GENETIC ASSOCIATION OF ABCG2 GENE POLYMORPHISMS WITH TAC CHEMOTHERAPY RESPONSE IN TRIPLE NEGATIVE BREAST CANCER (TNBC) PATIENTS

By

DR. ROSHAIDIE BIN ABDUL RASHID

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF PATHOLOGY (MEDICAL GENETICS)

UNIVERSITI SAINS MALAYSIA
2020

ACKNOWLEDGEMENT

First and foremost, Alhamdulillah. Praise be to Allah, The Most Merciful and The Most Compassionate for the guidance and the aid in completing this task, without Whom, this would not have been possible.

My enormous gratitude to my main supervisor, Prof. Dr. Ravindran Ankathil who was a perpetual figure throughout every step of this daunting task, from conception to its final product. His patience, dedication, and vast knowledge were among the few things I would forever be indebted to him for. He provided me with ample time and flexibility in conducting my dissertation, as well as space and freedom without sacrificing the quality of this dissertation. His guidance and opinions were of utmost benefit at every step of this dissertation's production and hopefully this dissertation would attest to that.

I would like to express my thanks to Dr. Andee Dzulkarnaen and Dr. Maya Mazuwin Yahya from Department of Surgery for their help in patient recruitment. To Associate Professor Dr Md Salzihan Md Salleh from Department of Pathology whose contribution as a histopathologist was of utmost importance in this study.

Not forgetting Dr. Najib Majdi Bin Yaacob from the Department of Biostatistics and Research Methodology whose opinion helped in many ways during the analysis part of this dissertation.

I would also like to sincerely thank Dr. Ahmad Aizat Abdul Aziz, then a fellow postgraduate student, and now a lecturer in Human Genome Centre for his patience and countless help at every stage of this dissertation. Thank you so much.

My gratitude also goes out to Associate Prof. Dr. Sarina Sulong, Director of Human Genome Centre, Prof. Dr. Zilfalil Alwi and Dr Nazihah Mohd Yunus for their kind guidance and input.

This thanksgiving excerpt will not be complete without me thanking the staff of Human Genome Centre, namely Puan Siti Mariam, Puan Norhidayah, Puan Nurul Alia, Puan Zulaikha Bakar, Encik Nik Mohd Zulfikri, and Puan Fatin from Cytogenetics Unit, Dr Fadhlina, Puan Fiza and Encik Zaki from Molecular Unit. Their contribution, be it small was of a great deal to me.

I would also like to thank my fellow Mpath, MSc and PhD comrades, whose knowledge, experience, joy and laughter were both the driving force and the life of this journey, as postgraduate students. Thank you to Universiti Sains Malaysia for approving the proposal of this study and providing me with a safe and well-equipped institution to conduct this study in. Thank you to my loved ones for their distant prayers and words of encouragement throughout the process of this study.

Finally, many thanks to those people who have helped me in any way or form but whose names are not mentioned here, do know that your contributions will not go unthanked.

TABLE OF CONTENTS

LIST OF SYMBOLS, ABBREVIATIONS OR NOMENCLATURE

ABCG2: ATP-binding cassette super-family G member 2

BCRP: Breast cancer resistance protein

BP: Base pairs

CML: Chronic Myeloid Leukemia

CT: Computerized tomography

DNA: Deoxyribonucleic acid

EDTA: Ethylenediaminetetraacetic acid

EMT: Epithelial mesenchymal transition

ER: Estrogen receptor

FFPE: Formalin-fixed paraffin-embedded

HER2: Human epidermal growth factor receptor 2

HRPZ: Hospital Raja Perempuan Zainab

HUSM: Hospital Universiti Sains Malaysia

IDC: Invasive Ductal Carcinoma

IM: Imatinib Mesylate

MMR: Major molecular response

MREC: Medical records

MRI: Magnetic resonance imaging

NOS: Not otherwise specified

NPI: Nottingham prognostic index

OR: Odds ratio

OS: Overall survival

PCR: Polymerase chain reaction

PR: Progesterone receptor

RFLP: Restriction fragment length polymorphism

SD: Standard deviation

SERM: Selective Estrogen Receptor Modulator

SNP: Single nucleotide polymorphism

SPSS: Statistical Product and Service Solutions

TAC: Taxane-Adriamycin-Cyclophosphamide

TNBC: Triple negative breast cancer

UV: Ultraviolet

WHO: World Health Organization

ABSTRAK

KEHUBUNG-KAITAN GENETIK DI ANTARA VARIASI GEN ABCG2 DAN RESPON KEMOTERAPI TAC DI KALANGAN PESAKIT KANSER PAYUDARA TIGA NEGATIF

Pengenalan

Kanser payudara tiga negatif (TNBC) adalah di antara jenis kanser payudara yang sangat agresif. Disebabkan oleh kekurangan reseptor-reseptor hormone, kemoterapi menggunakan rejim *Taxane-Adriamycin-Cyclophosphamide* (TAC) telah menjadi rawatan utama untuk kanser payudara tiga negative, dan ia juga adalah satu-satunya pilihan rawatan untuk kanser ini. Namun begitu, TNBC mempunyai karakter yang berkait rapat dengan risiko yang tinggi untuk pengulangan, metastasis, dan juga rintangan kepada kemoterapi. Ia juga berhubung kait dengan variasi di kalangan individu berbeza dari segi respon terhadap rawatan. Walaupun banyak mekanisme telah diutarakan tentang rintangan terhadap ubat antikanser, mekanisme yang tepat tentang rintangan kemoterapi di kalangan pesakit TNBC masih tidak jelas. Pengangkut yang cacat merupakan satu mekanisme utama dalam rintangan kemoterapi. Protin pengangkut antara membran seperti protin rintangan kanser payudara (BCRP) yang dikod oleh gen ABCG2 boleh menyebabkan rintangan kepada kemoterapi dengan cara menaikkan transportasi ubat-ubatan ke luar sel, lalu mencegah keberkesanan molekular ubat ke atas sel kanser. Hipotesis kami adalah, variasi genetik dalam gen

ABCG2 yakni G34A dan C421A, yang mampu untuk menghindarkan transportasi keluar substrat, mungkin berhubung kait dengan rintangan kemoterapi di kalangan pesakit-pesakit TNBC. Dan kami telah mereka penyelidikan ini untuk menguji hipotesis tersebut.

Metodologi

Sampel darah daripada 76 pesakit TNBC di Malaysia yang telah melalui rejim kemoterapi *Taxane-Adriamycin-Cyclophosphamide* (TAC) telah diambil dan disimpan di dalam tiub EDTA. DNA daripada sampel-sampel darah tersebut diekstrak dan kemudian diklasifikasikan mengikut genotip dengan menggunakan teknik Tindakbalas Berantai Polimeras-Polimorfisme Panjang Jalur Terpotong (PCR-RFLP). Genotip-genotip ini kemudiannya dikategorikan kepada homozigot jenis liar, heterozigot dan homozigot varian bergantung kepada saiz jalur di dalam gel elektroforesis. Perbezaan frekuensi genotip antara rintangan kemoterapi dan respon kemoterapi dikira menggunakan analisis tabulasi seberang. Hubungan antara varian-varian dan respon kemoterapi TAC di kalangan pesakit kanser payudara tiga negatif (TNBC) dinilai dengan menggunakan analisis regresi logistik dan memperolehi purata kebarangkalian (OR) dengan 95% selang keyakinan (CI) menggunakan perisian SPSS versi 20.2. Frekuensi haplotip antara dua kumpulan, rintangan kemoterapi dan respon kemoterapi dan juga hubungan di antara haplotip dan rintangan kemoterapi dinilai menggunakan perisian Haploview v4.2.

Keputusan

Genotip heterozigot (GA) dan homozigot (AA) untuk variasi gen *ABCG2* G34A telah menunjukkan kehubung-kaitan yang rendah dengan rintangan kemoterapi yang signifikan [masing-masing (OR= 0.303, p=0.029) dan (OR= 0.151, p=0.011)]. Manakala genotip heterozigot (CA) dan homozigot (AA) untuk variasi gen *ABCG2* C421A telah menunjukkan kehubung-kaitan yang rendah dengan rintangan kemoterapi walaupun tidak signifikan [masing-masing, (OR=0.481, p=0.251; OR=0.412, p=0.113)]. Alel A untuk kedua-dua variasi [masing-masing (OR=0.320, p=0.002 dan OR=0.487, p=0.039)] dan juga haplotip GA, CA dan AA [masing-masing, (ORs=0.020,0.0002, dan 0.00004)] telah menunjukkan kehubung-kaitan yang rendah dengan rintangan kemoterapi yang signifikan. Adalah wajar jika dicadangkan bahawa genotip varian (AA) dan alel varian (A) untuk polimorfisme G34A dan C421A boleh mengurangkan ekspresi dan aktiviti pengangkut *ABCG2*, mengurangkan aktiviti pengeluaran oleh ubat kemoterapi dan oleh itu menambahbaikkan keberkesanan kemoterapi dan sekaligus mengurangkan risiko rintangan kemoterapi.

Kesimpulan

Penemuan kami mencadangkan yang variasi *ABCG2* G34A mungkin mempunyai potensi untuk dijadikan sebagai penanda biologi untuk meramalkan respon kemoterapi di kalangan pesakit-pesakit TNBC.

ABSTRACT

Introduction

Triple negative breast cancer (TNBC) is one of the most aggressive subtypes of breast cancer. Due to the lack of hormone receptors, chemotherapy using Taxane-Adriamycin-Cyclophosphamide (TAC) regimen remains the mainstay of TNBC treatment, and is the only choice of treatment available for TNBC. However, TNBC is characterized by high risk of recurrence, metastasis, and chemoresistance, and is associated with inter-individual variability in treatment response. Despite several mechanisms have been implicated in anticancer drug resistance, the exact mechanism of chemoresistance in TNBC remains unclear. A defective transport is a major mechanism of resistance. Transmembrane transport proteins such as breast cancer resistance protein (BCRP) encoded by *ABCG2* can cause chemoresistance by means of increased efflux transportation of the drug out of the cell, hindering its molecular action on cancer cells. It was hypothesized that, the genetic variations G34A and C421A of ABCG2 gene, that impair substrate efflux could be associated with chemoresistance in TNBC patients undergoing chemotherapy and designed this study to test this hypothesis.

Methodology

Blood samples from 76 Malaysian TNBC patients who had undergone Taxane-Adriamycin-Cyclophosphamide (TAC) chemotherapy regimen were collected and stored in EDTA tube. DNA was extracted from these blood samples and genotyped using PCR-RFLP technique. The genotypes were categorized into homozygous

wildtype, heterozygous and homozygous variant based on the band sizes on gel electrophoresis. The difference in genotype frequencies between chemoresistant and chemoresponsive groups was determined by using crosstabs analysis. The association between the genotypes and alleles with TAC chemotherapy response was determined by using binary logistic regression analysis deriving Odds Ratio (OR) with confidence interval (CI) of 95% on SPSS version 20.2. Haplotype frequencies between the chemoresistant and chemoresponsive groups of TNBC patients and their association with treatment response was calculated using Haploview software v4.2.

Results

The heterozygous (GA) and homozygous variant (AA) genotypes of the *ABCG2* G34A polymorphism showed a significantly lower risk association with chemoresistance [(OR= 0.303, p=0.029) and (OR= 0.151, p=0.011) respectively]. Whereas the heterozygous (CA) and homozygous variant (AA) genotypes of C421A showed statistically insignificant low risk association with chemoresistance, [OR=0.481, p=0.251; OR=0.412, p=0.113 respectively]. The variant A allele of both SNPs [with OR=0.320, p=0.002 and OR=0.487, p=0.039 respectively] and haplotypes GA, CA and AA [with ORs=0.020, 0.002 and 0.00004, respectively] also showed significantly lower risk association with chemoresistance. It is reasonable to suggest that the variant genotype (AA) and variant allele (A) of G34A and C421A decrease the expression and transporter activity of *ABCG2*, lower the efflux activity of chemotherapeutic drugs, resulting in increased intracellular accumulation of the

drug and thereby improve the efficacy of chemotherapy and lower the risk of chemoresistance.

Conclusion

These findings suggest that the *ABCG2* G34A polymorphism may be useful as a potential biomarker to predict chemotherapy response in TNBC patients.

Keywords: *ABCG2*, TNBC, G34A, C421A, polymorphisms, TAC chemotherapy.

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Breast cancer is one of the most common cancers in the world ranking in second globally with an estimated one million newly diagnosed cases and with more than 400,000 deaths globally[1]. It is the most common cancer among women globally and even locally. The number of cases in Malaysia according to data from the Malaysian National Cancer Registry (MNCR) report compiled from 2007 to 2011 is estimated to be 18,343 cases, constituting 17.7% of all total cancer cases in Malaysia with the incidence of 31.1% per 100,000 population [The Malaysian National Cancer Registry Report (MNCR) 2007-2011, 2016]. This statistic varies among the three major ethnic groups in Malaysian population, where Malaysian Chinese and Malaysian Indians show an almost similar incidence of 1 in 19 women while Malays show the lowest incidence of 1 in 28 women. However, the overall survival rate is in the opposite manner due to the latter group presenting in later stages with higher tumor burden than the former groups[1].

Breast cancer can be classified according to its histology. According to the latest WHO classification released in 2012, breast cancer can be histologically divided into invasive carcinoma of no special type (NOS), invasive lobular carcinoma, tubular carcinoma, mucinous carcinoma, carcinoma with medullary features, cribriform carcinoma, and a few other rare categories[2].

During pathological assessment of a breast cancer tissue sample, apart from the histological type, histological grading and immunohistochemistry evaluations are also done to decide on the management as well as prognosis[3]. The immunohistochemical tests commonly done are estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2) and Ki67. While Ki67 is a proliferation indicator, the other three hormone receptors predict the responsiveness of the cancer to selective estrogen receptor modulator (SERM) such as Tamoxifen and immune-targeted drug Herceptin which indirectly predicts the prognosis[3]. All these parameters along with other parameters such as the stage of the cancer are compiled and given a score according to the Nottingham Prognostic Index (NPI) to give an estimation of the probability of death from breast cancer or recurrence[3].

Triple-Negative Breast Cancer (TNBC), a subtype of breast cancer, is characterized by absence of all three hormone receptors, namely the ER, PR and HER2. The incidence of TNBC among breast cancer patients in the world population is estimated to be around 10-20% of annually diagnosed breast cancer cases and the incidence in Malaysia stands at 17.6%[4].TNBC is known to recur and metastasize widely at diagnosis. The overall survival rate is also poor because of the limited choice of treatment available which is only chemotherapy[3], due to limited use of SERM and other hormonal therapy. Due to the absence of hormone receptors, there are no targeted drugs available for TNBC. Chemotherapy using taxane-adriamycin-cyclophosphamide (TAC) regimens represent the mainstay in TNBC therapy. Still, due to inter-individual variability,

a significant proportion of TNBC patients on TAC chemotherapy become resistant or intrinsically less susceptible while only a subset of TNBCs are sensitive to chemotherapy. The differences in clinical outcomes following chemotherapy imply that a subset of TNBC patients are sensitive to chemotherapy while a major subset become resistant and are intrinsically less susceptible. Those TNBC patients who do not achieve good response suffer an early recurrence due to development of resistance to chemotherapy and die from metastatic disease. It is one of the most aggressive subtypes of breast cancer with high risk of chemo-resistance and hence, treatment failure. So, elucidating the mechanism behind chemoresistance and developing a cost-effective predictive marker for better therapeutic response with TNBC patients is the need of the hour[5]. A predictive factor is any measurable characteristic associated with a response or lack of response to a specific treatment[6].

Numerous mechanisms can lead to the development of chemoresistance. The mechanisms that promote direct or indirect resistance against anti-cancer drugs include drug target alteration, drug inactivation, inhibition of apoptosis, DNA damage repair, alterations in drug efflux transport, epigenetic alterations and epithelial mesenchymal transition (EMT) [7]. Poor pharmacology is a major type of resistance although poor pharmacology can sometimes be turned into an advantage in therapy. Transporter mediated drug efflux, specifically a defective transport has been implicated as a major mechanism of anticancer drug resistance[8]. Adenosine triphosphate binding cassette (ABC) transporters are efflux transporters, which extrude drugs often against a concentration gradient.

The increased expression of ABC transporters on plasma membranes will result in increased efflux and decreased intracellular accumulation of anticancer drugs and this can lead to development of anticancer drug resistance. On the contrary, decreased expression of ABC transporters will result in decreased efflux and increased intra-cellular accumulation of anticancer drugs resulting in better response to the drug. As ABC transporters influence the pharmacokinetics and intracellular or systemic level of anticancer drugs, genetic variations such as single nucleotide polymorphisms (SNPs) of these transporter genes could be potential determinants of variability in drug disposition and efficacy[9].

ABCG2, the second member of the G family of ABC transporters, is located on chromosomal region 4q22 and encodes the breast cancer resistance protein (BCRP). ABCG2 has been found to facilitate the efflux of a variety of anti-cancer agents such as Adriamycin, Daunorubicin, 7-ethyl-10-hydroxycamptothecin, Topotecan, and Mitoxantrone, mediating multidrug resistance[10]. BCRP expression and function can be altered by SNPs in ABCG2 gene. Genetic variations in *ABCG2* are closely related to inter-individual variations in therapeutic performance. Around 80 variants have been found to exist in the ABCG2 gene across many ethnic groups[11]. Out of the many variants found under *ABCG2*, the two most common variants found in Asian population are the 34G>A [12Val>Met in exon 2, (rs2231137)] and 421C>A [141Gly>Lys in exon 5, (rs2231142)], seen in 19.3% and 28.9% of the population, respectively[12,13].These two SNPs are associated with decreased expression and thus reduced transporter activity of the ABCG2 protein[14]. Therefore, *ABCG2*

SNPs may be associated with inter-individual variability in drug response to anti-cancer agents and clinical outcome.

Interestingly, out of these two variants, the latter has been shown in few studies to have an association with increased bioavailability of some orally administered substrate drugs and lower susceptibility to cancer risk which are its desirable effects[12,15]. While some studies on the former variant, showed the opposite association[16–18].

Studies investigating the association between the different polymorphisms in the ABCG2 gene have outlined the possible association, both in terms of clinical outcome and treatment outcome. For example, a study by Wu *et al*., (2015) demonstrated the correlation between the different genotypes of these two polymorphisms and other parameters such as family history of breast cancer, clinical stage at diagnosis, histological type, tumor size at diagnosis, hormone receptor status, *BRCA* mutation status, *p53* mutation status, and also lymph node metastasis[16]. Two parameters (clinical stage and ER/PR statuses) showed a statistically significant coherent agreement that 421C>A is associated with better outcome compared to 34G>A[16]. The other parameters did not give any significant finding[16]. In this study by Wu *et al* on Chinese breast cancer patients, the 421 AA genotype showed the best outcome (statistically significant) in terms of clinical response to neo-adjuvant therapy as well as disease progression as compared to other genotypes. The variant AA genotype of 34G>A on the other

hand, showed the worst outcome in treatment response and disease progression although not statistically significant[16].

The allelic frequencies of *ABCG2* polymorphisms vary widely among ethnic groups. Few studies had investigated the association between the different polymorphisms in *ABCG2* in terms of treatment response and clinical outcome[12,16,19]. However, there are no reports available on the frequencies and the impact of genetic variations in ABCG2 genes on TAC chemotherapy response in Malaysian TNBC patients, as no previous studies have been undertaken. In this regard, use of pharmacogenetics to identify patients who possess good response or resistance marker, before treatment will be a good option. So it was of interest to assess the potential impact of inter-individual variations in chemotherapy efficacy due to these two *ABCG2* polymorphisms in TNBC patients.

Hypothesis: There is an association between the two SNPs, G34A and C421A of ABCG2 gene with variation in TAC chemotherapy response in Malaysian TNBC patients.

To test this hypothesis, this study was designed with the objectives shown in chapter 2.

CHAPTER 2: OBJECTIVES OF THE STUDY

2.1. General objective

To investigate the frequencies of alleles, genotypes and haplotypes of *ABCG2* SNPs, G34A and C421A in TNBC patients undergoing TAC chemotherapy and to determine their potential association with TAC chemotherapy response.

2.2. Specific objectives

1. To investigate the allele, genotype and haplotype frequencies of *ABCG2* 34G>A and 421C>A polymorphisms in TNBC patients undergoing TAC chemotherapy.
2. To determine the association of alleles, genotypes and haplotypes of the above SNPs in TNBC patients with TAC chemotherapy response.
3. To evaluate whether the SNP data generated could serve as predictive biomarkers of chemotherapy response in TNBC patients.

CHAPTER 3: MANUSCRIPT

3.1 TITLE

GENETIC ASSOCIATION OF ABCG2 GENE POLYMORPHISMS WITH TAC CHEMOTHERAPY RESPONSE IN TRIPLE NEGATIVE BREAST CANCER (TNBC) PATIENTS.

ROSHAIDIE RASHID[1], MD SALZIHAN MD SALLEH[2], AHMAD AIZAT ABDUL AZIZ [1], and RAVINDRAN ANKATHIL1[1]

[1]Human Genome Centre, and [2]Department of Pathology, School of Medical Sciences, Universiti Sains Malaysia Health Campus, 16150, Kubang Kerian, Kelantan, Malaysia

*For correspondence, e-mail: rankathil@hotmail.com

3.2 ABSTRACT

ABSTRACT

GENETIC ASSOCIATION OF ABCG2 GENE POLYMORPHISMS WITH TAC CHEMOTHERAPY RESPONSE IN TRIPLE NEGATIVE BREAST CANCER (TNBC) PATIENTS.

Introduction

Triple negative breast cancer which is one of the most aggressive subtypes of breast cancer with high risk of recurrence, metastasis, and chemoresistance, is associated with inter-individual variability in treatment response. Transmembrane transport proteins such as breast cancer resistance protein (BCRP) encoded by *ABCG2* can cause chemoresistance by means of increased efflux transportation of the drug out of the cell, hindering its molecular action on cancer cells. We hypothesized that, the genetic variations G34A and C421A of ABCG2 gene that impair substrate efflux could be associated with chemoresistance in TNBC patients undergoing chemotherapy and designed this study to test this hypothesis.

Methodology

Blood samples from 76 Malaysian TNBC patients who had undergone chemotherapy were collected and stored in EDTA tube. DNA was extracted from these blood samples and genotyped using PCR-RFLP technique. The association between the genotypes and alleles with TAC chemotherapy

response was determined using binary logistic regression analysis deriving Odds Ratio (OR) with confidence interval (CI) of 95% on SPSS version 20.2. Haplotype frequencies were determined between the chemoresistant and chemoresponsive groups and haplotypes with >0.03 % were presented. The association between haplotypes and chemotherapy response was calculated using Haploview software v4.2.

Results

Heterozygous (GA) and homozygous variant (AA) genotypes of the *ABCG2* G34A polymorphism showed a significantly lower risk association with chemoresistance [(OR= 0.303, p=0.029) and (OR= 0.151, p=0.011) respectively]. Whereas the heterozygous (CA) and homozygous variant (AA) genotypes of C421A showed statistically insignificant low risk association with chemoresistance, [OR=0.481, p=0.251; OR=0.412, p=0.113 respectively]. The variant A allele of both SNPs [with OR=0.320, p=0.002 and OR=0.487, p=0.039 respectively] and haplotypes GA, CA and AA [with ORs=0.020, 0.002 and 0.00004, respectively] also showed significantly lower risk association with chemoresistance.

Conclusion

Our findings suggest that the *ABCG2* G34A polymorphism may be useful as a potential biomarker to predict chemotherapy response in TNBC patients.

Keywords: *ABCG2*, TNBC, G34A, C421A, polymorphisms, TAC chemotherapy.

3.3 INTRODUCTION

Triple-Negative Breast Cancer (TNBC), a subtype of breast cancer, is characterized by absence of all three hormone receptors, namely the estrogen receptor, progesterone receptors and human epidermal growth factor receptor 2 (HER2). The incidence of TNBC among breast cancer patients in the world population is estimated to be around 10-20% of annually diagnosed breast cancer cases and the incidence in Malaysia stands at 17.6%.[1] It is one of the most aggressive subtypes of breast cancer with high risk of chemo-resistance and hence, treatment failure. TNBC is also known to recur and metastasize widely at diagnosis. The overall survival rate is also poor because of the limited choice of treatment available which is only chemotherapy,[2] due to limited use of SERM and other hormonal therapy. Chemotherapy using taxane-adriamycin-cyclophosphamide (TAC) regimens represent the mainstay in TNBC therapy. Still, due to inter-individual variability, a significant proportion of TNBC patients on TAC chemotherapy become resistant or intrinsically less susceptible while only a subset of TNBCs are sensitive to chemotherapy. Those TNBC patients who do not achieve good response suffer an early recurrence due to development of resistance to chemotherapy and die from metastatic disease. So, elucidating the mechanism behind chemoresistance and developing a cost-effective predictive marker for better therapeutic response with TNBC patients is the need of the hour.[3]

The mechanisms that promote direct or indirect resistance against anti-cancer drugs include drug target alteration, drug inactivation, inhibition of apoptosis, DNA damage repair, alterations in drug efflux transport, epigenetic alterations

and epithelial mesenchymal transition (EMT).[4] Adenosine triphosphate binding cassette (ABC) transporters are efflux transporters which extrude drugs often against a concentration gradient. The increased expression of ABC transporters on plasma membranes will result in increased efflux and decreased intracellular accumulation of anticancer drugs and this can lead to development of anticancer drug resistance. As ABC transporters influence the pharmacokinetics and intracellular or systemic level of anticancer drugs, genetic variations such as single nucleotide polymorphisms (SNPs) of these transporter genes could be potential determinants of variability in drug disposition and efficacy.[5]

ABCG2, the second member of the G family of ABC transporters, is located on chromosomal region 4q22 and encodes the breast cancer resistance protein (BCRP). *ABCG2* has been found to facilitate the efflux of a variety of anti-cancer agents such as Adriamycin, Daunorubicin, 7-ethyl-10-hydroxycamptothecin, Topotecan, and Mitoxantrone, mediating multidrug resistance.[6] BCRP expression and function can be altered by SNPs in ABCG2 gene. Out of the many variants found under *ABCG2*, the two most common variants found in Asian population are the 34G>A (12Val>Met in exon 2 rs2231137) and 421C>A (141Gly>Lys in exon 5 rs2231142), seen in 19.3% and 28.9% of the population, respectively.[7,8] These two SNPs are associated with decreased expression and then reduced transporter activity of the *ABCG2* protein.[9] Therefore, *ABCG2* SNPs may be associated with inter-individual variability in drug response to anti-cancer agents and clinical outcome.

The allelic frequencies of *ABCG2* polymorphisms vary widely among ethnic groups. Few studies had investigated the association between the different polymorphisms in *ABCG2* in terms of treatment response and clinical outcome.[7,10,11] However, there are no reports available on the frequencies and the impact of genetic variations in *ABCG2* on TAC chemotherapy response in Malaysian TNBC patients, as no previous studies have been undertaken. The potential inter-individual variations in efficacy due to *ABCG2* polymorphisms must be assessed for precision medicine. It was hypothesized that there is an association between the two SNPs, G34A and C421A of ABCG2 gene with variation in chemotherapy response in TNBC patients. To test this hypothesis, this study was designed to investigate the frequencies of alleles, genotypes and haplotypes of *ABCG*2 SNPs, G34A and C421A in TNBC patients undergoing TAC chemotherapy and to determine their potential association with chemotherapy response.

3.4 METHODOLOGY

Study subjects

This study was approved by the Research Review Board and Ethics Committee of Universiti Sains Malaysia (USMKK/PPP/JEPeM[260.39210]) and Ministry of Health Malaysia (NMRR-15-1200-25230) which complies with Declaration of Helsinki. The study subjects were recruited from Hospital Raja Perempuan Zainab II (HRPZII) and Hospital USM (HUSM) while experimental analyses were done at Human Genome Centre, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan.

TNBC patients who were histologically confirmed, had undergone mastectomy and had completed six cycles of taxane-adriamycin-cyclophosphamide (TAC) chemotherapy were recruited in this study. Peripheral blood samples were collected from 76 TNBC patients after getting informed consent. Clinicopathological data of the patients such as age, histological subtype and grade, stage and lymph node status were also recorded.

Evaluation of treatment response

TNBC patients who had undergone mastectomy and six cycles of TAC chemotherapy regimen were evaluated after one year. Those patients who developed disease progression, local recurrence or distant metastases which were based on radiographic findings such as CT and MRI scan, were categorized into chemoresistant group. Those patients who did not develop any of the above signs were categorized into chemoresponsive group.

DNA extraction and SNPs genotyping

Using the 3.0 ml whole blood sample collected in EDTA tube, DNA was extracted using QIAamp DNA mini kit (Qaigen, Hilden, Germany). Polymerase Chain Reaction- Restriction Fragment Length Polymorphism (PCR-RFLP) technique was employed to determine the genotype using the primers obtained from Primer-BLAST Analysis. The forward and reverse primer sequences for ABCG2 G34A were 5'- CAGTAATGTCGAAGTTTTTATCGCA-3' and 5'-AAATGTTCATAGCCAGTTTCTTGGA-3' respectively. While for C421A, the

forward and reverse primer sequences were 5'-GTTGTGATGGGCACTCTGATGGT-3' and 5'- CAAGCCACTTTTCTCATTGTT-3' respectively. Each PCR mixture consisted of 1x MyTaq Reaction buffer (Bioline Ltd, London, UK), 1 unit of MyTaq DNA Polymerase (Bioline ltd, London, UK), 50 ng genomic templates and ddH2O in a total volume of 20.0 µl. PCR conditions involved denaturation at 95^{o} C for 1 min, and repeated 35 cycles consisting of 3 steps: denaturation at 95^{o} C for 15 seconds,; annealing at 56°C and 58°C for C421A and G34A for 15 seconds, and extension at 72^{o} C for 10 seconds, followed by 3 minutes of final extension at 72^{o} C. The PCR products were subsequently digested with restriction enzymes Taal for 10 minutes at 65^{o} C (C421A) and BseMI for 20 minutes at 65 ^{o} C (G34A). Digested products were then electrophoresed on 3% agarose gel, stained with SyBr green and viewed under UV transilluminator imaging system (Alpha DigiDoc™, USA). Based on the different bands, the genotypes were categorized into homozygous wildtype, heterozygous and homozygous variant. The different genotype patterns observed for G34A and C421A polymorphisms of ABCG2 gene are shown in Figure 1 & 2 respectively. Following genotyping, 10% of samples from each different genotype were randomly selected for sequencing to confirm the expected sequences of each genotype. The selected PCR products were purified by using a QIAquick PCR purification kit (QIAGEN) before sending them to First Base Laboratories, Kuala Lumpur for sequencing. Representative sequencing results are shown in Figures 3 and 4.

Statistical analysis

The difference in genotype and allele frequencies of the SNPs among the chemoresistant and chemoresponsive groups of TNBC patients was determined using chi-square test ($\chi 2$). The association between genotypes and alleles with chemotherapy response was determined using binary logistic regression analysis deriving Odds Ratio (OR) with confidence interval of 95%. All these statistical tests were carried by using SPSS v24 (SPSS, Chicago, USA). P–values less than 0.05 were considered as statistically significant. Haplotype frequencies were determined between the chemoresistant and chemoresponsive groups using LD block analysis of the two SNPs, expanded 8kb, as shown in Figure 5 and haplotypes with >0.03 % were presented. Both investigated SNPs showed low linkage disequilibrium ($r^2<1$). The association between haplotypes and chemotherapy response was calculated using Haploview software v4.2 (Broad Institute Cambridge, USA).

3.5 RESULTS

3.5.1 *Clinicopathological data of TNBC patients*

For this study, a total of 76 histopathologically confirmed TNBC patients were recruited. The mean age of study subjects at diagnosis was 48.9±9.67 years. Clinicopathological data of the TNBC patients included are shown in Table 1. Based on chemotherapy response, 47 patients (61.8%) were chemotherapy responders while 29 (38.2%) were chemotherapy resistant. The clinicopathological data of these 76 patients are shown in Table 1.

3.5.2 *Genetic association of ABCG2 polymorphisms with chemotherapy response*

Table 2 shows the genotype and allele frequencies of *ABCG2* G34A and C421A polymorphisms in chemoresistant and chemoresponsive groups of TNBC patients. For *ABCG2* G34A polymorphism, the frequency of homozygous wildtype (GG) was significantly higher in chemoresistant group (58.6%) than in chemoresponder group (25.5%) (P = 0.004). The heterozygous (GA) and homozygous variant (AA) genotypes were more frequent in chemoresponsive group compared to chemoresistant group, but the difference was not statistically significant. There was also a significant difference in frequency of variant A allele which was significantly higher (p=0.001) in the chemoresponsive group (52.1%) than in the chemoresistant group (25.9%). For *ABCG2* C421A polymorphism, the frequency of CC genotype was higher in chemoresistant group whereas the frequencies of CA and AA genotypes were higher in the chemoresponsive group. But the difference in frequency was not statistically significant. However, the frequency of variant A allele was significantly higher (P=0.037) in the chemoresponsive group (50%) than in the chemoresistant group (32.8%).

Binary logistic regression analysis was performed to determine the association of these SNPs with chemotherapy response as shown in Table 2. For G34A polymorphism, both heterozygous (GA) and homozygous variant (AA) showed significantly lower risk (protective) association (P = 0.029 and 0.011 respectively) against chemoresistance with odds ratios (ORs) of 0.303 and 0.151, respectively. The A allele of G34A polymorphism also showed a significant lower

risk (protective) association (P = 0.002) with chemoresistance with odds ratio (OR) of 0.320. For C421A polymorphism, both heterozygous (CA) and homozygous variant (AA) genotypes showed lower risk association with chemoresistance, but was statistically insignificant. However, A allele of C421A polymorphism showed a significant lower (protective) risk association against chemoresistance (OR: 0.487, p=0.039).

3.5.3 *Haplotype analysis*

Haplotype analysis as per Table 3 showed that the frequency of *ABCG2* haplotype 34A/421A was significantly higher in the chemoresponsive group (21.8%) than in chemoresistant group (7.3%) (P = 0.018). The association analysis of these three haplotypes (GA, CA and AA) with chemoresistance showed significantly lower risk (protective) against chemoresistance with OR; 0.363, 0.245, 0.137 and p=0.02, 0.002 and 0.00004 respectively.

3.6 DISCUSSION

TNBC is an aggressive subtype of breast cancer plagued with recurrence due to drug resistance, treatment failure, low survival rate, and thus worse clinical outcome. Conventional prognostic markers such as patient's age, tumor size, and lymph node involvement are rarely associated with recurrence and distant metastasis in TNBC patients.[12] Furthermore, no light has been shed on the exact mechanism of drug resistance seen in TNBC, further complicating its management. Although several studies have hypothesized several mechanisms

behind chemoresistance in TNBC patients, the exact mechanism still remains unclear. It is plausible that resistant genotypes or genes prone to being cancer-treatment-resistant might be pre-existing in the host and that adaptively gets selected by the initial treatment. Genetic variations in *ABCG2* are closely related to inter-individual variations in therapeutic performance.[13] As *ABCG2* is a major efflux transporter responsible for multidrug resistance in cancer cells,[14] this study aimed to investigate the impact of variations in *ABCG2* as an altered drug efflux transport mechanism mediating chemoresistance and to identify potential biomarkers that could predict clinical outcome in TNBC patients. To the best of available knowledge, this is the first study to investigate the association of *ABCG2* SNPs G34A and C421A with TAC chemotherapy response in Malaysian TNBC patients.

The mean age of diagnosis of TNBC among Malaysian population is 48.9+-9.67 years which is fairly similar to other cohorts such as in one study on Turkish breast cancer patients with 52.4±12.5 years.[15] Out of 76 TNBC patients recruited in this study, 29 (38.2%) were resistant to TAC chemotherapy regime. Qiu *et al* (2016) reported a recurrence frequency of 27.9% in Chinese TNBC patients compared to 13.38% in non-TNBC patients.[16] Another study on Chinese TNBC patients by Wu *et al* showed a slightly higher distribution of patients (55%) in the resistant group after anthracycline-based neoadjuvant chemotherapy regime.[11] All these reports indicate a higher rate of recurrence in TNBC patients, which is in agreement with our findings.

In the present study, the *ABCG2* 34GG wild type genotype was significantly higher (p= 0.004) in the chemoresistant group of TNBC patients. In a study on Malaysian CML patients undergoing Imatinib treatment by Au *et al*, there was no significant difference in the distribution of the homozygous wildtype (GG) between Imatinib responsive and resistant groups of CML patients.[17] In terms of association of G34A SNP with chemotherapy response, homozygous variant (AA), heterozygous (GA) genotypes and variant A allele of *ABCG2* G34A showed a significantly lower risk association (protective) with chemotherapy resistance (OR: 0.303, 0.151 and p=0.029, 0.011 respectively). Also, the variant A allele of G34A polymorphism, similar to its counterpart variant A allele in C421A polymorphism, showed a protective effect against chemoresistance as compared to the respective wildtype G and C alleles (OR:0.320, p=0.002). TNBC patients harboring these genotypes are more likely to experience favorable response to TAC chemotherapy and less likely to develop resistance against TAC chemotherapy regime.

There were not many previous studies associating *ABCG2* G34A polymorphism with chemoresistance in TNBC patients. But there were a few studies testing the association of this polymorphism with other cancers in terms of susceptibility, treatment response or survival. One study by Tandia *et al* reported that heterozygous (GA) genotype of the G34A polymorphism was significantly associated with the lowest Sorafenib plasma level indicating a possible better therapeutic response in hepatocellular carcinoma (HCC) patients.[18]

Another study by Au *et al* also showed a similar protective effect association of the heterozygous (GA) and homozygous variant (AA) genotypes of G34A polymorphism against IM resistance in CML patients albeit being not statistically significant.[17] In that study, heterozygous (GA) and homozygous variant (AA) of G34A were also associated with good cytogenetic and molecular responses in CML patients undergoing Imatinib treatment.

With regard to cancer susceptibility, a study by Wu *et al*, reported the homozygous variant (AA) and heterozygous (GA) genotypes of G34A to be associated with increased risk of developing hormone receptor positive (ER and PR) breast cancer.[11] These authors also reported that hormone positive breast cancer patients with *ABCG2* 34AA genotype and undergoing chemotherapy had a worse outcome, which is contradictory with the present study findings on TNBC patients. However, in terms of overall survival (OS) of breast cancer patients with positive ER and PR, those with homozygous variant (AA) genotype showed longer OS, which is in agreement with our findings on TNBC patients.

In the present study, the homozygous wildtype (CC) genotype of *ABCG2* C421A was also more frequently found in chemoresistant patients than in chemoresponsive patients although the difference was not significant (p=0.084). However, the A allele of this SNP showed a significantly higher distribution among the chemoresponsive group than the chemoresistant group (p=0.037) and the variant A allele of *ABCG2* C421 was associated with a lower risk (protective effect) against chemotherapy resistance (OR: 0.487, p=0.037). This

is in agreement with the findings by Wu *et al* on Chinese breast cancer patients[11] and Ghafouri *et al* on Kurdish breast cancer patients[19] undergoing chemotherapy which stated that the risk of resistance correlated with the wildtype C allele. These authors reported that breast cancer patients who carried A allele of C421A had complete response whereas carriers of C allele showed weaker response to treatment with anthracyclines and paclitaxel. In our study, TNBC patients carrying the heterozygous CA and homozygous variant AA genotypes of C421A showed lower risk against chemoresistance with OR<1.0. But it was not statistically significant. Lack of significant association could be due to the unequal distribution in the number of patients included in the two groups. The studies by Wu *et al* and Ghafouri *et al* also showed that the wild type CC genotype was more frequent in chemoresistant group while homozygous variant (AA) was more frequent in the chemoresponsive group.[11,19] In these studies, the homozygous variant AA of C421A was associated with good response to anthracycline based chemotherapy, a view coherent with our findings in the present study. Chinese breast cancer patients with *ABCG2* 421AA genotype were reported to have better therapeutic response , best outcome and longer survival,[11] which is in agreement our findings.

There are not many reports available on TNBC patients to compare with and authenticate our results. However, few other earlier studies investigated the impact of this polymorphism on treatment response in CML patients undergoing Imatinib therapy. Au *et al* and Jiang *et al* also showed that the A allele of *ABCG2* C421A polymorphism was associated with favourable treatment response and major molecular response (MMR) respectively in CML patients which is in

agreement with the present study.[17,20] A meta-analysis by Jiang *et al* revealed that the variant allele A of *ABCG2* C421A was significantly associated with a higher rate of MMR and overall response especially in Asian CML patients undergoing Imatinib treatment.[20] On the contrary, in another study on lung cancer patients which focused on 6 SNPs including the two SNPs of interest in this present study, Mueller *et al* (2009) reported that A allele carrier of the C421A polymorphism treated with platinum-based chemotherapy was associated with worse overall survival (OS).[21]

Genetic polymorphism in *ABCG2* may alter the expression of transporter proteins and affect the drug metabolic disposition. Wild type genotype and allele would render increased efflux transporter efficiencies whereas the variant genotypes and alleles would render reduced chemotherapy drug efflux.[22] In G34A polymorphism, the amino acid changes from valine to methionine whereas in C421A polymorphism, the change of amino acid is from glutamine to lysine. These genetic alterations could lead to a decreased protein expression that might render the sensitivity to a certain drug and vice versa.[10,11,23,24] It is possible that the efflux activity of the BCRP might be altered in such a way that the mutant BCRPs transport out chemotherapy drugs lesser than the wildtype.[11,24] This could result in increased intracellular drug accumulation because of reduced chemotherapy drug efflux and better treatment response.[22] In other words, variant genotype (AA) of G34A and C421A show lower expression of BCRP proteins, leading to low level of drug resistance. *ABCG2* G34A or C421A variant alleles decrease the expression and the transporter activity of *ABCG2* protein and thereby lower the efflux activity of

substrates such as chemotherapy drugs.[10,24–26] This will result in accumulation of chemotherapy drugs for prolonged time inside the malignant cells. Subsequently, this will improve the efficacy of chemotherapy in patients with variant genotypes of ABCG2 gene, thus posing a lower risk for resistance and thereby resulting in a better response.

Haplotype analysis of the two SNPs also showed a significantly higher frequency of AA haplotypes in chemoresponsive patients (p=0.018) which is in agreement with the findings by Au *et al* on Malaysian CML patients.[17] In our study, the frequencies of CA and GA haplotypes were not significantly different among the two groups of TNBC patients. When associating these haplotypes with chemotherapy response, all three, GA, CA and AA haplotypes showed significantly lower risk for chemoresistance with OR: 0.363, 0.245 and 0.137 and p=0.02, 0.002 and 0.00004 respectively. Hence, TNBC patients carrying these haplotypes might show better response to treatment. This is in agreement with the result of haplotype association with IM response in CML patients by Au *et al* that showed AA haplotype was associated with good IM response.[17]

Our findings are contradictory with few other studies. Possible explanation for the inconsistent findings of association between variations in and chemoresistance across other studies might be due to the differences in frequencies of these SNPs in other populations, genetic background of the study subjects and the treatment employed. Likewise, the interaction between various other genes and proteins involved in this mechanism might also be contributing, which is consistent with current understanding of chemotherapy resistance.[27]