ROLE OF GENETIC VARIATIONS OF ABC (*ABCB1*, *ABCC1* AND *ABCG2*) ON TRIPLE NEGATIVE BREAST CANCER SUSCEPTIBILITY RISK

YEOH HAO ING

UNIVERSITI SAINS MALAYSIA

2024

ROLE OF GENETIC VARIATIONS OF ABC (*ABCB1*, *ABCC1* AND *ABCG2*) ON TRIPLE NEGATIVE BREAST CANCER SUSCEPTIBILITY RISK

by

YEOH HAO ING

Thesis submitted in fulfilment of the requirements
for the Degree of
Master of Science

September 2024

ACKNOWLEDGEMENT

First and foremost, I would like to extend my deepest appreciation to my supervisor, Dr Ahmad Aizat, for his guidance, expertise, and patience throughout this endeavour. Your mentorship has not only shaped my academic and professional growth but has also inspired me to strive for excellence. To my esteemed lab members, Farah, Mimi, Elyn, Fasihah, Kak Sha, Kak Ina, Dr Izzati, Dr Ridzuan, Dr Hidayati and Dr Durar, your support and encouragement have been a constant source of inspiration and joy. Your friendship has provided me courage and strength during the most challenging times, and I am truly blessed to have you in my life. To the patients who participated in my research in HUSM and the support staff, I am greatly grateful for your willingness to contribute to the advancement of knowledge in our field. Your participation is invaluable, and I am humbled by your trust and cooperation. I am also deeply grateful to our HGC lecturers, Prof Zilfalil, Assoc Prof Dr Sarina Sulong, Dr Marahaini, Dr Nazihah, Dr Nurfadhlina and Dr Aizati, for your valuable comments and discussion. To the support staff of HGC, En Zaki, En Md Ros, En Ariff, Kak Nor, Kak Ann, Kak Nisah and whose I left out the name, I am truly thankful for your continuous support. To my family, whose unconditional love, encouragement, and understanding have been my source of strength and motivation, I owe a debt of gratitude. Your unwavering belief in my abilities has propelled me forward during both the triumphs and trials of this journey. Finally, I would like to dedicate this acknowledgment to my late father, whose unwavering trust and unconditional love supported me along every step of my life's journey. Though you are no longer with us, your love continues to resonate in every milestone I achieve. To each of you, I offer my deepest appreciation for your support, encouragement, and belief in me. This accomplishment would not have been possible without your unwavering support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\chi 2$	:	Chi Square
^{o}C	:	Degree Celsius
=	:	Equal to
<	:	Less than
≤	:	Less than or equal to
>	:	More than
%	:	Percentage
±	:	Plus-minus
+	:	Positive
-	:	Negative

LIST OF ABBREVIATIONS

µL	:	Microlitre
µM	:	Micromolar
ABC	:	ATP-binding cassette
ABCA	:	ATP binding cassette subfamily A
ABCB1	:	ATP binding cassette subfamily B member 1
ABCC1	:	ATP binding cassette subfamily C member 1
ABCD	:	ATP binding cassette subfamily D
ABCF	:	ATP binding cassette subfamily F
ABCG2	:	ATP binding cassette subfamily G member 2
AE	:	Elution buffer
AJCC	:	American Joint Committee on Cancer
ARMS	:	Amplification refractory mutation system
ASCO	:	American Society of Clinical Oncology
ASR	:	Age-standardised rate
ATP	:	Adenosine triphosphate
AW1	:	Washing buffer 1
AW2	:	Washing buffer 2
BARD1	:	BRCA1 Associated Ring Domain 1
BCS	:	Breast-conserving surgery
BI-RADS	:	Breast imaging-reporting and data system
BL1	:	Basal-like 1
BL2	:	Basal-like 2
bp	:	Base pair
BRCA1/2	:	Breast cancer 1/2 gene
BRIP1	:	BRCA1 interacting protein
CDK4/6	:	Cyclin-dependent kinase 4 and 6
CI	:	Confidence interval
CSC	:	Cancer stem cells
DBT	:	Digital breast tomosynthesis
DCIS	:	Ductal carcinoma in situ
df	:	Degree of freedom
DFI	:	Disease-free interval
DFS	:	Disease free survival

DNA	:	Deoxyribonucleic acid
EBCTCG	:	Early Breast Cancer Trialists' Collaborative Group
ECIBC	:	European Commission Initiative on Breast Cancer
ECM	:	Extracellular matrix
EDTA	:	Ethylenediaminetetraacetic acid
ER	:	Estrogen receptor
ERBT	:	External beam radiation therapy
ESR1	:	Estrogen Receptor 1
FANCM	:	Fanconi anemia complementation group M
FEC	:	Fluorouracil, epirubicin hydrochloride, and cyclophosphamide
g	:	Gram
GLOBOCAN	:	Global Cancer Observatory
GPR55	:	G protein-coupled receptor 55
H&E	:	Hematoxylin and eosin stain
HER2	:	Human epidermal growth factor receptor 2
HR	:	Hazard ratio
HR+	:	Hormonal receptor positive
HRM	:	High-resolution melting analysis
ICI	:	Immune checkpoint inhibitor
IDC	:	Invasive ductal carcinoma
IHC	:	Immunohistochemistry
ILC	:	Invasive lobular carcinoma
IM	:	Immunomodulatory
ISH	:	in situ hybridization
ITC	:	Isolated tumour cells
kb	:	Kilobase
kDa	:	Kilodalton
LAR	:	Luminal androgen receptor
LD	:	Linkage disequilibrium
M	:	Mesenchymal
mA	:	Milliampere
MDR1	:	Multidrug resistance protein 1
min	:	Minute
mL	:	Millilitre

MOH	:	Ministry of Health
MRI	:	Magnetic resonance imaging
MRP1	:	Multidrug resistance protein 1
MSL	:	Mesenchymal stem–like
N	:	Number of individuals
NA	:	No available data
NAFLD	:	Non-Alcoholic Fatty Liver Disease
ng	:	Nanogram
OR	:	Odds ratio
ORR	:	Objective response rate
OS	:	Overall survival
P_0	:	Frequency of a variant allele in control
P_1	:	Expected variant frequency in the study population
PALB	:	Partner and localizer of BRCA
PARP	:	Poly (ADP-ribose) polymerases
PCR	:	Polymerase chain reaction
pCR	:	Pathologic complete response
PD-1	:	Programmed cell death-1
PD-L1	:	Programmed cell death ligand 1
PFS	:	Progression-free survival
PGP	:	P-glycoprotein
PIK3CA	:	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PR	:	Progesterone receptor
PTEN	:	Phosphatase and tensin homolog
RAD51C	:	RAD51 homolog C
RAD51D	:	RAD51 Paralog D
ref	:	Reference
RFLP	:	Restriction fragment length polymorphism
RNA	:	Ribonucleic acid
ROS	:	Reactive oxygen species
rpm	:	Revolutions per minute
RR	:	Relative risk
s	:	Second

S1P	:	Sphingosine-1-phosphate
SBR	:	Modified Scarff–Bloom–Richardson grading system
SD	:	Standard deviation
SEER	:	National Cancer Institute Surveillance, Epidemiology, and End Results
SNP	:	Single nucleotide polymorphisms
SP	:	Side population
SPSS	:	Statistical Package for Social Sciences
TBE	:	Tris-borate-EDTA
T-DM1	:	Ado-trastuzumab emtansine
TEC	:	Taxotere-epirubicin-cyclophosphamide
TIL	:	Tumour-infiltrating lymphocytes
TNBC	:	Triple negative breast cancer
TNM	:	Tumour, Node, Metastasis
TP53	:	Tumour protein p53
US	:	Ultrasound
USM	:	Universiti Sains Malaysia
V	:	Voltage
VAF	:	Variant allele frequency
WHO	:	World Health Organization

LIST OF APPENDICES

PERANAN VARIASI GENETIK ABC (*ABCB1*, *ABCC1* DAN *ABCG2*) TERHADAP RISIKO KECENDERUNGAN KANSER PAYUDARA TIGA KALI NEGATIF

ABSTRAK

Kanser payudara tiga kali negatif (TNBC) dikenali dengan sifatnya yang agresif dan dikaitkan dengan prognosis yang kurang baik. Gen pengangkut ABC telah dikaji secara meluas kerana hubungannya dengan pelbagai rintangan ubat pada pesakit kanser dan penyakit lain. Namun begitu, bukti mengenai polimorfisme gen pengangkut ABC dan kerentanan terhadap TNBC masih terhad dan tidak konklusif. Objektif utama kajian ini adalah untuk menyiasat hubungan polimorfisme gen pengangkut ABC iaitu *ABCB1*, *ABCC1*, dan *ABCG2* dengan kerentanan terhadap TNBC dalam populasi Malaysia. Kajian kawalan kes ini dijalankan di Hospital Universiti Sains Malaysia, Kubang Kerian, Kelantan. Sebanyak tiga mL sampel darah periferal telah dikumpul daripada 79 pesakit TNBC dan 100 individu kawalan yang sihat, diikuti dengan pengekstrakan DNA genom daripada darah. Seterusnya, penentuan genotaip dijalankan dengan menggunakan tindak balas rantai polimerase-polimorfisme panjang serpihan sekatan (PCR-RFLP) menggunakan enzim seperti *TaaI*, *BseMI*, *RsaI*, *EcoO1091*, *MboI* dan *PspF1* serta kaedah tindak balas rantai polimerase sistem mutasi penentangan penguatan (ARMS-PCR). Genotaip diperiksa dengan memerhatikan saiz jalur serpihan yang dicerna dan produk PCR pada gel agarosa 3% melalui elektroforesis. Alel, genotaip dan haplotaip polimorfisme telah dinilai, dan ujian $\chi 2$ bebas dijalankan untuk menjelaskan hubungannya dengan kerentanan terhadap TNBC dan pemboleh ubah klinikopatologi lain manakala regresi logistik dijalankan untuk mengukur kekuatan hubungan. Analisis alel dan genotaip

menunjukkan bahawa alel T dan pembawa genotip TT bagi *ABCB1* 1236 C>T dikaitkan dengan peningkatan risiko kerentanan TNBC (p = 0.026 dan 0.035, masing-masing) dengan OR sebanyak 1.628 (95% CI: 1.060-2.501) dan 3.034 (95% CI: 1.171-7.866) masing-masing. Selain itu, *ABCG2* 421 C>A juga dikaitkan dengan peringkat tumor yang lebih lanjut (p = 0.004) dan OR untuk alel A dan pembawa genotaip AA ialah 3.464 (95% CI: 1.687-7.111) dan 11.625 (95% CI: 2.187-61.804) masing-masing. Begitu juga, hubungan yang signifikan juga diperhatikan dalam pembawa genotaip AA bagi *ABCG2* 421 C>A dengan subtipe histologi yang lebih jarang seperti karsinoma metaplastik dan medulari (p = 0.016), dengan OR sebanyak 6.171 (95% CI: 1.467-25.961). Sementara itu, analisis haplotaip menunjukkan hubungan yang signifikan antara pembawa *ABCB1* 1236C/3435T/2677G dan kerentanan terhadap TNBC (p = 0.011), menunjukkan kesan perlindungan dengan OR yang lebih rendah sebanyak 0.120 (95% CI: 0.015-0.952). Selain itu, pembawa haplotaip *ABCG2* 34G/421A dikaitkan dengan peningkatan risiko peringkat lanjut (p = 0.030, OR: 2.333, 95% CI: 1.005-5.417) dan jenis histologi yang lebih jarang seperti karsinoma metaplastik dan medulari (p = 0.009, OR: 2.599, 95% CI: 1.027-6.576). Kesimpulannya, kajian ini mencadangkan bahawa polimorfisme *ABCB1* dan *ABCG2* dikaitkan dengan kerentanan terhadap TNBC, peringkat lanjut, dan jenis histologi yang lebih jarang. Kajian ini menyokong hipotesis bahawa varian terpilih dalam gen pengangkut ABC menyumbang kepada risiko kerentanan terhadap TNBC dan boleh dipertimbangkan sebagai penanda biomarker dan faktor prognosis dalam pengurusan TNBC.

ROLE OF GENETIC VARIATIONS OF ABC (*ABCB1*, *ABCC1* AND *ABCG2*) ON TRIPLE NEGATIVE BREAST CANCER SUSCEPTIBILITY RISK

ABSTRACT

Triple-negative breast cancer (TNBC) is known for its aggressive behaviour and is associated with poor prognosis. ABC transporter genes were widely studied for their association with multiple drug resistance in cancer patients and other diseases. However, the evidence on ABC transporter gene polymorphisms and TNBC susceptibility remains limited and inconclusive. The main objective of this study was to investigate the association of ABC transporter genes *ABCB1*, *ABCC1*, and *ABCG2* polymorphisms with TNBC susceptibility in the Malaysian population. This case-control study was conducted at Hospital Universiti Sains Malaysia, Kubang Kerian, Kelantan. A volume of three mL of peripheral blood samples were collected from 79 TNBC patients and 100 healthy controls, followed by genomic DNA extraction from the blood. Next, the genotyping was performed by employing polymerase chain reaction-restriction fragment length polymorphisms (PCR-RFLP) using enzymes such as *TaaI*, *BseMI*, *RsaI*, *EcoO1091*, *MboI* and *PspF1* as well as amplification-refractory mutation system-polymerase chain reaction (ARMS-PCR) methods. The genotypes were examined by observing the band sizes of digested fragments and PCR products on 3% agarose gel through electrophoresis. The allele, genotype and haplotype of polymorphisms were evaluated, and the independent χ^2 test was carried out to elucidate their association with TNBC susceptibility and other clinicopathological variables while the logistic regression was performed to measure the strength of association. The allele and genotype analysis showed that the T allele and TT genotype carrier of *ABCB1* 1236 C>T were associated with the increased risk of TNBC

susceptibility (p = 0.026 and 0.035, respectively) with OR of 1.628 (95% CI: 1.060-2.501) and 3.034 (95% CI: 1.171-7.866) respectively. Besides, *ABCG2* 421 C>A was also associated with the advanced staging of the tumour (p = 0.004) and the OR for the A allele and AA genotype carrier was 3.464 (95% CI: 1.687-7.111) and 11.625 (95% CI: 2.187-61.804) respectively. Likewise, a significant association was also observed in the AA genotype carrier of *ABCG2* 421 C>A with rarer histologic subtypes of metaplastic and medullary carcinoma (p = 0.016), with an OR of 6.171 (95% CI: 1.467-25.961). Meanwhile, the haplotype analysis demonstrated a significant association between the *ABCB1* 1236C/3435T/2677G carrier and TNBC susceptibility (p = 0.011), showing the protective effect with reduced OR of 0.120 (95% CI: 0.015-0.952). Besides, the haplotype carrier of *ABCG2* 34G/421A was associated with an increased risk of advanced staging (p = 0.030, OR: 2.333, 95% CI: 1.005-5.417) and rarer histologic type of metaplastic and medullary carcinoma (p = 0.009, OR: 2.599, 95% CI: 1.027-6.576). In conclusion, the present study suggests that *ABCB1* and *ABCG2* polymorphisms were associated with TNBC susceptibility, advanced staging, and rarer histologic types of carcinomas. This study supports the hypothesis that selected variants in ABC transporter genes contribute to TNBC susceptibility risk and could be considered candidate biomarkers and prognostic factors in TNBC management.

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Breast cancer is the most prevalent cancer type in female worldwide, accounting for 23.8% among 9.7 million cases in females, according to the report from Global Cancer Observatory (GLOBOCAN) 2022 (Bray et al., 2024). While in Malaysia 21, 634 cases were reported from 2012 to 2016, which amounted to the most of all cancer types in females (Azizah et al., 2019). Among all subtypes of breast cancer, triple-negative breast cancer (TNBC) was known for its aggressive behaviour and poor prognosis, signified by the higher histologic grade, higher risk for lymph node positivity, shorter recurrence time and lower 5-year overall survival (Gonçalves et al., 2018; Kulkarni et al., 2020).

TNBC is defined by the absence of estrogen receptor (ER), progesterone receptor (PR) and no amplification of human epidermal growth factor receptor 2 (HER2) under immunohistochemistry (IHC) staining (Kumar & Aggarwal, 2016). Overall, the prevalence of TNBC could amount from 7.7% to 16.0% among all breast cancers worldwide (Acheampong et al., 2020; Al-thoubaity, 2020; Su et al., 2011; Tada et al., 2023). While in Malaysia, the incidence could range between 12.4% to 17.6% (Azman et al., 2019; Kanapathy Pillai et al., 2012; Tan et al., 2009) among all breast cancer cases. The treatment options for TNBC patients are generally restricted, primarily limited to chemotherapy as the primary systemic treatment. This limitation arises mainly from a lack of hormonal and HER2 receptors in cancer cells, which the endocrine therapy that is widely used in breast cancer treatment do not achieve targeted efficacy.

Several risk factors have been observed to relate to the risk of TNBC, such as younger age, pre-menopausal status, and lack of history of breastfeeding (Horakova et al., 2018). Meanwhile, attention should be given to the genetic factors since several studies have suggested a positive correlation between breast cancer family history and TNBC susceptibility (Anderson et al., 2014; Gomes et al., 2022; Liu et al., 2021). One of the famous genetic predisposition factors is a mutation in *BRCA1/2*, in which the functional proteins play a crucial role in homologous recombination (Gudmundsdottir & Ashworth, 2006). Approximately 70% of breast cancer patients who inherit the mutation demonstrated low or absent expression of hormonal receptors and HER2 (Karim et al., 2023). *BRCA1* mutation carriers are also more likely to develop TNBC than non-carriers, according to the study by Chen et al. (2018). Other predisposition genes related to TNBC include *PALB* and *FANCM* (Hahnen et al., 2017; Kiiski et al., 2014; Zhou et al., 2020).

ATP-binding cassette (ABC) transporter is a group of transmembrane proteins that translocate various compounds in couple with ATP. As for now, a total of 48 genes that have been classified into seven groups have been discovered (Dean et al., 2001). Most of the members function to translocate various compounds such as xenobiotics, metabolites, lipids and hormones (Holland, 2011). This study included three prominent genes, the *ABCB1*, *ABCC1* and *ABCG2*, that played a vital role in the excretion of xenobiotics.

ABCB1 gene encodes for MDR1 (multidrug resistance protein 1), aka P-glycoprotein (PGP). This protein serves various functions, such as restricting xenobiotic absorption, translocating signalling peptides in immune cells, and participating in hormonal regulation within reproductive organs (Brinkmann, 2001; Efferth & Volm, 2017). The genetic variations might affect the proteins' functionality,

which we postulated might affect the individual susceptibility towards cancer. The study by Salama et al. (2006) demonstrated the reduction of protein activity from 80% to 100 % due to the presence of the *ABCB11236T/2677T/3435T* variant.

The most prominent variants that could be found in the Asian population are *ABCB1* 1236 C>T (rs1128503), 2677 T>G/A (rs2032582) and 3435 C>T (rs1045642), which accounted for 31.2%, 52.2%, 15.2% and 37.5% respectively (Ieiri, 2012). Several studies have illustrated their correlations with cancers such as colorectal cancer (Yue et al., 2013), lung cancer (Zawadzka et al., 2020), as well as breast cancer (Wu et al., 2012). Besides that, its pharmacogenetics value has also been demonstrated in several studies involving breast cancer patients (Li et al., 2017) and TNBC patients (Abdul Aziz et al., 2018).

ABCC1 encodes for multidrug resistance-associated protein (MRP1), which involves the exportation of physiological substrates, organic anions and xenobiotics across the membrane and is well-known for causing multidrug resistance in cancerous cells (Conseil et al., 2006; Stride et al., 1997). Two of the variants, the 2012 G>T (rs45511401) and 825 T>C (rs246221), were selected for our study. The 825 T>C is one of the prominent single nucleotide polymorphisms (SNP) in the Asian population, accounting for 37.0% to 53% (Fukushima-Uesaka et al., 2007; Yang et al., 2021). On the other hand, although the prevalent of 2012 G>T is low in the Asian population, their effect on protein translation (replace glycine to valine) might cause functional alteration, which eventually causes the reduction in transport activity of the protein (Gao et al., 2000; Szakács et al., 2000). The studies on *ABCC1* and cancer were mainly focusing on pharmacogenetics prediction, which aimed to investigate the response of cancer patients to anticancer drug regimens (Kunická & Souček, 2014; Pfeil et al., 2014; Vulsteke et al., 2013).

ABCG2 encodes for a transmembrane transporter protein widely studied for its function to cause multidrug resistance in cancer cells. It can extrude a variety of anticancer drugs, such as Adriamycin, Daunorubicin, Topotecan, and Mitoxantrone (Ross & Nakanishi, 2010). Among all the variants, 421 C>A (rs2231142) and 34 G>A (rs2231137) represented the most in the Asian population, accounting for 28.9% and 19.3%, respectively (Ieiri, 2012). The substitution of glycine to lysine by 421 C>A will reduce the expression and ATPase activity of the protein (Furukawa et al., 2009; Mizuarai et al., 2004). Meanwhile, substituting methionine to valine by 34 G>A will interfere with localising the protein on the plasma membrane (Heyes et al., 2018). On the other hand, the 376 C>T (rs72552713) that was selected in this study could promote pre-mature stop codon (glutamine to stop codon) that eventually impairs the function of the transporter (Kobayashi et al., 2005). As for now, *ABCG2* polymorphisms have been investigated to associate with different kinds of cancers, including breast cancer (Wu et al., 2015), multiple myeloma (Niebudek et al., 2019) and myeloid leukaemia (Salimizand et al., 2016).

1.2 Problem statements

TNBC is known for its aggressive phenotype and behaviour, and the patients usually exhibited a poor prognosis compared to the patients of other subtypes. For instance, TNBC patients have a worse overall survival (OS) and shorter disease-free survival (DFS) (Agarwal et al., 2016; Li et al., 2017). Also, the patients were more likely to develop a higher tumour grade (Kulkarni et al., 2020), with positive lymphocytic infiltration and a shorter recurrence time (Gonçalves et al., 2018). Since the early stage of TNBC has a better OS and DFS than later stage TNBC (Agarwal et

al., 2016; Chen et al., 2020), preventive measures and early detection are crucial, especially for high-risk individuals.

The study of genetic predisposition factors could help to identify the TNBC-related biomarkers so that the high-risk individuals from the population could be identified. Despite the well-studied high penetrance genes such as *BRCA1*/2, the present study aims to investigate the ABC transporter genes as a new biomarker for TNBC susceptibility assessment. ABC transporter genes were widely studied for their association with multiple drug resistance in cancer patients, mainly due to their function that can extrude anticancer drugs. The study by Abdul Aziz et al. (2018) showed that the variant genotype and allele of *ABCB1* 3435 C>T and haplotype 1236T/3435T/2677T and 1236G/3435T/2677T were associated with chemoresistance.

Despite their role in causing multidrug resistance, much literature has included their association with various cancers and diseases. However, the evidence on ABC transporter gene polymorphisms and TNBC susceptibility remains limited and inconclusive. Since there are no reports available in the Malaysian population and the results vary across populations, a case-control study was undertaken to investigate the association of ABC transporter gene polymorphisms, particularly *ABCB1* (2677 G>T/A, 1236 C>T and 3435 C>T), *ABCC1* (2012 G>T and 825 T>C) and *ABCG2* (34 G>A, 421 C>A and 376 C>T) polymorphisms, with TNBC susceptibility in the Malaysian population. The study aims to advance knowledge on genetic predisposition factors influencing TNBC susceptibility in Malaysia. Other than that, the present may also help identify high-risk individuals in the population so that appropriate preventive measures and surveillance programs can be taken.

1.3 Research hypotheses

Null hypothesis: There is no association between *ABCB1, ABCC1*, and *ABCG2* polymorphisms on TNBC susceptibility risk.

Alternate hypothesis: There is an association of *ABCB1*, *ABCC1* and *ABCG2* polymorphisms on TNBC susceptibility risk.

1.4 Objective (s) of the study

The main objective of this study was to investigate the contribution of genetic variations of ABC transporter genes (*ABCB1*, *ABCC1* and *ABCG2*) on TNBC susceptibility risk.

The specific objectives were:

1. To assess the genotype, allele, and haplotype frequencies of *ABCB1* (2677 G>T/A, 1236 C>T and 3435 C>T), *ABCC1* (2012 G>T and 825 T>C) and *ABCG2* (34 G>A, 421 C>A and 376 C>T) polymorphisms in TNBC patients and healthy normal controls.
2. To associate the *ABCB1* (2677 G>T/A, 1236 C>T and 3435 C>T), *ABCC1* (2012 G>T and 825 T>C) and *ABCG2* (34 G>A, 421 C>A and 376 C>T) polymorphisms with TNBC susceptibility risk.
3. To evaluate whether data generated on the SNPs of *ABCB1*, *ABCC1* and *ABCG2* could serve as predictive biomarkers for TNBC susceptibility risk.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of breast cancer

2.1.1 Breast cancer carcinogenesis

Breast cancer is a malignancy that originates from the uncontrolled growth and proliferation of abnormal cells within the breast tissue, particularly in the mammary glands' ducts or lobules. Figure 2.1 has illustrated the classic multistep carcinogenesis model, recognising three main phases: initiation, promotion, and progression (Devi, 1991). The initiation of breast cancer was thought to be started from transformation (genetic or epigenetic alteration) arising spontaneously or induced by carcinogens at the cellular level (Polyak, 2007). For instance, the functionality alteration caused by BRCA1 gene mutation is known to compromise the DNA repair function of the cells which further lead to the transformation due to the accumulation of mutations (Rosen et al., 2003). On the other hand, the epigenetic control such as hypermethylation event in BRCA1 gene promoter region would silence the expression of gene, which in turn reducing or causing the complete loss of protein translation (Esteller et al., 2000).

While the onset of transformation is complex and involves complicated steps, several factors, such as heredity, genetic factors, hormonal exposure, and environmental exposure to carcinogens, are among the contributing factors. For instance, the inherited *BRCA1/2* mutation is known for its association with breast cancer. They function as tumour suppressor genes that play a crucial role in DNA repair and maintaining chromosomal stability, and the mutated genes promote breast cell transformation through genomic instability (Yoshida & Miki, 2004). Meanwhile, prolonged exposure to hormones, especially estrogens, is also a significant risk factor

as the initiator for genomic changes and transformation in breast tissues (Russo & Russo, 2006).

However, the initial transforming event is not enough to form cancer. The neoplastic development is highly influenced by intra and extracellular environments. The subsequent recruitment of host cell partners enables niche construction in a stromal climate, facilitating survival and promoting proliferation and malignant behaviour in initiator cells (Barcellos-Hoff et al., 2013). Promoter signals such as cytokines and chemokines or pro-inflammatory substances (Okumura et al., 2010), as well as physiological conditions such as the composition of tissue at the time of initiation (Bemis & Schedin, 2000), strongly influence the promotion of tumour.

The following clonal expansion further promotes the progression to the stage of neoplasia and, subsequently, carcinoma *in situ*. This expansion of niches involves more components, such as immune cells and fibroblasts, that can promote chronic inflammation and alter cellular adhesion early in carcinogenesis (Schor & Schor, 2001). This interplay of different entities helps the survival of transformed cells and the progression of carcinoma. Further niche maturation occurs when the angiogenesis happens, thus forming a stable tumour microenvironment (Barcellos-Hoff et al., 2013).

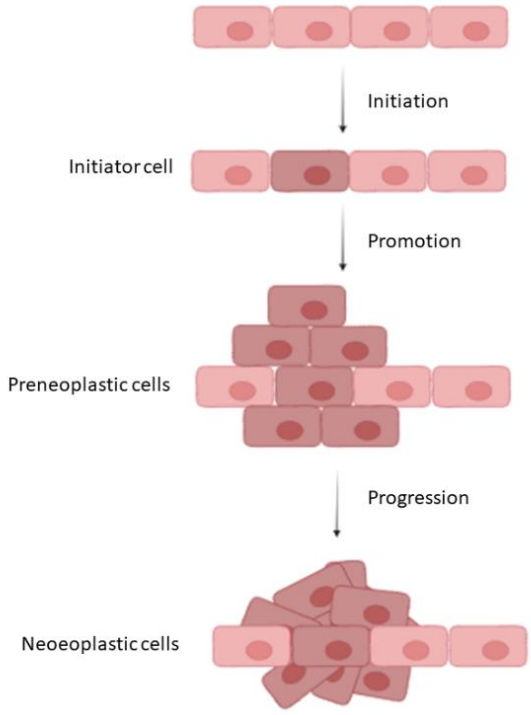

Figure 2.1. Schematic diagram of multistep carcinogenesis model of initiation, promotion, and progression. Created with BioRender.com.

2.1.2 Breast cancer incidence

Breast cancer is the most dominant cancer in women worldwide. According to the GLOBOCAN 2022, the prevalence of breast cancer could reach up to 23.8% among 9.7 million cancer cases in females, with an estimated mortality rate of about 15.4% among 4.3 million cancer deaths (Bray et al., 2024). This made breast cancer to be the top among all the cancer types in females. The incidence of breast cancer varies according to the region of the world. For instance, an 88% higher incidence rate was observed in developed countries (54.1 per 100 000) compared to the transitioning countries (30.8 per 100 000). The highest incidence rates were notably found in Australia, western Europe and Northern America, while the lowest rate was found in Central America, Africa and Asia (Bray et al., 2024).

In Malaysia, breast cancer cases were recorded in 21,634 females from 2012 to 2016, accounting for the most of all cancers. The age-standardised rate (ASR) was

34.1 per 100,000 people, with the highest incidence rate to be observed among Chinese (40.7 per 100,000), followed by Indians (38.1 per 100,000) and Malays (31.5 per 100,000) (Azizah et al., 2019). According to World Health Organization (WHO), 8,418 breast cancer cases were recorded in females, accounting for 32.9% of all new cases of cancer in the year 2020. The ASR was 49.3 per 100, 000 population with mortality rates of 20.7 per 100, 000 (https://gco.iarc.fr/today/data/factsheets/populations/458-malaysia-fact-sheets.pdf). Unsurprisingly, this made breast cancer ranked top again among all the cancers in women. Thus, the impact of breast cancer on public health is significant and cannot be ignored.

2.1.3 Classification of breast cancer subtypes

Breast cancer is a heterogeneous disease that can be distinguished by different schemata such as morphology, histopathology, molecular manifestation, and gene profiling (Makki, 2015). With the advance of genomic and expression profiling techniques, the classification system focusing on the molecular expression of breast cancer cells has reached a consensus. It is now widely used for subtyping breast cancer.

The current molecular classification divides the subtypes into four main groups with different extents of involvement of three receptors: the ER, PR and HER2. They are the Luminal A subtype that has high expression of ER and PR with lacking HER2 expression; the Luminal B subtype with positive ER and negative HER2 expression plus varies expression of PR; HER2 with positive expression of HER2 and negative in hormonal receptors, followed by basal-like or triple-negative with negative in hormonal receptors and low expression of HER2 (Eliyatkin et al., 2015).

The prevalence of each subtype differs according to region and country, but in general, Luminal A dominated among all subtypes, ranging from 37 to 73.2% worldwide. The next prominent subtype is Luminal B, whose prevalence could range from 6.9 to 31.3% worldwide. Meanwhile, for the HER2 subtype, the number varies from 4.4 to 20.6%, and lastly, for the triple-negative, 7.7 to 23.9% was recorded worldwide (Acheampong et al., 2020; Al-thoubaity, 2020; Azman et al., 2019; Carey et al., 2006; Hjerkind et al., 2022; Pandit et al., 2020; Spitale et al., 2009; Su et al., 2011; Tada et al., 2023).

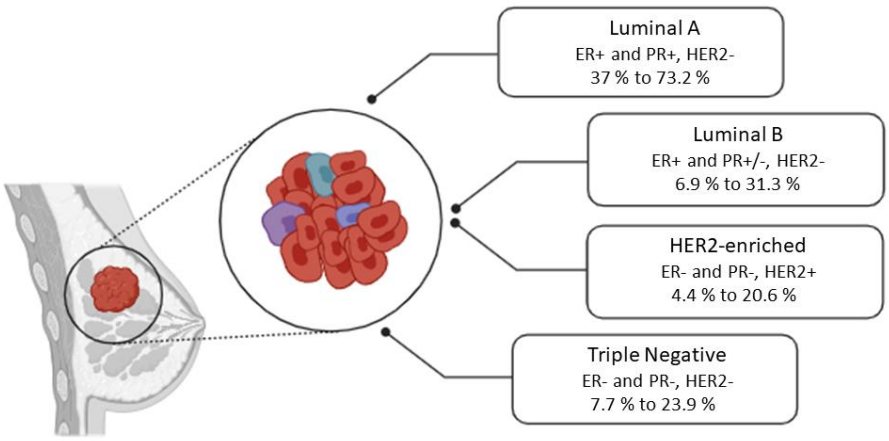

Figure 2.2. Subtypes of breast cancer based on the molecular classification. Created with BioRender.com.

2.1.4 Risk factors of breast cancer

Breast cancer is a multifactorial disease, contributing to various factors such as genetics, age of menarche and menopause, breast density, and other modifiable risk factors, including parity and breastfeeding history, as well as oral contraceptive use.

2.1.4(a) Genetics

Genetics is a significant contributing factor to breast cancer development. Family history linked to breast cancer cases is estimated to make up 5% to 10% of all cases worldwide, while in Malaysia, the incidence was accounted for 13.6% to 14.4% (Liaw et al., 2020;Tan et al., 2018; Wen et al., 2018). A large cohort study conducted by Brewer et al. (2017) in the UK population estimated the increased risk of breast cancer by 2.52 fold (95% CI: 1.83-3.47) in women that have two or more first-degree relatives with breast cancer and their younger age of diagnosis could serve as a powerful predictor for risk calculation. Likewise, the study by Shiyanbola et al. (2017) highlighted the importance of first-degree family history as a factor for breast cancer risk.

In Malaysia, a case-control study by Tan et al. (2018) reported an increased odd ratio (OR) for breast cancer risk by 1.19 (95% CI: 1.02-1.38) if the individual has a first-degree family history of breast cancer within a cohort of 7663 individuals. Similarly, Mohd Razif et al. (2011) investigated the effect of family history towards premenopausal breast cancer risk among Malaysian women. The study reported that breast cancer history in any relative could raise the risk of premenopausal breast cancer by 4.81-fold (95% CI: 2.41-9.58) while those that have a first-degree family history of breast cancer recorded a higher OR by 5.45 (95% CI: 2.10-14.13).

Therefore, individuals with a significant family history should be considered with germline genetic testing for multigene panels consisting of clinically validated hereditary breast and ovarian cancer syndrome genes. This includes the prominent BRCA1 and BRCA2 gene mutations that increase mutant carriers' lifetime risk by approximately 60 % (Sessa et al., 2023). The functional BRCA proteins play a crucial role in maintaining genomic stability by repairing the double-strand breaks through homologous recombination (Gudmundsdottir & Ashworth, 2006). They are a pivotal tumour suppressor gene whose loss or impairment of function can promote genomic instability and compromise DNA damage repair. Other predisposition gene markers include *PALB*, *PTEN*, and *TP53* (Sessa et al., 2023).

2.1.4(b) Age at menarche and menopause

Other than that, the age at menarche and menopause are also considered essential predisposition factors. They mark the onset and cessation of ovarian activity associated with women's reproduction. A meta-analysis involving 118 964 breast cancer women demonstrated that, for every year younger at menarche, the risk of breast cancer increased by 1.050 (95% CI: 1.044-1.057) while every year older at menopause slightly increased the risk by 1.029 (95% CI: 1.025-1.032). When comparing the identical age women with pre- and post-menopausal status, premenopausal women had a greater risk by 1.43 (95% CI: 1.33-1.52) (Hamajima et al., 2012). Likewise, the early age of menarche (<12 years old) was found to be positively correlated with breast cancer susceptibility (HR: 1.10, 95% CI: 1.01-1.20) (Goldberg et al., 2020). Meanwhile, menopausal status also plays a significant role in influencing breast cancer susceptibility. For instance, studies have found an association between premenopausal status and breast cancer (Tan et al., 2018), with particularly early onset at age ≤40 years old (Yang et al., 2022).

2.1.4(c) Breast density

Breast density refers to the proportion of fibro-glandular tissues relative to the fat within the breast. According to the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS), it is separated into four categories, ranging from Category A (the least density, primarily composed of fat) to Category D (extremely dense), as assessed through mammography (Winkler et al., 2015). Thick breast density is considered a risk factor for cancer susceptibility, mainly due to the high fibro-glandular tissues that contain more epithelial and stromal cells, which are at risk of carcinogenic transformation (Soguel et al., 2017).

Several studies provided significant evidence to illustrate the positive association between dense breasts and breast cancer risk. The nested case-control study that involved 1112 matched pair subjects demonstrated a strong correlation between dense breast and breast cancer, showing an OR of 4.7 (95% CI: 3.0-7.4) in women with breast density more than 75% (Sun et al., 2007). Likewise, a meta-analysis correlated dense breast (Category D) with a high risk of breast cancer, with a 2.11-fold (95% CI: 1.84-2.42) increase in the risk (Bodewes et al., 2022). A similar result was also observed in the southeast Asian population, with an OR of 3.32 (95% CI: 2.44-4.52) for individuals having breast density from 26.04%-100% based on the mammographic screening (Ho et al., 2020).

2.1.4(d) Parity and breast-feeding history

Parity and breastfeeding history are both factors that can influence women's risk of developing breast cancer. Parity is associated with a reduced risk overall, but the timing and age at which the pregnancies occur could influence the risk of breast cancer (Albrektsen et al., 2005). A pooled analysis showed a positive association of parous women with breast cancer risk, peaked about five years after birth (HR: 1.80, 95% CI: 1.63-1.99), which would then decrease to 0.77 (95% CI: 0.67-0.88) after 34 years (Nichols et al., 2019). Overall, women with a history of more than one full-term pregnancy had a reduced risk, especially hormonal receptor positive (HR+) subtypes of breast cancer, by 13% (95% CI: 0.78-0.37) and 19% (95% CI: 0.71-0.91) for ER+ and ER+/PR+ subtypes respectively (Ma et al., 2010).

Other than that, breastfeeding is also consistently associated with a reduced risk of breast cancer. Although not fully understood, the reduced menstrual cycles and exposure to specific hormones during pregnancy and breastfeeding could be the underlying hormonal factors. Meanwhile, the differentiation of mammary cells during lactation is more resistant to carcinogenesis and at the same time, the cells with DNA damage could be removed from breast tissue during the process (Anstey et al., 2017). Overall, the relative risk decreased with the increasing breastfeeding lifetime months in parous women, based on epidemiological studies in 30 countries (CGHFBC, 2002). In Malaysia, the parous women who have breastfeeding experience were protected from breast cancer with a reduced OR of 0.56 (95% CI: 0.48-0.65) (Tan et al., 2018). Likewise, the study by Ho et al. (2020) in other southeast Asian populations illustrated an elevated risk by 1.49 (95% CI: 1.22-1.82) in parous women without breastfeeding experience.

2.1.4(e) Hormonal contraceptive use

The global investigation into the relationship between hormonal therapy, including oral contraceptive use, and breast cancer risk has yielded valuable insights suggesting that such treatment is associated with an increase in breast cancer risk among women. This association is greatly influenced by usage duration and formulation types. For instance, the lifetime study among 1.2 million young women (age 24 to 43 years old during enrolment) showed that the current use of oral contraceptive was associated with increased risk by 1.33 times (95% CI: 1.03-1.73), particularly for the user of Triphasic (RR: 3.05, 95% CI: 2.00-4.66) and Norgestrel (RR: 1.89, 95% CI: 1.05-3.41) (Hunter et al., 2010). Likewise, the use of hormonal contraception with a duration of more than six months previously slightly elevated the risk by 1.08 times (95% CI: 1.03-1.13) and the highest risk was recorded in individuals that having more than ten years of use of any hormonal contraception (RR: 1.38, 95% CI: 1.26-1.51). A significant association was shown between the oral combined ethinyl oestradiol with progestin such as Levonorgestrel, Norgestimate, Desogestrel and Gestodene, and the single use of Levonorgestrel (Mørch et al., 2017).

2.1.5 Breast cancer diagnosis

Breast cancer is typically identified either through screening procedures or the manifestation of symptoms, such as pain or the detection of a palpable mass, which initiates diagnostic evaluation. Early diagnosis offers a better survival and prognosis (Tan et al., 2021) , and several imaging techniques are widely used to address diverse clinical scenarios.

2.1.5(a) Mammography

Mammography is one of the standard imaging techniques for screening and diagnosis of breast cancer. The mammogram is obtained using a low-energy X-ray to visualise the breast structure post-compressed between two plates. Mammography screening aims to detect breast cancer early, often before symptoms appear. On the other hand, diagnostic mammography helps diagnose breast cancer when symptoms such as breast lumps are present (Bhushan et al., 2021).

In 2019, Malaysia's Ministry of Health (MOH) issued clinical practice guidelines recommending the consideration of screening from 30 - 39 years of age in high-risk populations and annually from age 40 and older. Meanwhile, biannual screenings were suggested for women aged 50 – 74 in the general population (MOH, 2019). Likewise, the European Commission Initiative on Breast Cancer (ECIBC) recommended a biannual screening for average-risk women aged 50-69 years old with conditional recommendation for women of younger and older age (Loibl et al., 2023).

However, high breast density, which could obscure underlying cancer, is one of the major contributing factors to false-negative results. Thus, supplementary imaging assessments such as ultrasound (US) or advanced digital breast tomosynthesis (DBT) should be considered for women with high breast density (Zhang et al., 2023). Apart from mammography, further assessments are required for diagnostic purposes

in clinical settings, including clinical assessment (Clinical Breast Examination) and pathology (histology and/or cytology) as suggested by clinical practice guidelines of Malaysia's Ministry of Health (MOH, 2019).

2.1.5(b) Magnetic resonance imaging (MRI)

In the case of uncertainties and under certain clinical situations, magnetic resonance imaging (MRI) is recommended following the standard imaging procedure (Loibl et al., 2023; MOH, 2019). MRI uses strong magnets and low-frequency radio waves to generate detailed images within the body. It is non-invasive and non-ionizing (Van Geuns et al., 1999), hence providing a safer approach for imaging.

MRI offers several advantages over other medical imaging techniques, including non-destructive and non-invasive examination, absence of ionising radiation, high soft tissue contrast, and multiple imaging parameters. It provides detailed views of tissue morphology and anatomical structure, allowing for a comprehensive analysis of physiological functions. MRI's superior soft-tissue resolution results in higher sensitivity in breast cancer screening and diagnosis than conventional imaging methods (Ruan & Sun, 2023). Although MRI is highly sensitive (Cho et al., 2017; Niell et al., 2017), its reduced specificity (Warner et al., 2008) necessitates the involvement of an experienced reader (Niell et al., 2017), highlighting the subjective nature of qualitative justification. Also, the need for intravenous contrasting agents, as well as the limited accessibility and substantial cost, would need to be considered (Zhang et al., 2023).

2.1.5(c) Ultrasound (US)

Ultrasound (US) is an imaging approach that uses sound waves above the upper limit of human hearing to penetrate tissues and produce images by receiving the reflected echoes. It allows a very safe approach and does not cause injury to human biological tissues (Evans et al., 2018). US is also an indispensable tool and could be adjunct and complementary for mammography and MRI for breast imaging. It provides a real-time and dynamic approach to analysing the lesion and detecting subtle findings compared to static images. For instance, it allows the direct assessment of palpable lesions and real-time access to the mobility and location of lesions together with their adjacent structures (Hooley et al., 2013). Furthermore, the US also acts as an excellent supplementary screening after mammography for women with high breast density, increasing the detection rate by 40% compared to mammography alone (Rebolj et al., 2018). Meanwhile, axillary triage and assessment are best performed using US techniques, showing a high specificity of 100% (95% CI: 99–100%) and sensitivity of 51% (95% CI: 43–59%), especially for high metastasis burden (Boulc'h et al., 2021). However, the reliability of the result might be intensely dependent on the experience and expertise of the examiner and patient posture habitus, provided the highly operator-dependent nature of the US (Evans et al., 2018).

2.1.5(d) Pathology assessment

Apart from imaging techniques, the pathology assessment of biopsy tissues is also critical for the triple evaluation in breast cancer diagnosis (MOH, 2019). Core needle biopsy is an established tool to obtain the breast lesion tissues following the histomorphological examination and molecular and biomarker assessment. While hematoxylin and eosin (H&E) staining remain the gold standard for routine pathological diagnosis (Fischer et al., 2012), IHC staining has been implemented as a standard in most of the clinical practice guidelines to assess the biomarker status including ER, PR and HER2 status (Allison et al., 2019; Loibl et al., 2023; MOH, 2019; Wolff et al., 2022).

2.1.5(e) Tumour staging

The staging and grading of tumours provide vital information to guide treatment following the diagnosis and have an impact on the patient's prognosis. Currently, the protocol published by the American Joint Committee on Cancer (AJCC) for the specimen examination in 2018 (8th edition) has been standardised by the Ministry of Health Malaysia for the staging of patients (MOH, 2019).

The staging takes into consideration primary tumour characteristics (T), lymph node metastases (N) and the presence of distant metastasis (M). Individual scores are attributed to each T (0-4), N (0-3), and M (0-1) and combined to form the overall clinical or pathological stages (0 to IV). While clinical staging provides a guide on treatment decisions and prognosis before the treatment and intervention based on the findings of clinical examination and radiology, pathological staging is considered more accurate because of the direct information provided by the histopathological assessment based on the tissue specimen resected from patients. The definitions for all aspects and TNM classification were summarised in Tables 2.1 to 2.4.

Table 2.1. Pathological definition of primary tumour (T) based on AJCC cancer staging manual (8th edition).

Category	Criteria
T0	No evidence of a primary tumour
Tis	Ductal carcinoma in situ (DCIS) or Paget disease
T1	Tumour ≤20 mm in greatest dimension T1mi: Tumour ≤1 mm in greatest dimension T1a: Tumour >1 mm but ≤5 mm in greatest dimension T1b: Tumour >5 mm but ≤10 mm in greatest dimension T1c: Tumour >10 mm but ≤20 mm in greatest dimension
T2	Tumor >20 mm but ≤50 mm in greatest dimension
T3	Tumor >50 mm in greatest dimension
T4	Tumor of any size with direct extension to the chest wall and/or to the skin T4a: Extension to the chest wall T4b: Extension to the skin that does not meet the criteria for inflammatory T4c: Both T4a and T4b are present T4d: Inflammatory carcinoma

Table 2.2. Pathological definition of regional lymph nodes (N) based on AJCC cancer staging manual (8th edition).

Category	Criteria
pN0	No regional lymph node metastasis was identified or ITCs only pN0(i+): ITCs only (malignant cell clusters ≤0.2 mm)
pN1	Micrometastases or metastases in 1–3 axillary lymph nodes and/or clinically negative internal mammary nodes with micrometastases or macrometastases by sentinel lymph node biopsy pN1mi: Micrometastases (approximately 200 cells, 0.2 mm - 2.0 mm) pN1a: Metastases in 1–3 axillary lymph nodes, at least one metastasis > 2.0 mm pN1b: Metastases in ipsilateral internal mammary sentinel nodes pN1c: pN1a and pN1b combined
pN2	Metastases in 4–9 axillary lymph nodes or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary lymph node metastases pN2a: Metastases in 4–9 axillary lymph nodes (at least one tumor deposit > 2.0 mm) pN2b: Metastases in clinically detected internal mammary lymph nodes, pathologically negative axillary nodes
pN3	Metastases in 10 or more axillary lymph nodes; or in infraclavicular (Level III axillary) lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive Level I, II axillary lymph nodes; or in more than three axillary lymph nodes and micrometastases or macrometastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes; or in ipsilateral supraclavicular lymph nodes pN3a: Metastases in 10 or more axillary lymph nodes (at least one tumor deposit > 2.0 mm); or metastases to the infraclavicular (Level III axillary lymph) nodes pN3b: pN1a or pN2a, positive internal mammary nodes by imaging or pN2a in the presence of pN1b pN3c: Metastases in ipsilateral supraclavicular lymph nodes

Abbreviation: ITC, isolated tumour cells.

Table 2.3. Pathological definition of distant metastasis (M) based on AJCC cancer staging manual (8th edition).

Category	Criteria
M0	No distant metastases
M1	Any histologically proven metastases in distant organs; or if in non-regional nodes, metastases > 0.2 mm

Table 2.4. Anatomic stage group as defined by AJCC cancer staging manual (8^{th} edition).

Stage	T score	N score	M score
0	Tis	N0	M0
IA	T1	N0	M0
IB	T0	N1mi	M0
	T1	N1mi	M0
IIA	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
IIB	T2	N1	M0
	T3	N0	M0
IIIA	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1	M0
	T3	N2	M0
IIIB	T4	N0	M0
	T4	N1	M0
	T4	N2	M0
IV	Any T	Any N	M1