Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 04, pp. 29728-29734, April, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

HIGH RISK FOR LUNG AND BLADDER CANCERS IN THE ARABIAN GULF POPULATION ASSOCIATED WITH EGFR GENETIC VARIANTS REGULATING EXPRESSION

[1,☯]Raed H. Qaddourah, [1]Sonia Bourguiba-Hachemi, [1]Maryam Marzouq, [2]Abdel Halim Salem, [3]Hamdi Jarjanazi, [4]Miral Mashhour, [5]Wassim F. Raslan, and *,[1]Dahmani Fathallah, M.

[1]Life Sciences Department, Health Biotechnology Program & King Fahd Chair for Biotechnology, College of Graduate Studies, Arabian Gulf University, PO BOX 26671 Manama- Kingdom of Bahrain
[2]Department of Anatomy, College of Medicine and Medical Sciences, Arabian Gulf University, PO Box 26671, Manama, Kingdom of Bahrain
[3]Environmental Monitoring and Reporting Branch, Ontario Ministry of the Environment, 125 Resources Road, Toronto, ON M9P 3V6, Canada
[4]Department of Pathology and Laboratory Medicine, King Fahad Specialist Hospital, P.O. Box 15215, Dammam, Kingdom of Saudi Arabia
[5]Johns Hopkins Saudi Aramco Medical Services Organization, Kingdom of Saudi Arabia

ARTICLE INFO

Article History:
Received 20th January, 2016
Received in revised form
23rd February, 2016
Accepted 28th March, 2016
Published online 26th April, 2016

Key words:

EGFR, IHC,
Expression,
Gene polymorphism,
Cancer.

ABSTRACT

The Epidermal growth factor receptor (EGFR) expression is controlled by a polymorphic CA simple sequence repeat 1 [CA-SSR1] in intron one, and two single nucleotide polymorphisms (SNPs) in the promoter region (-216 G/T,-191C/A). We investigated the EGFR cell surface expression level and analyzed these genomic markers in six types of adenocarcinoma in Arabian patients. The highest expression of EGFR was observed in lung (93%) and bladder (95%) cancers. The CA-SSR1 long allele (OR=1.641, 95% CI 1.01-2.66, p=0.044) and the SNP-191C/A **A** allele (OR=3.87, 95% CI 1.51-9.87, p=0.004) were significantly associated with lung adenocarcinoma (LA). Haplotype analysis revealed three haplotypes within *EGFR* gene were significantly associated with lung and bladder cancer risk. However, after Bonferroni correction only haplotype SSR1L/-216G/-191A remained significantly associated with increased lung cancer risk (OR=4.652, 95% CI 1.47-14.71, p=0.012). This is the first observation of a significant association between genetic markers influencing EGFR expression level and the risk of developing lung cancers in Arabs. Furthermore, the molecular event that favors the selection of the long SSR1 allele in Arab lung cancer patients seems to be different from the somatic mutation that favors the selection of the short allele in the East Asian lung cancer patients.

Citation: Raed H. Qaddourah, Sonia Bourguiba-Hachemi, Maryam Marzouq *et al*. 2016. "High risk for lung and bladder cancers in the Arabian Gulf population associated with EGFR genetic variants regulating expression", *International Journal of Current Research*, 8, (04), 29728-29734.

INTRODUCTION

The human epidermal growth factor receptor [EGFR, HER1, erbB1] member of the ErbB receptor tyrosine kinase family, known to induce tumor growth and progression (Olayioye *et al*., 2000), has emerged as promising target for cancer therapy. It has been shown that the EGFR is frequently overexpressed in a large variety of human tumors such as, head and neck, esophageal, breast, non-small cell lung cancer (NSCLC), colorectal, gastric, pancreatic, renal, bladder, prostate, and ovarian cancers with an expression level varying from low to high according to the tumor type and the population studied (Nicholson *et al*., 2001, Salomon *et al*., 1995). Immunotherapy interfering with the functions of growth factors receptors has introduced a breakthrough in cancer therapy and achieved significant leap forward for response rate, culminating in complete remission in some cases. Antibodies targeting the EGFR such as Cetuximab (Erbitux) are being used successfully for close to a decade now (Karapetis *et al*., 2008, Martinelli *et al*., 2009, Van Cutsem *et al*., 2009, Pillay *et al*., 2011). However several important issues still need to be addressed with EGFR such as how to best evaluate EGFR expression and how gene amplification, mutations and genomic variability affect EGFR transcription level or cell surface expression. The recent data on the molecular basis of

Corresponding author: Dahmani Fathallah, M.
Life Sciences department, Health Biotechnology Program & King Fahd Chair for Biotechnology, College of Graduate Studies, Arabian Gulf University, PO BOX 26671 Manama- Kingdom of Bahrain.
☯Contributed equally to this work

EGFR expression in cancer pathophysiology suggest that the success of immunotherapy using antibodies against EGFR depends heavily on a good assessment of EGFR expression. Standardized IHC protocols, with harmonized scoring methods backed by exhaustive genetic information on the EGFR expression pattern are required to make reliable interpretation.
The gene coding the 170 kDa membrane-spanning EGFR protein is located on the short arm on human chromosome 7 (7p12.1-12.3) (Gao *et al*., 2008, Davies *et al*., 1980), and it is organized in 28 exons. The genomic polymorphism of the *EGFR* loci has been extensively investigated and was shown to display interethnic genetic variability (Liu *et al*., 2003, Nomura *et al*., 2007). It has been reported that transcription of the *EGFR* gene is modulated by a highly polymorphic CA simple sequence repeat 1 (CA-SSR1) located in intron 1 in proximity to a transcriptional enhancer element (Gebhardt *et al*., 2000). The length of this CA-SSR1varied from 14 to 21 in Caucasian population, with the most common allele containing 16 (42%), 18 (20%) and 20 (26%) repeats. An inverse correlation has been reported between the number of CA-SSR1presents and the *EGFR* gene transcriptional activity (Desai *et al*., 2003).

In addition at least 12 single nucleotide polymorphisms (SNP) were described in the 5' untranslated/regulatory region of the *EGFR* gene (Liu *et al*., 2005). Two of these SNPs, found within the promoter sequence, were reported to regulate the *EGFR* gene expression and hence the EGFR cell surface expression (Liu *et al*., 2005). The 216 G/T (rs712829) is located in the Sp1 binding site and the 191 C/A (rs712830) four bp upstream of one of six transcription initiation sites (Johnson *et al*., 1988, Kageyama *et al*., 1988). Despite the high number of studies performed on this subject, it still not definitely established whether the EGFR expression level correlates with patient prognosis and can be used to predict response to therapy. This is probably due to the fact that EGFR cell surface expression is affected by a number of factors including the natural genetic variability, gene amplification and mutations in the EGFR coding region and activation of cell proliferation through an alternative pathway such KRAS mutations (Han *et al*., 2007, Buerger *et al*., 2000, Nomura *et al*., 2007). Therefore good knowledge of the genetic characteristics of the patients not only will help in designing a better medication but will also help in establishing more efficient therapeutic procedures to improve the clinical outcome of cancer therapies targeting EGFR. Within this context, little data is available on the relevance of *EGFR* gene polymorphism to cancer in Arabian Gulf people.

To study the expression level status of EGFR and the genetic variability of EGFR in cancer patients of Arab origin; we analyzed, using a standardized automated IHC procedure, the IHC staining profile of EGFR in six different types of carcinomas in a total of 182 cancer patients from the Arabian Gulf region. Paralleled to this study, we carried out the same genetic study on an ethnically-matched healthy control group. Our study focused on the EGFR locus three polymorphic markers (CA-SSR1, 216 G/T, 191 C/A) shown to regulate EGFR expression level in experimental studies. This observation was validated in some cancer types and not others and varied between studies in patients of different ethnic origin.

MATERIAL AND METHODS

Patients

We investigated a total of 182 cancer patients from the Arabian Gulf region distributed as follow: Colorectal Adenocarcinoma (n=20); Breast Infiltrating Ductal Carcinoma (n=20); Lung Adenocarcinoma (n=63), Ovarian Serous Carcinoma (n=20); Bladder Urothelial Carcinoma (n=50) and Thyroid Papillary Carcinoma (n=9). The patients and 114 healthy control individuals were selected on the basis of their Arabic name and all participants signed the required approval consent. Access to patient's data was strictly confined to the study senior investigators. The patient's personal information's were encoded and the clinical and histological characteristics along with the genetic data were compiled in a database. All database files were encrypted and password protected. The Institution Review Board (IRB) approved the study.

Tumors Grading

As grading systems differ based on cancer type, we used the grading systems recommended by the world Health Organization (WHO) and listed in the latest edition of the cancer staging manual of the American Joint Committee on Cancer (AJCC) (Edge *et al*., 2010).

Immunohistochemical Analysis

We carried out the IHC study on Formalin-fixed Paraffin embedded tissues sections on a total of 104 cancer patients of Arabic descend from the Gulf region. All biopsies were taken before the patients undergo any therapeutic course. We standardized IHC by using a Ventana Benchmark slide staining automated platform (Ventana Medical Systems Inc. Strasbourg, France) to carry out IHC based on the manufacturer's recommendations using anti-EGFR (3C6) primary antibody.

Two different readings of the immunostained sections were performed under light microscope by one pathologist at an interval of at least six months. No interpretational discrepancies were encountered between the two readings. The pathologist was blinded about the molecular results of the study. The staining cutoff for positive results was 10% of cells showing complete strong membranous stain. To standardize the IHC data ,the scoring system used for the determination of EGFR expression and the staining pattern of tumor cell membranes was the metric 0, 1+, 2+ and 3+ recommended by the American Society of Clinical Oncology/ College of American Pathologists (ASCO/CAP) reporting guidelines (Wolff *et al*., 2007)and used for Her2neu. Score zero, no staining (negative), 1+ score (negative): faint incomplete membranous staining in more than 10 % of the tumor cells. 2+ score (positive): weak complete membranous staining in 10-30% of the tumor cells. 3+ score (positive): strong complete membranous staining in more than 30 % of the tumor cells.

Genomic DNA extraction

DNA was extracted from peripheral blood samples and PFFE tissues using Blood & Cell Culture DNA Mini Kit (Cat. # 13323, Qiagen, USA) and QIAamp DNA FFPE Tissue Kit (Cat. # 51306, Qiagen, USA) respectively according to the manufacturer's instructions.

EGFR Genetic typing

EGFR CA-SSR1 genotyping

Analysis of EGFR CA-simple sequence repeat 1 (CA-SSR1) length polymorphism was performed using GenoScreen DNA typing technical platform (Lille, France). Briefly, the repeat region was ampli ed using the following primers (ThermoFisher Scientific GmbH, Germany): forward 5'-GGGCTCACAGCAAACTTCTC-3'and reverse 5'HEX-AAGCCAGACTCGCTCATGTT-3'. Polymerase chain reaction (PCR) (25 µl) for each sample was set in 0.2 mL thin-walled tube using 20 ng of DNA, 1 pmol of each primer, 6 pmol dNTPs, 37.5 pmol $MgCl_2$ and 1 unit of Fast start Taq DNA polymerase (Roche). PCR reaction was carried out in MJ research Thermal Cycler (Waltham, MA, USA) using the following conditions: initial denaturation of 95°C for 10 min followed by 40 cycles at 95°C for 30 sec, 55°C for 30 sec and 72°C for 1 min and a nal extension at 72°C for10 min. 1µl of each ampli ed product was diluted at 1/50 then migrated on sequencer ABI 3730 XL (Applied Biosystems) and analyzed by GeneScan software. For GeneScan, 1 µL of diluted PCR product mixed with 0.25 µL of LIZ 500 (Applied Biosystems) and 8.75 µL of HiDi-Formamide (Applied Biosystems). Migration conditions on sequencer were as follows: injection time: 15 sec, run voltage: 15 kV and run time: 1600 sec. The raw data was analyzed using the GeneMapper Software (Applied Biosystems) to estimate the alleles.

Single Nucleotide Polymorphism Genotyping

Two SNPs rs71289 (-216 G/T) and rs712830 (-191 C/A) were genotyped using Genoscreen technical platform KBioscience Competitive Allele-Specific PCR genotyping technique (KASP) version 4.0 SNP (Lille, France). Briefly, KASP detects SNPs via FRET (Fluorescence Resonance Energy Transfer). For each SNP, one common reverse primer and two allele-specific forward primers (that differ at their 3' ends). The 296 samples and 2 negative controls (water) were genotyped following KASP PCR protocol. Genotyping reactions were performed in 2xHC16 thermal cycler in a final volume of 5 µl containing 2.5 µl of genomic DNA (12 ng/µl), 2.5 µl of 2X KASP reaction mix (Taq polymerase enzyme and the passive reference dye ROX, KBioscience) and 0.07 µl of the KASP assay primer mix (consisting of 12 µM of each allele-specific forward primers and 30 µM reverse primer). The following cycling conditions were used: a hot-start Taq polymerase activation step (94°C for 15 min) followed by a 10 touchdown cycles of 94°C for 20sec, 65-57°C for 1 min (decreasing 0.8°C per cycle), and 26 cycles of 94°C for 20 sec and 57°C for 1 min. Fluorescence detection of the reactions was performed using a FRET reading instrument; BMG PHERAstar and genotyping data was analyzed by KlusterCaller software (Kbioscience).

Statistical analysis

For statistical data analysis, we used the SPSS statistical software package version 19.0 (SPSS Chicago Illinois, USA). We applied the Chi square and Pearson tests when appropriate. In the control group, we systematically checked the Hardy Weinberg equilibrium for all the genetic markers studied. We also used the SHEsis software platform (Shi and He, 2005, Li *et al*., 2009), and the Haploview software package version 4.2 to construct the haplotypes and to study linkage disequilibrium and polymorphic loci genetic association. P value < 0.05 was considered significant.

RESULTS

EGFR expression analysis by Immunohistochemistry

We generated immunohistochemistry EGFR data on 182 Arabian cancer patients grouped in six consecutive types of carcinomas.

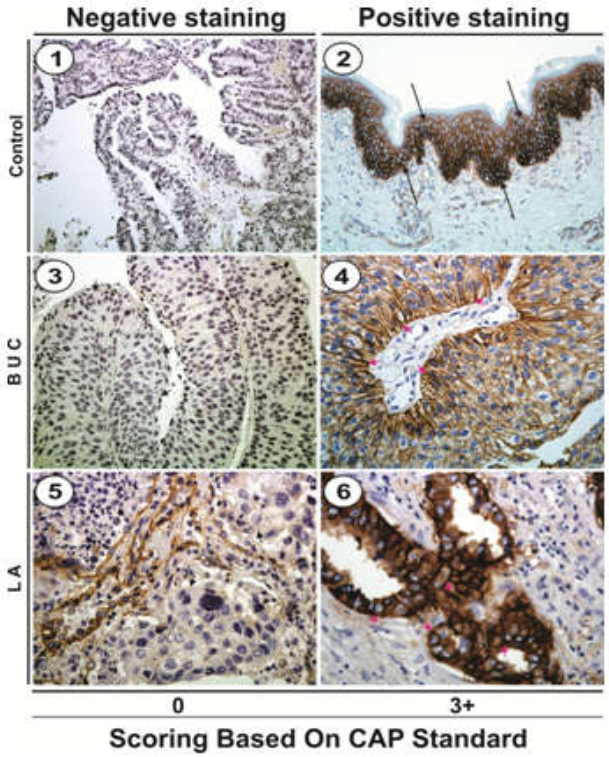

Figure 1. Immunohistochemical staining for EGFR detection in Bladder Urothelial Carcinoma (BUC) and Lung Adenocarcinoma (LA). EGFR IHC staining in tissue samples from BUC (Panels 3 and 4); LA (Panels 5 and 6). The staining of tumor cell membranes with the Ventana EGFR 3C6 antibody was graded as negative or positive based on CAP standard scoring 0, 1+, 2+ & 3+. 0 was scored when negative or unspecific staining (black arrow head). 3+ was scored when >30% of tumor cells exhibit a strong immune-reactivity and complete membrane staining (fuchsia arrow head). Ovarian serous carcinoma tissue was stained in the absence of primary antibody as a negative control (1). Skin tissue was used as positive control (2) exhibiting an intense complete membranous staining of >30% (black arrows). Magnification (X 200)

As Shown in Figure 1, EGFR positive tissues sections showed brown or golden brown membranous immunostaining of variable intensity. By applying the scoring system for HER-2/neu expression accredited by the college of American pathologist for use in breast cancer diagnosis, we found the overall frequency of positive staining for the EGFR (HER1) to be 50%. The estimation of EGFR expression for each type of cancer showed higher rates of positive staining in lung and bladder carcinomas, respectively 93 and 95% while breast infiltrating ductal carcinoma cases had the lowest expression rate, 8% (data not shown). All Lung Adenocarcinoma and Bladder Urothelial Carcinoma samples with high tumor grade (stage III-IV) had elevated EGFR expression level (score 3+) (Fig.1).

Genetic Analysis

Based on the results obtained by IHC for the level of EGFR expression on the six types of tumors investigated, we focused our genetic studies on lung adenocarcinoma and bladder urothelial carcinoma groups of patient.

EGFR CA-SSR1 genetic typing

Short and long alleles were calculated according to the different alleles repeat length and the mean was taken as a cut off value: Short ≤ 35 >Long. A total of eight alleles were found in both the control and patients groups (LA and BUC). In the control group, the number of CA-SSR1 varied from 15 to 22. Allelic distribution was trimodal with predominance of allele 16 (51%), allele 17 (10%) and allele 20 (20%), the rest being minor alleles.

No deviations from Hardy-Weinberg equilibrium (HWE) were observed in the control group (χ^2=1.94, df=1, Pearson's p=0.163). We compared the differences in frequency distributions of CA-SSR1 marker alleles (short and long) between cases and controls by χ^2 test (Table 1) and identified the minor long allele >35 repeat associated with the risk of lung adenocarcinoma (χ^2=4.026, OR=1.641, 95% CI 1.010-2.667 and p=0.044). No such association was observed in the BUC group (Table1).

Table 1. Association of EGFR CA-SSR1 Alleles with Lung Adenocarcinoma (LA) and Bladder Urothelial Carcinoma (BUC)

	Control (n=114)	LA (n=63)	BUC (n=50)
EGFR CA-SSR1	L0.349/S0.651	L0.468/S0.532	L0.398/S0.602
χ^2		4.026	0.622
OR		1.641	1.234
[95% CI]		[1.010-2.667]	[0.730-2.085]
p value		0.044	0.43

LA: Lung Adenocarcinoma; BUC: Bladder Urothelial Carcinoma; χ^2: chi square; OR: Odd Ratio, CI: Confidence Interval.

SNPs rs712829 (-216G/T) and rs712830 (-191C/A) genotyping

The observed genotype frequencies of rs712829 (-216G/T) and rs712830 (-191C/A) complied with the HWE in the control subjects (χ2=1.005, df=1, Pearson's p=0.315 and χ2=0.121, df=1, Pearson's p=0.727, respectively). For SNP rs712830 the minor allele A showed a strong allelic association with LA (Table 2).

Table 2. Allelic association of SNPs re712829 and rs7122830 with Lung Adenocarcinoma (LA) and Bladder Urothelial Carcinoma (BUC)

	Control (n=114)	L A (n=63)	B U C (n=50)
SNP_ID rs712829	G 0.537 / T 0.463	G 0.643 / T 0.357	G 0.622 / T 0.378
χ^2		3.407	1.757
OR [95% CI]		1.553 [0.971-2.484]	1.420 [0.844-2.388]
p value		0.06	0.184
p value adj.*		0.12	0.368
SNP_ID rs712830	C 0.968 / A 0.032	C 0.885 / A 0.115	C 0.970 / A 0.030
χ^2		9.072	0.012
OR [95% CI]		3.870 [1.517-9.874]	0.923 [0.233-3.647]
p value		0.002	0.909
p value adj.*		0.004	1

LA: Lung Adenocarcinoma; BUC: Bladder Urothelial Carcinoma; χ^2: chi square; OR: Odd Ratio, CI: Confidence Interval; *p* value adj.*: *p* value was adjusted by Bonferroni correction.

Table 3. EGFR Haplotype Association with Bladder Urothelial Carcinoma in the Arabian Gulf population

Haplotypes	Freq (case)	Freq (control)	χ^2	*p* value	*p* value adj.*	OR [95%CI]
SSR1.S/-216 G/-191 C	0.256	0.275	0.193	0.660	1	0.869 [0.46-1.62]
SSR1.S/-216 T/-191 C	0.307	0.353	0.716	0.397	1	0.775 [0.43-1.39]
SSR1.L/-216 G/-191 C	0.369	0.232	4.187	0.040	0.12	1.866 [1.02-3.40]
SSR1.L/-216 T/-191 C	0.031	0.111	4.603	0.031	0.093	0.245 [0.06-0.97]

S: short, L: Long, χ2: Chi square; OR: Odd Ratio, CI: Confidence interval. *p* value adj.
*: *p* value was adjusted by Bonferroni correction.

Table 4. EGFR Haplotype Association with Lung Adenocarcinoma in the Arabian Gulf population

Haplotypes	Freq (case)	Freq (control)	χ2	p value	*p* value adj.*	OR [95% CI]
SSR1. S/-216 G/-191 C	0.211	0.275	1.329	0.248	0.744	0.706 [0.39-1.28]
SSR1. S/-216 T/-191 C	0.326	0.353	0.210	0.646	1	0.883 [0.52-1.50]
SSR1. L/-216 G/-191 C	0.307	0.232	1.770	0.183	0.549	1.468 [0.83-2.59]
SSR1. L/-216 G/-191 A	0.120	0.029	8.044	0.004	0.012	4.652 [1.47-14.71]
SSR1. L/-216 T/-191 C	0.036	0.111	4.811	0.028	0.084	0.295 [0.09-0.93]

However for SNP rs712829 the major allele G showed only a tendency toward a significant association with LA. No allelic association was observed with BUC (Table 2).

Haplotype Analysis

The analysis as a mini haplotype of the EGFR markers within the regulatory region that controls transcription and protein cell surface expression; CA-SSR1 variant, rs712829 (-216G/T), and rs712830 (-191C/A); showed the existence of 4 and 5 out of 8 possible haplotypes in BUC and LA, respectively. The results of the associations between these EGFR haplotypes and the risk of LA and BUC cancers are listed in table 3 and table 4. The data show that haplotype SSR1L/-216G/-191C is significantly associated with an increased risk of BUC (p=0.040, OR=1.866, 95%CI 1.02-3.40, Table 3), while SSR1L/-216T/-191C is significantly a protective haplotype (p= 0.031, OR=0.245, 95%CI 0.06-0.97, Table 3). However, after Bonferroni correction for multiple comparisons, the statistical significance was lost (Table 3). Interestingly the SSR1L/-216T/-191C is also a protective haplotype against LA (p= 0.028, OR=0.295, 95%CI 0.09-0.93, Table 4), while haplotype SSR1L/-216G/-191A is significantly associated with an increased risk of lung adenocarcinoma (p= 0.004, OR=4.652, 95% CI 1.47-14.71, Table 4). After adjusting for multiple comparisons, only the later (SSR1L/-216G/-191A) remained significant (p= 0.012, Table 4).

DISCUSSION

EGFR is an important cell surface signaling molecules and deregulation of the EGFR signaling pathway via elevated expression level or through mutations results in cell proliferation and loss of apoptosis. Therefore EGFR is critically involved in the molecular pathogenesis of cancer (Jorissen *et al*., 2003, Yarden and Sliwkowski, 2001). EGFR is being successfully used as a target in cancer therapy. However the mechanism by which the EGFR germline genetic variability contributes to carcinogenesis is not fully elucidated despite the elevated number of genetic studies carried out in different cohort. In this case-control study, we investigated whether EGFR genetic variants are associated with carcinogenesis in the Arabian Gulf population. The Arabian gulf population has been undergoing dramatic lifestyle changes in the last two to three decades and in which prevalence of cancer is significantly rising according to reports from the International Agency for Research on Cancer (IARC), and the specialized cancer agency of the World Health Organization WHO (Ferlay *et al*., 2012, Bray *et al*., 2013). We have focused our study first on standardizing the IHC method to investigate EGFR expression, then we carried out genetic analysis of three markers located at the 5'end of the *EGFR* gene previously shown to modulate EGFR expression level in *in vitro* experimental studies using cell lines (Liu *et al*., 2007). However this effect was not consistently confirmed in all cancers and varied between the cancer types and the ethnic origin of the patients studied (Brandt *et al*., 2004, Shitara *et al*., 2012, Jung *et al*., 2012, Kim *et al*., 2009, Nomura *et al*., 2007, Etienne-Grimaldi *et al*., 2005). Investigation of this issue in the Arabian Gulf population provided new data on the determination of EGFR expression status in six different types of cancer using IHC as well as interesting data on the genetic polymorphism of potentially functional genetic variants. Standardization of IHC relies heavily on the adjustment of the antibody preparations and detection methods, as well as the scoring system necessary to produce an assay for EGFR evaluation that provides consistent, comparable results. To standardize IHC for the determination of tumor EGFR, we attempted to overcome the limitations of the method by first using the Ventana Benchmark slide staining automated platform and the commercial anti-EGFR Monoclonal antibody 3C6 that recognize even EGFR variant III. We couldn't find a study that demonstrates the superiority of an antibody versus another in measuring EGFR Cell surface expression by IHC. In addition we adopted a scoring method validated by the college of American pathologists (CAP) for the diagnosis of breast cancer patients through measuring the Her2Neu cell surface expression. The data we obtained for the six types of cancer we studied is consistent with what was reported using IHC on EGFR expression level in cancer patients from populations other than Arabs. Furthermore our data show a very high prevalence of EGFR expression in lung and bladder adenocarcinoma.

Our observations combined with those reported in the literature suggest that standardization of IHC for the study of EGFR expression could make it the most recommended method for the clinical evaluation of EGFR expression status in cancer setting. Since EGFR has been reported as being a major driver of malignancy in several types of cancer and that polymorphism in the *EGFR* gene 5 'regulatory region was shown to control EGFR expression level, and based on the results obtained by IHC for the level of EGFR expression, we investigated the association of the genetic variability of three markers in the gene's regulatory region rs712829 (-216G/T) and rs712830 (-191C/A) and intron 1 CA-SSR1 in Lung adenocarcinoma and Bladder urothelial carcinoma groups of patient. Our study showed no significant difference in the EGFR expression according to the length of CA-SSR1. The length of CA-SSR1 polymorphism differs by ethnicity and tends to be longer in Asians than in Europeans or African-Americans (Jung *et al*., 2012). In the current study, we show that in Arabs the short CA-SSR1 allele (up to 16 repeats) is predominant and that the pattern of the CA-SSR1 allele's frequency is closer to that of the Caucasians (Liu *et al*., 2003). On the other hand the association's study of CA-SSR1 alleles with cancer showed a significant association of the long allele (L) with lung adenocarcinoma. Interestingly in a study on NSCLC in patients from various ethnic backgrounds Nomura *et al*. (2007) reported a selective amplification of the shorter CA-SSR1 allele in tumors harboring a mutation resulting in higher EGFR expression in East Asian where the long allele is predominant (Nomura *et al*., 2007). This situation is at the opposite of our observation in the Arab population where the CA-SSR1 short allele is predominant while the long allele seems to be selected in patients with lung tumors even though we do not have yet exhaustive information's on the EGFR somatic mutation status of the Arab patients. This suggests that the molecular event that favors the selection of the long allele in the Arab patients is different from the somatic mutation that favors the selection of the short allele in the East Asian lung cancer patients.

In addition, assuming that the minor allele of the SNP was a risk allele compared to the wild type major allele, the genotype analysis has shown for the first time in the Arabian Gulf population, a strong association of variant A of SNP rs712830 (-191 C/A), as a single marker, with lung adenocarcinoma (p=0.002, OR=3.870, 95%CI 1.517-9.874). This association was further confirmed with a haplotype combining the two SNPs rs712829 and rs712830 and the CA-SSR1. We have shown that LGA haplotype was associated with a significantly increased risk of lung cancer in the Arabian Gulf population. Moreover, this finding suggests that haplotype analysis could be an appropriate tool for evaluating the risk association than the single marker. Our data confirm that the natural genetic variability affects EGFR cell surface expression and is linked to the susceptibility to develop specific types of cancer; this in addition to the known intrinsic tumorogenicity-related factors including, gene amplification and mutations in the EGFR coding region and activation of cell proliferation through an alternative pathway such KRAS mutations (Stella *et al*., 2012). In conclusion our study has shed some light on the role of EGFR polymorphism in cancer setting in the Arab Gulf populations and confirmed that germline genetic variations in the EGFR 5' regulatory region might play a role in the development of cancer particularly lung and bladder adenocarcinoma in addition to other cancer types. Testing for the potentially harmful or protective genetic patterns identified in Arabs can be clinically very useful. It can help identifying people with high risk of LA or BUC whom would become eligible for screening for early detection of these two types of cancer. Meanwhile it can prevent people with the protective genetic signature undergoing unnecessary costly medical check-ups. In addition determination of EGFR expression level may help match patients with individualized treatments. Genetic testing based on this data might lead to an earlier, accurate diagnosis and prevention of LA and BUC in the Arabian Gulf region people. These positive outcomes preclude a reduction of the financial burden on the health care system.

Acknowledgments

We are grateful to all the patients and individuals for their participation. We also thank the clinicians and other hospital staff in respective centers who contributed to data collection for this study. This work was supported by a competitive grant from the Arabian Gulf University Scientific Board and by King Fahd Chair for Health Biotechnology.

Competing interests

The authors have declared that no competing interests exist.
Ethics approval The Institution Review Board (IRB).

REFERENCES

Brandt, B., Hermann, S., Straif, K., Tidow, N., Buerger, H., and Chang-Claude, J. 2004. "Modification of breast cancer risk in young women by a polymorphic sequence in the egfr gene." *Cancer Res.* 64 (1):7-12.

Bray, F., Ren, J.S., Masuyer, E., and Ferlay, J. 2013. "Estimates of global cancer prevalence for 27 sites in the adult population in 2008." *Int J Cancer.* 132 (5):1133-45.

Buerger, H., Gebhardt, F., Schmidt, H., Beckmann, A., Hutmacher, K., Simon, R., Lelle, R., Boecker, W., and Brandt, B. 2000. "Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression." *Cancer Res.* 60 (4):854-7.

Davies, R.L., Grosse, V.A., Kucherlapati, R., and Bothwell, M. 1980. "Genetic analysis of epidermal growth factor action: assignment of human epidermal growth factor receptor gene to chromosome 7." *Proc Natl Acad Sci U S A.* 77 (7):4188-92.

Desai, A.A., Innocenti, F., and Ratain, M.J. 2003. "Pharmacogenomics: road to anticancer therapeutics nirvana?" *Oncogene.* 22 (42):6621-6628.

Edge, S., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L., and Trotti, A., eds. 2010. *AJCC Cancer Staging Manual*. Edited by 7th. Vol. XV: Springer.

Etienne-Grimaldi, M.C., Pereira, S., Magne, N., Formento, J.L., Francoual, M., Fontana, X., Demard, F., Dassonville, O., Poissonnet, G., Santini, J., Bensadoun, R.J., Szepetowski, P., and Milano, G. 2005. "Analysis of the dinucleotide repeat polymorphism in the epidermal growth factor receptor (EGFR) gene in head and neck cancer patients." *Ann Oncol.* 16 (6):934-41.

Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D., Forman, D., and Bray, F. 2012. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer.

Gao, L.B., Zhou, B., Zhang, L., Wei, Y.S., Wang, Y.Y., Liang, W.B., Lv, M.L., Pan, X.M., Chen, Y.C., and Rao, L. 2008. "R497K polymorphism in epidermal growth factor receptor gene is associated with the risk of acute coronary syndrome." *BMC Med Genet.* 9 (74):1471-2350.

Gebhardt, F., Burger, H., and Brandt, B. 2000. "Modulation of EGFR gene transcription by a polymorphic repetitive sequence--a link between genetics and epigenetics." *Int J Biol Markers.* 15 (1):105-10.

Han, S.W., Jeon, Y.K., Lee, K.H., Keam, B., Hwang, P.G., Oh, D.Y., Lee, S.H., Kim, D.W., Im, S.A., Chung, D.H., Heo, D.S., Bang, Y.J., and Kim, T.Y. 2007. "Intron 1 CA dinucleotide repeat polymorphism and mutations of epidermal growth factor receptor and gefitinib responsiveness in non-small-cell lung cancer." *Pharmacogenet Genomics.* 17 (5):313-9.

Johnson, A.C., Ishii, S., Jinno, Y., Pastan, I., and Merlino, G.T. 1988. "Epidermal growth factor receptor gene promoter. Deletion analysis and identification of nuclear protein binding sites." *J Biol Chem.* 263 (12):5693-9.

Jorissen, R.N., Walker, F., Pouliot, N., Garrett, T.P., Ward, C.W., and Burgess, A.W. 2003. "Epidermal growth factor receptor: mechanisms of activation and signalling." *Exp Cell Res.* 284 (1):31-53.

Jung, M., Cho, B.C., Lee, C.H., Park, H.S., Kang, Y.A., Kim, S.K., Chang, J., Kim, D.J., Rha, S.Y., Kim, J.H., and Lee, J.H. 2012. "EGFR polymorphism as a predictor of clinical outcome in advanced lung cancer patients treated with EGFR-TKI." *Yonsei Med J.* 53 (6):1128-35.

Kageyama, R., Merlino, G.T., and Pastan, I. 1988. "Epidermal growth factor (EGF) receptor gene transcription.

Requirement for Sp1 and an EGF receptor-specific factor." *J Biol Chem.* 263 (13):6329-36.

Karapetis, C.S., Khambata-Ford, S., Jonker, D.J., O'Callaghan, C.J., Tu, D., Tebbutt, N.C., Simes, R.J., Chalchal, H., Shapiro, J.D., Robitaille, S., Price, T.J., Shepherd, L., Au, H.J., Langer, C., Moore, M.J., and Zalcberg, J.R. 2008. "K-ras mutations and benefit from cetuximab in advanced colorectal cancer." *N Engl J Med.* 359 (17):1757-65.

Kim, M.J., Shin, K.C., and Lee, K.H. 2009. "Intron 1 Polymorphism, Mutation and the Protein Expression of Epidermal Growth Factor Receptor in Relation to the Gefitinib Sensitivity of Korean Lung Cancer Patients." *Korean J Pathol.* 43 (6):509-516.

Li, Z., Zhang, Z., He, Z., Tang, W., Li, T., Zeng, Z., He, L., and Shi, Y. 2009. "A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (http://analysis.bio-x.cn)." *Cell Res.* 19 (4):519-23.

Liu, W., Innocenti, F., Chen, P., Das, S., Cook, E.H., Jr., and Ratain, M.J. 2003. "Interethnic difference in the allelic distribution of human epidermal growth factor receptor intron 1 polymorphism." *Clin Cancer Res.* 9 (3):1009-12.

Liu, W., Innocenti, F., Wu, M.H., Desai, A.A., Dolan, M.E., Cook, E.H., Jr., and Ratain, M.J. 2005. "A functional common polymorphism in a Sp1 recognition site of the epidermal growth factor receptor gene promoter." *Cancer Res.* 65 (1):46-53.

Liu, W., Wu, X., Zhang, W., Montenegro, R.C., Fackenthal, D.L., Spitz, J.A., Huff, L.M., Innocenti, F., Das, S., Cook, E.H., Jr., Cox, N.J., Bates, S.E., and Ratain, M.J. 2007. "Relationship of EGFR mutations, expression, amplification, and polymorphisms to epidermal growth factor receptor inhibitors in the NCI60 cell lines." *Clin Cancer Res.* 13 (22 Pt 1):6788-95.

Martinelli, E., De Palma, R., Orditura, M., De Vita, F., and Ciardiello, F. 2009. "Anti-epidermal growth factor receptor monoclonal antibodies in cancer therapy." *Clin Exp Immunol.* 158 (1):1-9.

Nicholson, R.I., Gee, J.M., and Harper, M.E. 2001. "EGFR and cancer prognosis." *Eur J Cancer.* 37 (4):S9-15.

Nomura, M., Shigematsu, H., Li, L., Suzuki, M., Takahashi, T., Estess, P., Siegelman, M., Feng, Z., Kato, H., Marchetti, A., Shay, J.W., Spitz, M.R., Wistuba, II, Minna, J.D., and Gazdar, A.F. 2007. "Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers." *PLoS Med.* 4 (4).

Olayioye, M.A., Neve, R.M., Lane, H.A., and Hynes, N.E. 2000. "The ErbB signaling network: receptor heterodimerization in development and cancer." *Embo J.* 19 (13):3159-67.

Pillay, V., Gan, H.K., and Scott, A.M. 2011. "Antibodies in oncology." *N Biotechnol.* 28 (5):518-29.

Salomon, D.S., Brandt, R., Ciardiello, F., and Normanno, N. 1995. "Epidermal growth factor-related peptides and their receptors in human malignancies." *Crit Rev Oncol Hematol.* 19 (3):183-232.

Shi, Y.Y., and He, L. 2005. "SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci." *Cell Res.* 15 (2):97-8.

Shitara, M., Sasaki, H., Yokota, K., Okuda, K., Hikosaka, Y., Moriyama, S., Yano, M., Kawaguchi, T., Kubo, A., Takada, M., Kitahara, N., Okumura, M., Matsumura, A., Iuchi, K., and Fujii, Y. 2012. "Polymorphisms in intron 1 of the EGFR gene in non-small cell lung cancer patients." *Exp Ther Med.* 4 (5):785-789.

Stella, G.M., Luisetti, M., Inghilleri, S., Cemmi, F., Scabini, R., Zorzetto, M., and Pozzi, E. 2012. "Targeting EGFR in non-small-cell lung cancer: Lessons, experiences, strategies." *Respiratory Medicine.* 106 (2):173-183.

Van Cutsem, E., Kohne, C.H., Hitre, E., Zaluski, J., Chang Chien, C.R., Makhson, A., D'Haens, G., Pinter, T., Lim, R., Bodoky, G., Roh, J.K., Folprecht, G., Ruff, P., Stroh, C., Tejpar, S., Schlichting, M., Nippgen, J., and Rougier, P. 2009. "Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer." *N Engl J Med.* 360 (14):1408-17.

Wolff, A.C., Hammond, M.E., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A., McShane, L.M., Paik, S., Pegram, M.D., Perez, E.A., Press, M.F., Rhodes, A., Sturgeon, C., Taube, S.E., Tubbs, R., Vance, G.H., van de Vijver, M., Wheeler, T.M., and Hayes, D.F. 2007. "American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer." *Arch Pathol Lab Med.* 131 (1):18-43.

Yarden, Y., and Sliwkowski, M.X. 2001. "Untangling the ErbB signalling network." *Nat Rev Mol Cell Biol.* 2 (2):127-37.

****