and testing at a ratio of 3:1, where the training data were used to construct a classifier and the testing data were used to evaluate its performance. Details on the three methods are described in the next three subsections.

After a classification model was constructed, we applied it to the testing data to obtain a predicted probability for each sample to be a certain subtype. We applied 0.5 as the cutoff point of the predicted probability to classify samples into the two subtypes. A set of statistics including overall misclassification rates, misclassification rates for LUAD and LUSC, and AUC values were calculated. In addition, we believed that genes used in the selected models were informative. Therefore, we tracked those genes from our selected models for PCR and LASSO. Due to method difference, KNN doesn't return information on genes. For PCR, we selected 10 genes (5 with the largest positive and 5 with the largest negative loading values) from each of the first two principal components. For LASSO, we recorded all genes included in the selected model. The frequency of those genes will be reported in the Results section.

For each method, we repeated the process 500 times using 500 randomly generated seed numbers. Within each replication, we applied function "set.seed" with fixed seed number. In this way, we got the same split datasets for all three methods to make the final results more comparable. All data processing and analyses were performed in R (version 3.3.2).

PCR: Logistic Regression on Principal Components

The "prcomp" function from R Stats package was used to obtain principal components from training data. We fitted the training data with a different number of principal components and recorded the AIC value for each model. The model with the smallest AIC value was identified as the best model.

LASSO: Logistic Regression with Lasso Shrinkage

The R package "glmnet" is used to fit the LASSO model. To select the shrinkage parameter lambda in the LASSO model, we further split the training dataset at a ratio of 2:1. We fit 100 models with different lambda values using the first portion of the data and calculate the overall misclassification rate of each model using the second portion of the data. The cutoff point to calculate the overall misclassification rate was set as 0.5. The best lambda value was chosen as the model that yields the smallest overall misclassification rate.

KNN: Kth Nearest Neighbors

The R package "class" was used to build a KNN-based classifier. We applied the same strategy as we did for LASSO to select the best number of nearest neighbors (K) in the KNN model. Specifically, the training data were split at a ratio of 2:1, where the first portion was used to fit 20 KNN models with K ranging from 1 to 20. Then the 20 models were applied to the second portion of the data to calculate the overall misclassification rate (cutoff point=0.5) for each model. The best model (best number for K) was selected as the one that achieved the smallest overall misclassification rate.

CANCER STUDIES AND MOLECULAR MEDICINE
Open Journal
Openventio PUBLISHERS
ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-120

RESULTS

Table 1 shows the mean value and 95% confidence intervals for overall misclassification rate and misclassification rates for LUAD and LUSC. As we mentioned in the Methods section, all misclassification rates were obtained by using 0.5 as the cut-off point. LASSO has the smallest overall misclassification rate 3.42% (95% CI: 3.25%-3.60%), followed by PCR 5.64% (95% CI: 5.51%-5.76%) and KNN 8.70% (95% CI: 8.57%-8.83%). Compared to PCR and KNN, LASSO also has the smallest misclassification rate for both LUAD (2.55%) and LUSC (4.36%).

Receiver operating characteristic (ROC) curves with area under the curve (AUC) values for all three methods are displayed in Figure 1. LASSO has the highest average AUC value (0.993), compared to PCR (0.987) and KNN (0.965).

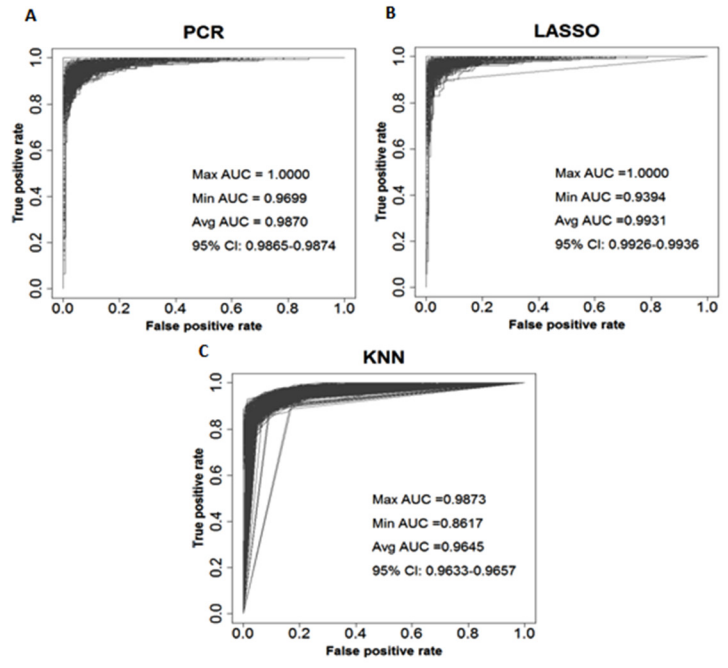
Figure 2 shows the classification result from one replication. The test dataset for this replication included 145 LUAD and 137 LUSC cases. Among them, 7LUAD cases (red triangle) were classified into LUSC and 12LUSC cases (green plus) were classified into LUAD. The AUC value was 0.974 if only the first two principal components (PCs) were used. To explore whether increasing number of PCs in the model would help to reduce the misclassification rates and increase the AUC value, we considered a model including the first thirteen PCs. We got the same overall misclassification rate and same misclassification rates for LUAD and LUSC from the selected model using the same cutoff point. The AUC value was 0.985, slightly higher than the AUC value from using only two PCs. Since by only using the first two PCs, LUAD and LUSC were well separated, we only recorded genes with large effects from the first two PCs.

Frequencies of genes identified by the first two principal components from PCR and by the selected model from LASSO are shown in Figure 3. There are 12 genes identified by the first principal component, 16 genes identified by the second one, and 408 genes identified by LASSO. We only display 20 genes with the highest frequency from LASSO. Among all identified genes, 21 genes have a 90% or higher frequency rate. Three genes, *CALML3*, *DSG3*, and *KRT5* were identified by both PCR and LASSO.

Table 1: Summary Statistics for PCR, LASSO, and KNN. Mean Misclassification Rates and 95% Confidence Intervals are shown in Percentages. All Statistics Presented in the Table are based on 500 Replications.

	PCR	LASSO	KNN
Misclassification Rate (%)	**Mean (95% CI)**	**Mean (95% CI)**	**Mean (95% CI)**
LUAD	4.36 (4.23-4.49)	2.55 (2.41-2.69)	7.25 (7.01-7.49)
LUSC	7.02 (6.84-7.20)	4.36 (4.13-4.59)	10.16 (9.93-10.39)
Overall	5.64 (5.51-5.76)	3.42 (3.25-3.60)	8.70 (8.57-8.83)

Figure 1: ROC Curves for PCR (plot A), LASSO (plot B), and KNN (plot C). Max AUC: Maximum AUC Value; Min AUC: Minimum AUC Value; Avg AUC: Mean AUC Value; 95% CI: 95% Confidence Interval for AUC Value. All Statistics Presented in the Figures are based on 500 Replications.



ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-120

Figure 2: Classification Result from the First Two Principal Components. FLUAD: LUSC Cases Classified as LUAD; FLUSC: LUAD Cases Classified as LUSC; TLUAD: LUAD Cases Classified as LUAD; TLUSC: LUSC Cases Classified as LUSC. This Figure Shows the Result of one Replication. For this Replication, the First Principal Component Explains 19.1% of the Variance and the Second One Explains 7.1% of the Variance.

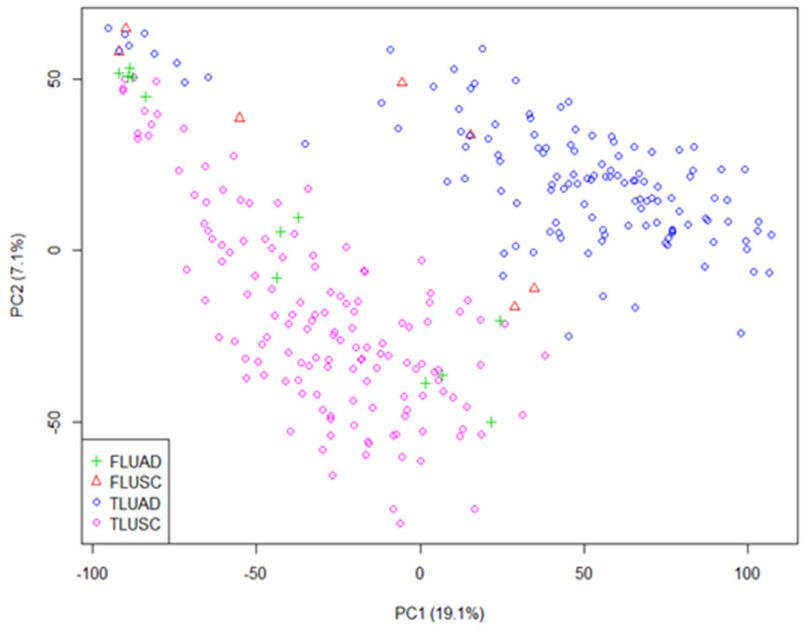
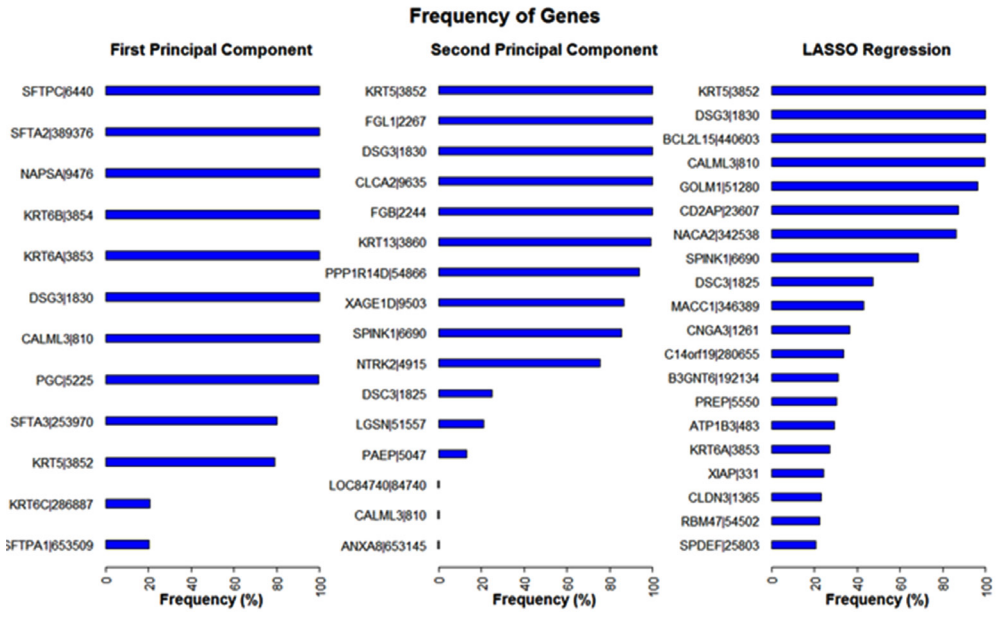


Figure 3: Frequency of Genes Identified by PCR and LASSO. Frequencies are shown in Percentages. We use Frequency Rates to Describe. Results are based on 500 Replications. A 100% Frequency Rate Means the Gene was Recorded in all 500 Replications.



DISCUSSION

In this paper, we considered three different statistical methods to classify LUSC and LUAD patients based on their gene expression profiles. All the three methods have a low overall misclassification rate of less than 9% and high AUC of greater than 0.96.

Our analyses demonstrate that gene expression data can accurately discriminate LUSC and LUAD samples, which does not specifically depend on the choice of statistical meth-

ISSN 2377-1518 http://dx.doi.org/10.17140/CSMMOJ-3-120

od. Therefore, gene expression profile may potentially be used in clinicly or in clinics to enhance the diagnosis of LUSC and LUAD.

One issue we had was to standardize the data. Some genes have many zero values. After we split the whole data into subsets, the subsets may contain all zeros for certain genes, and those genes could be different in training and testing data, and they could also be different for each replication. To confirm the results that we got for all three methods, we removed all genes that have constant zeros for all subsets. This further reduced the number of genes included in the analysis. Then we applied three methods again on the standardized data. No significant changes were seen after the standardization. Therefore, we chose to process the data without standardization.

The model fitting based on the KNN method was sometimes unstable in our analysis. It might have been caused by the high dimensions of the data. Different algorithms on distance metric could be used to improve the classification. Also, we only tested a number of neighbors (K) from 1 to 20. Larger K may improve the diagnostic performance. In contrast, PCR and LASSO were very stable. As for computational time, both PCR and LASSO were considerably less than it was for KNN. On average, one replication took less than 1 minute and less than 10 seconds for PCR and LASSO, respectively. One KNN replication for K from 1 to 20 took more than 10 minutes. When we increased K, the computation time increased significantly. Since PCR and LASSO performed better on sub-typing LUAD and LUSC and had shorter computation time, we consider PCR and LASSO to be better methods than KNN.

Another advantage of PCR and LASSO is that we are able to get information on the contribution of each gene. Among those genes identified by our methods, *SFTA3*, *DSG3*, *DSC3*, and *CALML3* were found to be useful to distinguish LUAD and LUSC from each other using an earlier version of TCGA data.[4] In addition, *DSG3*, *NAPSA*, *KRT5*, *KRT6A*, *KRT6B*, and *SFTA2* were identified as potential biomarkers for distinguishing between the two subtypes using different data sources.[3,9,10] Unlike our study, those studies applied various differently expressed gene screening methods to identify potentially informative genes. Although, our main aim is to discriminate LUAD *versus* LUSC, we successfully identified many genes which were found in other studies. We believe that those high dimension reduction methods can help to discover potential biomarkers to distinguish between the two subtypes. Those methods also can be applied to other disease types known to have different molecular profiling.

ACKNOWLEDGEMENTS

The authors wish to thank Richard Charnigo for helpful comments. This research was supported by the Biostatistics and Bioinformatics Shared Resource Facility of the University of Kentucky Markey Cancer Center (P30CA177558).

CONFLICTS OF INTEREST

The authors have no conflicts to declare.

REFERENCES

1. American Cancer Society. Cancer Facts & Figures 2015. 2015. Atlanta, GA, USA. Web site. http://oralcancerfoundation.org/wp-content/uploads/2016/03/Us_Cancer_Facts.pdf. Accessed September 1, 2017.

2. Travis WD. Pathology of lung cancer. *Clin Chest Med*. 2011. 32(4): 669-692. doi: 10.1016/j.ccm.2011.08.005

3. Kim MJ, Shin HC, Shin KC, Ro JY. Best immunohistochemical panel in distinguishing adenocarcinoma from squamous cell carcinoma of lung: Tissue microarray assay in resected lung cancer specimens. *Ann Diagn Pathol*. 2013. 17(1): 85-90. doi: 10.1016/j.anndiagpath.2012.07.006

4. Zhan C, Yan L, Wang L, et al. Identification of immunohistochemical markers for distinguishing lung adenocarcinoma from squamous cell carcinoma. *J Thorac Dis*. 2015. 7(8): 1398-1405. doi: 10.3978/j.issn.2072-1439.2015.07.25

5. Yu KH, Zhang C, Berry GJ, et al., Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun*. 2016; 7: 12474. doi: 10.1038/ncomms12474

6. Kosinski M. RTCGA.rnaseq (Software). R package version 20151101.6.0. Rna-seq datasets from The Cancer Genome Atlas; 2016. doi: 10.18129/B9.bioc.RTCGA.rnaseq

7. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014; 511(7511): 543-550. doi: 10.1038/nature13385

8. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012; 489(7417): 519-525. doi: 10.1038/nature11404

9. Jian Xiao XL, Chen X, Zou Y, et al. Eight potential biomarkers for distinguishing between lung adenocarcinoma and squamous cell carcinoma. *Oncotarget*. 2017. Web site. http://www.readcube.com/articles/10.18632/oncotarget.17606.

10. Savci-Heijink CD, Kosari F, Aubry M-C, et al. The role of desmoglein-3 in the diagnosis of squamous cell carcinoma of the lung. *Am J Pathol*. 2009; 174(5): 1629-1637. doi: 10.2353/ajpath.2009.080778