Sequence analysis of the *E-cadherin* (*CDH1*) gene in a cohort of gastric cancers seen in the Western Cape

By

ANELISA JACA

JCXANE001

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In (partial if the degree was by full dissertation) fulfilment of the requirements for the degree

MSc (Med) Anatomical Pathology

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

Date of Submission: 30th April 2013

Supervisor: A/Professor Richard Naidoo

Department: Clinical Laboratory Science

Declaration

I, Anelisa Jaca, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: ……

Date: …….

Contents

List of Figures v
Abstract vi
Acknowledgements vii
List of Abbreviations viii
CHAPTER 1 1
1. Introduction 1
1.1. Carcinogenesis 1
1.2. Tumour suppressor genes 3
1.3. Oncogenes 4
1.4. Environmental exposures and carcinogenesis 4
2. Background Review 5
2.1. Anatomy of the stomach 5
2.2. Histology of the stomach 5
3. Gastric Cancer 6
3.1. Worldwide Prevalence of gastric cancer 7
4. The risk factors of Gastric Cancer 8
4.1 *Helicobacter pylori* 8
4.2. Diet and its role in Gastric Cancer 11
4.3. Smoking and alcohol intake in gastric cancer 12
5. Epigenetic and Genetic factors 13
5.1. Epigenetic factors 13
5.2. Genetic factors 14
6. Cadherin Molecules 14
6.1. E-cadherin protein structure and function 14
7. The Wnt signalling pathway 15
7.1. β-catenin protein 16
7.2. Canonical or Wnt/β-catenin signalling pathway 16
7.3. The canonical signalling pathway and gastric cancer 18
7.4. The Non-canonical pathway 19
8. The *CDH1* gene structure 20
8.1. The *CDH1* gene and Familial Gastic Cancer 21
8.2. The *CDH1* gene and Sporadic Gastric Cancer (SGC) 22
8.3. *CDH1* Large genomic alterations and gastric cancer 24

8.4. Microsatellite instability in gastric cancer 24
8.5. *CDH1* Polymorphisms in gastric cancer 26
9. Aims and Objectives 27
CHAPTER 2 28
Materials and Methods 28
2.1. Ethical Approval 28
2.2. Case Selection 28
2.3. Pathologic Case Reports 28
2.3. Histology 29
2.4. Haematoxylin and Eosin (H & E) Staining 29
2.5. DNA Extraction 29
2.6. DNA quantification 30
2.7. Insulin PCR 30
2.8. Gel Electrophoresis 31
2.9. Primer Design 31
2.10. Gel Electrophoresis 33
2.13. Nested PCR 33
2.14. DNA Sequencing of PCR products 35
2.15. Statistical Analysis 35
CHAPTER 3 36
3. Results 36
3.1. Demographic and clinicopathologic characteristics of study participants 36
3.2. Descriptive statistics of demographic/clinicopathologic data 36
3.2. Microscopic examination of intestinal and diffuse gastric adenocarcinoma 40
3.3. Verification of DNA quality 40
3.4. E-cadherin (*CDH1*) gene amplification 41
3.5. Sequence analysis 44
3.6. Immunohistochemistry 45
CHAPTER 4 46
4. Discussion 46
4.1. Clinicopathologic features of tumours from gastric cancer patients 46
4.2. Sequence analysis of the *CDH1* gene 47
4.3. Association of *CDH1* mutations with Clinicopathologic data and protein expression 49
References 1

Appendices.....10
Appendix A: The chromatograms of the 26 cases 10
Appendix B: Preparation of reagents for Agarose Gel Electrophoresis 23
Appendix C: The demographic/clinicopathologic data of the 102 cases 24
Appendix D: Staging.....27

List of Figures

Figure 1: The H&E staining of the stomach walls, mucosa, submucosa and muscle layer. 6
Figure 2: A diagrammatic representation of the E-cadherin protein structure showing the different domains: extracellular (A), transmembrane (B) and cytoplasmic (C) domains. The diagram was amended from Mateus et al., 2009. 15
Figure 3: A diagrammatic representation of the Wnt signalling pathway. In the absence of the Wnt ligand, β-catenin binds the destruction complex (APC/Axin/GSK3), and then becomes phosphorylated and degraded. In the presence of the Wnt ligand, the destruction complex then becomes deactivated and permits the translocation of β-catenin to the nucleus (Catriona and Nusse, 2004). 18
Figure 4: Diagrammatic representation of the *E-cadherin* gene location on the long arm of chromosome 16 at region 2, band 2 and sub-band 1 (16q22.1) (http://ghr.nlm.nih.gov/gene=*CDH1*). 20
Figure 5: The distribution of different age groups in gastric cancer patients. 37
Figure 6: The relation between gender and the histologic types of gastric cancer is indicated. The diffuse type is commonly seen amongst female patients (61.11%) while the intestinal type is more frequent amongst the male patients (62.15%). 37
Figure 7: The distribution of histological types in gastric cancer patients. 38
Figure 8: A diagrammatic representation of the intestinal (**A**) and diffuse (**B**) types of gastric cancer. The intestinal type is characterized by glands and the diffuse type by single line dis-cohesive cells. 40
Figure 9: Representation of a 1.5% agarose gel showing amplification of exon 2 of the insulin gene which produces a PCR (236 bp) product. Lanes 1 to 10 represent the DNA samples; lane 11 positive (+) and lane 12 negative control (-); M, a 100 bp ladder. 41
Figure 10: Representation of a 1.5% agarose gel showing amplification of exon 8 of the CDH1 gene which produces a PCR (397 bp) product. Lanes 1 to 11 represent the DNA samples; lane 12 positive (+) and lane 13 negative controls (-); M, a 100 bp ladder. 42
Figure 11: Representation of a 1.5% agarose gel showing amplification of exon 9 of the CDH1 gene which produces a PCR (397 bp) product. Lanes 1 to 11 represent the DNA samples; lane 12 positive (+) and lane 13 negative controls (-); M, a 100 bp ladder. 42
Figure 12: A representative chromatogram of sample T68 showing sequencing data as obtained from Macrogene. 44
Figure 13: A representative chromatogram of sample T79 showing sequencing data as obtained from Macrogene. 44
Figure 14: The immunohistochemistry stains showing the E-cadherin protein expression (with DAB chromogen). Sample T68 (**A**) - showing membrane positivity. Sample T79 (**B**) - showing absence or loss of E-cadherin. 45

Abstract

Background: Gastric cancer is commonly seen in the Western Cape. There are numerous factors that contribute to the development of this cancer and these include both environmental and genetic factors. Amongst the genetic factors, the *E-cadherin* (*CDH1*) gene is said to play a major role in the development of gastric cancer. Both germline and somatic mutations have been identified in the *CDH1* gene. The germline mutations span the entire exon of this gene and are seen in approximately 29% to 56% of familial gastric cancers. On the other hand, *CDH1* somatic mutations are seen in approximately 31% of sporadic cancers. These mutations have been identified in exons 7 to 10, with exons 8 and 9 being the commonly mutated regions.

Objective: The primary objective of this study was to determine the prevalence of *CDH1* genetic mutations in gastric cancer in the Western Cape and to correlate these findings with the demographic/clinicopathologic data.

Design: A total of 102 gastric cancer cases were collected from the archives of the Division of Anatomical Pathology, University of Cape Town/NHLS Groote Schuur. DNA was then extracted from FFPE tissues and verified by amplifying exon 2 of the insulin gene. Following the insulin PCR, exon 8 and 9 of the *CDH1* gene were amplified. The PCR products were sequenced and the sequencing data was analysed using Bioinformatics tools (BioEdit and ClastalW).

Results: Good quality DNA was obtained from 44 cases insulin PCR, however, only 26 of the 44 samples were amplified for exon 8. *CDH1* mutations were identified in only 2 of the 26 (7.6%) cases for exon 8. Only 2 samples were amplified for exon 9, but the sequencing reaction was unsuccessful. The mutational data was correlated with clinicopathologic features; however no significant conclusions could be reached due to the low frequency of mutations in this study.

Conclusion: This study showed that there was a low prevalence of *CDH1* genetic mutations in our study cohort.

Acknowledgements

So many people played a very important role during my MSc. I would like to express my gratitude for their support, although I haven't mentioned them by name, I am extremely grateful for all their assistance.

I would like to express my appreciation to Professor Richard Naidoo for his wonderful supervision, his support and for believing in me. I also thank Dr Michael Locketz and Professor Dhiren Govender for helping me understand pathology.

I would like to express my gratitude to Mr Henry Carrara, a statistician whom I was consulting for the study. Not forgetting my friends and colleagues (Subash Govender, Phumza Phillips and Raymond Kriel) from Anatomical Pathology for their support, I appreciate you all so much.

I would like to express and extend my sincerest gratitude to Aron Abera for his support and mentorship with primer design and the sequencing assay.

Last but not least, I would like to extend my gratitude to NRF, UCT and CANSA for funding this study.

I dedicate this master's dissertation to my family (biological and spiritual). Thank you so much for your love and support, you kept me strong throughout. I really love and appreciate you.

Moreover, I would like to thank Almighty God, the author of wisdom, for his omnipotence.

List of Abbreviations

ºC – degrees Celsius

16q – long arm of chromosome 16

A – Adenine

ABL– Abelson murine leukemia viral oncogene homolog 1

APC–Adenomatous polyposis coli

Asn – Asparagine protein

BAX – BCL2-associated X protein

BCL2 – B-cell lymphoma 2

bp – base pairs

C – Cytosine

cagA – Cytotoxin-associated gene

CDK – Cyclin dependent kinase

CDH1 – E (epithelial)-cadherin gene

c-MYC – cellular myelocytomatosis oncogene

CpG– Cytocine-phosphate guanine

DGC – Diffuse gastric cancer

DNA – deoxyribonucleic acid

DNMTs – DNA methyltranferases

dNTPs – dinucleotide triphosphate

DSBs – Double strand breaks

EC – Extracellular repeats

E-cadherin – Epithelial-cadherin

ERBA – Erythroblastosis virus avian

ERBB2 – v-erb-b2 Erythroblasticleukemia viral oncogene

FFPE – Formalin-fixed paraffin-embedded tissue

G6PD - Glucose-6-phosphate dehydrogenase

G – Guanine

Gly – Glycine protein

H & E – Haematoxylin-eosin

HDGC - Hereditary diffuse gastric cancer

H.pylori –Helicobacter pylori

HREC – Human Research Ethics Committee

IM – Intestinal metaplasia

MSI – Microsatellite instability

µl – microliter

MET – Mesenchymal epithelial transition factor

MLH1 – MutL homolog 1

MLH2 – MutL homolog 2

MMP-7 - Matrix metalloproteinase-7

MMR – Mismatch repair

MUCI -Mucin 1

MYB – Myeloblastosis gene product

MYCN – v-mycMyelocytomatosisyiral related oncogene neuroblastoma derived avian

MYCL - v-mycMyelocytomatosisyiral related oncogene homolog 1, lung carcinoma derived avian

N-cadherin – Nueral-cadherin

NER – Nucleotide excision repair

NF1- Neurofibromatosis type 1

ng – nanogram

NHLS – National Health Laboratory Service

ONP –Osteopontin

p53 – Tumour suppressor gene product 53

P-cadherin – Placental-cadherin PCR – Polymerase Chain Reaction

pmol - Picomol

RAS – Rat Sarcoma

RET – Ret proto-oncogene

SGC- Sporadic gastric cancer

SNPs – Single nucleotide polymorphisms

STK11 – Serine/threonine kinase 11

T – Thymine

Taq – Thermusaquaticus

TBE – Tris-Borate-EDTA

THBS4 - Thrombospondin 4 glycoprotein

TRK – Tyrosine kinase

TSGs – Tumour suppressor genes

UV - Ultraviolet

V – Volts

vacA – Vacuolatingcytotoxin gene

Val – Valine

WT1–Wilms tumour 1 gene

CHAPTER 1

1. Introduction

1.1. Carcinogenesis

Carcinogenesis is an uncontrolled cell growth and survival of abnormal cells (Yamamoto et al 2011; Zhang et al., 2008), which consists of initiation, promotion and progression stages. The consequence is cell accumulation, loss of tissue differentiation, loss of cellular adhesion, nuclear enlargement, enhanced mitotic activity and eventually metastasis (Coup et al., 2004; O'Leary et al., 2011). Cancer cells can invade the blood vessels and metastasise to various organs in the body. Secondly, they can invade the lymphatic vessels and spread to the lymph nodes. Cancer cells are classified according to its appearance or shape on a cell surface. They could have a sessile, papillary, polypoid, fungating, ulcerated and/or annular appearance. Cancer is also distinguished in accordance with the type of cell or tissue it originates from, namely the epithelial and connective tissues. Epithelial malignant tumours are termed carcinomas whilst those of connective tissues are called sarcomas.

Furthermore, epithelial tissues are either glandular or non-glandular. The non-glandular cancerous tumours are named using the epithelial cell type e.g. squamous and transitional cell carcinomas. Conversely, the glandular tumours are specified as adenocarcinomas and are accompanied by the type of tissue they arise from. Tumours also resemble the cells or tissue type from which they originate. This is termed tumour differentiation. They could be well, moderately or poorly differentiated (Hottenrott, 2012). The well differentiated tumours almost resemble the tissues of their origination and are less aggressive. The poorly differentiated ones are totally different from their tissues of origination. The moderately differentiated malignancies lie in between the above two types. As a result, the poorly differentiated tumours are more difficult to diagnose than the well and moderately differentiated types.

Several factors play a significant role in the initiation, promotion and progression of cancer. These include deregulation of molecular pathways, genetic instabilities and environmental exposures (Cano, 2008). One of the molecular pathways is the cell cycle which is a mechanism by which cells divide. It is divided into gap phase 0 (G0), gap phase 1 (G1), DNA synthesis (S), gap phase 2 (G2), and mitosis (M). This pathway is composed of proteins i.e. cyclins (G1/S cyclins, S cyclins, M cyclins, G1 cyclins) and cyclin-dependent kinases (CDK). These molecules form complexes and modulate the progression from one stage of the cell cycle to another (Vermeulen et al., 2003). Checkpoints are identified within the cell cycle stages to ensure that events such as DNA synthesis proceed accordingly. If there is any DNA damage, the checkpoints arrest the cell cycle until the problem is repaired (Collins et al., 1997). Alterations or disturbances in this cycle may therefore lead to cancer development.

Molecules which regulate the G1/S phase such as proto-oncogenes and tumour suppressor genes may undergo alterations and hence contribute towards cancer formation. Additionally, checkpoints such as those responsible to check DNA damage and modulate DNA replication may be genetically compromised and then trigger the development of cancer (Bartek et al., 1999).

Apoptosis, also known as programmed cell death is a mechanism required for embryonic growth and tissue homeostasis. It is activated in response to DNA damage and abnormal cell growth and development (Renehan et al., 2001). There are a number of genes that are involved in the regulation of apoptosis. The key genes include the tumour suppressor gene *p53,* anti-apoptotic gene *bcl-2* and the pro-apoptotic gene *bax* (Sjostrom and Bergh, 2001). Furthermore, these genes may accumulate mutations and disrupt the apoptotic process, resulting to tumour initiation, progression and metastasis (Lowe and Lin, 2000; Carnero, 2002).

The majority of human malignancies develop due to genetic instabilities (Langauer et al., 1998). These genetic instabilities involve DNA damage and repair (Diaz, 2008). DNA may be damaged as a result of a number of abnormalities such as alterations in the

gene and chromosome number. Additionally, DNA may also be impaired due to radiation, namely, X-rays or gamma-rays. This results in DNA double strand breaks (DSBs) which in turn leads to chromosomal abnormalities (Okayasu, 2011). Chromosome translocation is another aberration that occurs due to DNA damage (Langauer et al., 1998). These DNA abnormalities may create mutations that promote carcinogenesis (Finn et al., 2012).

DNA repair is a mechanism that corrects errors such as repeat sequences, mismatches and strand breaks which may occur during DNA replication (Madhusudan and Middleton, 2005). The two important repair mechanisms are the nucleotide excision (NER) and (MMR) mismatch repair (Langauer et al., 1998). The NER system serves to eradicate the DNA damaged as a result of ultraviolet light. On the other hand, during DNA synthesis the MMR gets rid of mismatched bases or DNA loops that arise as a result of DNA slippage at microsatellites (Leibeling et al., 2006). When DNA repair genes accumulate mutations it gives rise to the development of cancer through this mechanism (Hoeijmakers, 2001).

1.2. Tumour suppressor genes

Tumour suppressor genes (TSGs) are a group of molecules which are involved in the modulation of biological processes such as cell proliferation, cell death and cell invasion. They can lead to the formation and progression of cancer when their function is lost (Weinberg, 1989; Macleod, 2000; Lai et al., 2012). Their inactivation involves a mechanism called the Knudson "two" hit hypothesis. Therefore, for tumour suppressor genes to give rise to cancer, two hits are required to totally deactivate both gene copies. Moreover, these genes are known to cause hereditary type of cancers (Collins et al., 1997).

The most important TSGs recognized in human cancers include the *p53*, *retinoblastoma* (*Rb*), *APC*, *WT1* and *NF1*. However, the most common gene involved in more than half of the cancers is the *p53* gene (Hussain and Harris, 1998). Other TSGs include *PTEN*, *BRCA1* and the *E-cadherin* gene (Christofori and Semb, 1999; Lai et al., 2012).

1.3. Oncogenes

Proto-oncogenes play an important role in cellular growth and proliferation. When activated, they act as oncogenes and therefore promote tumour formation and progression (Lehman et al., 1991). Only a single gene copy of the gene needs to be mutated in order to give rise to cancer, therefore the Knudson "two" hit hypothesis does not apply in this regard (Grander, 1998). A number of proto-oncogenes such as; *TRK*, *MET*, *RET*, and *KIT* have been identified in cancer formation (Sanchez-Carbayo and Cordon-Cardo, 2002). Other genes include the *c-MYC, ABL, RAS, MYB, ERBA, ERBB2, MYCN* and *MYCL* gene (Schwab, 1998).

1.4. Environmental exposures and carcinogenesis

There are various environmental factors that impact on carcinogenesis. These include; infections, lifestyle and chemical exposures. Pathogenic infections have been identified as a cause of malignant tumours. These include bacterial, viral and fungal infections (Correa, 2003). Some of the viral pathogens involved in carcinogenesis include the Epstein-Barr Virus (EBV), Human papillomavirus (HPV), Human herpes virus (HHV) and Hepatitis C virus (Stockfleth et al., 2004; Mclaughlin-Drubin and Munger, 2008; Hino et al., 2009). Bacterial pathogens such as *Helicobacter pylori, Mycoplasma*, *Chlamydia pneumonia*, *Haemophilus influenza* and *Staphylococcus* strains have also been implicated in carcinogenesis (Apostolou et al., 2011; Vogelmann and Amieva, 2007). Lifestyle factors include diet, smoking and alcohol consumption. Various chemicals have also been implicated as a causative agent for cancer development. These include; gasoline, glass, wool, ethylene oxide and radon gas (Huff, 1993).

2. Background Review

2.1. Anatomy of the stomach

The stomach is located in the upper compartment of the abdominal cavity. This organ lies beneath the diaphragm and liver and is sub-divided into four parts, namely the cardia, fundus, body and the pylorus. The pylorus is further partitioned into three areas including the pyloric antrum, pyloric canal and pylorus (Tortora and Derrickson, 2011). The stomach has two curvatures, namely the lesser and the greater curvatures. The right superior part of the stomach wall forms the lesser curvature, whereas the greater curvature is formed by the left inferior portion (Tortora and Derrickson, 2011).

There are sphincters that control the passage of food into and out of the stomach. The lower oesophageal or cardiac sphincter modulates the opening of the oesophagus into the stomach. The pyloric sphincter on the other hand, regulates the opening from the pyloric part into the first part of the duodenum (Tortora and Derrickson, 2011).

2.2. Histology of the stomach

The stomach wall has four layers: the mucosa which lies on the surface, submucosa, muscularis externa and serosa. The mucosa is composed of surface mucus secreting simple columnar epithelial cells also referred to as mucous cells. Furthermore, the mucosa consists of an areolar connective tissue called the lamina propria. The mucosa is separated from the submucosa by a smooth muscle layer called muscularis mucosae (Tortora and Derrickson, 2011). The gastric glands are located within the mucosa which extends down into narrow paths called gastric pits. These glands consist of three cell types depending on where they are located in the stomach. These are the mucous neck cells, chief cells and parietal cells. The submucosa of the stomach consists of areolar connective tissue. Within the muscularis externa layer, there are three layers of the smooth muscle: an outer longitudinal layer, middle circular layer and an inner oblique layer.

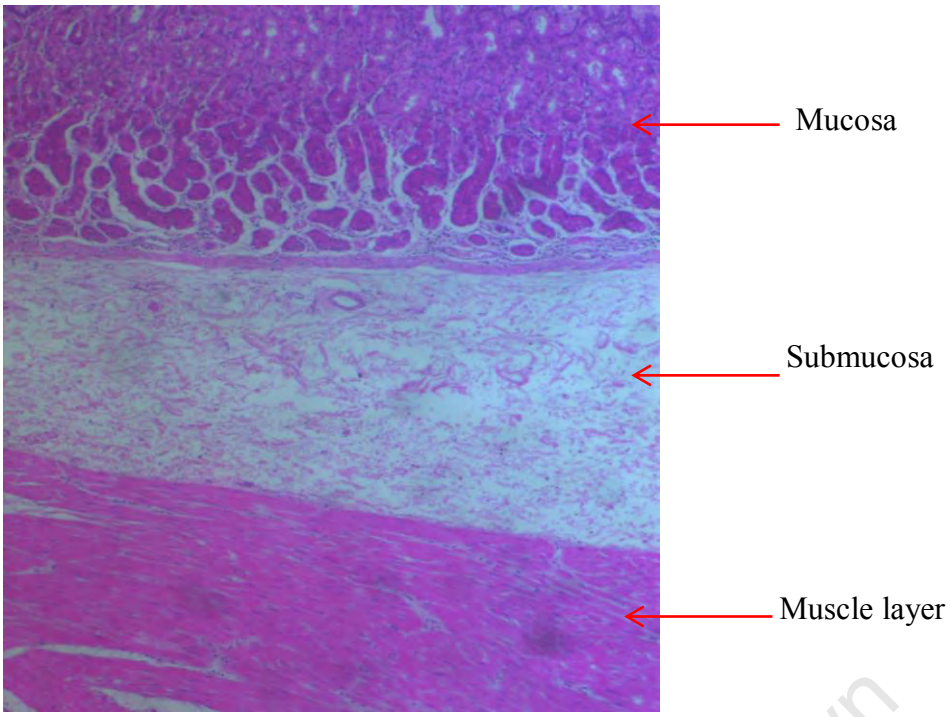


Figure 1: The H&E staining of the stomach walls, mucosa, submucosa and muscle layer.

3. Gastric Cancer

Gastric carcinoma is described as a malignant tumour that develops from the epithelial tissue of the stomach (Forx and Wang, 2007). This malignancy is histologically classified into two forms, namely the diffuse and intestinal types according to the Lauren Classification system (Hackenson et al., 2010).

The intestinal type developing from epithelial cells of the mucosa typically forms glands (Yamamoto et al., 2011). It is more prevalent in high risk areas and occurs most frequently in older people. It occurs primarily in males and is characterized by a slow growth, and is often highly differentiated (Corso et al., 2012). Further, it involves the distal stomach and proceeds from normal mucosa through atrophic gastritis, intestinal metaplasia, dysplasia and eventually progressing to cancer (Kudo et al., 2011). These cancer cells usually demonstrate cell adhesive properties and may resemble either a sheet like or nested structure (Padmavathy et al., 2011).

The diffuse type on the other hand has an early age of onset and commonly occurs in low risk areas (Hackenson et al., 2010; Chen et al., 2011; Corso et al., 2012). It occurs in younger people, commonly females, and is characterized by enlarged dispersed cells which have abundant cytoplasm and prominent nucleoli (Correa and Piazuelo, 2011). This type of gastric cancer invades the abdominal wall, and proximal stomach, especially the cardia. Symptomatic individuals present with abdominal symptoms such as epigastric pain, early satiety and weight loss (Chen et al., 2011; Foster et al., 2011). Additionally, it may present as signet ring adenocarcinoma, usually together with lymph node involvement and distant metastasis (Chen et al., 2011).

3.1. Worldwide Prevalence of gastric cancer

Gastric cancer is rated as the fourth most frequent malignancy in the world with 988,000 cases per annum. It is the second leading cause of cancer mortality with 736,000 deaths per year, worldwide. The global estimated new cases and deaths amongst men and women are (640,000 cases and 463,000 deaths); and (384,000 cases and 273,000 deaths), respectively. Incidence and death rates, however, differ greatly amongst various populations. Stomach cancer is less frequent in more developed regions such as Europe, Australia, New Zealand and North America with 274, 000 cases and 180,000 deaths per annum. High incident regions include developing areas; Africa, Eastern Asia and South America with 713,000 cases and 555,000 deaths per year (Globacan, 2008).

The highest incidence rates (713,000) are seen in developing countries with 467,000 cases and 353,000 deaths in men, whilst in women, there are 246,000 cases and 202,000 deaths. On the other hand, prevalence and mortality rates are low in developed countries (173,000 cases and 110,000 deaths in men; 101,000 cases and 70,000 deaths in women). In general, in countries like China there are 463,000 cases and 352,000 deaths; central and eastern Europe (165,000 cases and 133,000 deaths) and Western Pacific regions (621,000 cases and 432,000 deaths). Further, Canada and Korea have increased incidence and death rates per year (Globacan, 2008).

4. The risk factors of Gastric Cancer

There are several factors that contribute to the development of gastric adenocarcinoma and these include environmental, genetic and epigenetic agents. Environmental factors include *Helicobacter pylori* (*H. pylori*) infection, diet, smoking and alcohol consumption. Diet and *H. pylori* infection are the major contributing factors.

4.1 *Helicobacter pylori*

4.1.1. Helicobacter pylori (H. pylori) Morphology

H. pylori is a microaerophilic gram negative bacterium with a length of 2 to 4 µm and a width of 0.5 to 1 µm (Serrano et al., 2007). This bacterium is spiral in shape but can also resemble a rod (Kim et al., 2011). The outer membrane possesses 2 to 6 sheathed unipolar flagella with a length of about 3 µm, which allow motility (Correa and Piazuelo, 2011) especially in a mucus environment of the stomach (Kusters et al., 2006). *H. pylori* also has virulence factors which contribute to the pathogenesis of gastroduodenal disease such as peptic ulcers, atrophic gastritis and gastric cancer (Matsunari et al., 2012). These factors include the *Cag* pathogenecity island (PAI); CagA protein, vacuolating toxin (VacA); blood group antigen binding adhesin (BabA) and the (dupA) duodenal ulcer producing gene (Wen and Moss, 2009).

4.1.2. H. pylori Colonization Mechanism and Gastric cancer

H. pylori was categorized by the World Health Organization (WHO) as a class I carcinogen due to its significant role in the pathogenesis of gastric adenocarcinoma (Kandulski and Malfertheiner, 2008). This bacterium has been predicted to infect approximately half of the population, globally (Bartchewsky et al., 2009). It is transported through fecal and oral contamination amongst families during childhood (Ruggiero, 2012). The *H. pylori* bacterium surface receptors, namely, BabA attaches or binds to the fucosylated Lewis B-binding (Leb-binding) antigen on the mucosa or

epithelial tissue of the stomach (Stoicov et al., 2004; Forx and Wang, 2007).

Once it binds to the epithelium, the virulent factors namely; CagA and VacA colonise the gastric mucosa (Correa and Piazuelo, 2011). This is followed by the release of a urease enzyme which catalyses urea and then produces ammonia and carbon dioxide (Stingl et al., 2002). The ammonia creates an alkaline environment which allows the bacterium to thrive in the acidic stomach. The ammonium compound also helps the *H. pylori* to persistently colonize the cell surface of the gastric mucosa and initiate disease (Burne and Chen, 2000).

The bacterium persistently infiltrates gastric mucosa and induces a chronic inflammation. The gastric mucosa becomes wounded as a result of the inflammation, thereby reducing gastric glands and causing atrophy. A different type of epithelial tissue called the intestinal metaplasia (IM) is then regenerated by the damaged mucosa (Kudo et al., 2011). An atrophic gastritis together with IM is followed by dysplasia and eventually leads to the intestinal type of gastric cancer. Diffuse gastric cancer on the other hand develops from atrophic gastritis in the absence of IM (Kim et al., 2010).

4.1.3. H. pylori infection in gastric mucosa - associated lymphoid tissue (MALT) lymphoma

Gastric lymphoma develops in the stomach as a result of chronic gastritis induced by the *H. pylori* bacterium. Gastric lymphoma is categorised as the MALT, diffuse large B cell and T cell lymphoma which is not commonly seen in the stomach. The MALT lymphoma is specified as the low or high grade lymphoma. The low grade lymphoma is treated by eradicating the *H. pylori* bacterium as opposed to the high grade lymphoma, which is treated by chemotherapy. The genetic host factors of the *H. pylori* bacterium may play a significant role in the pathogenesis of this lymphoma. This bacterium has been identified in approximately 98% patients with gastric MALT lymphoma. However, only a small percentage of people with *H. pylori* infection develop gastric MALT lymphoma (Witkowska and Smolewski, 2013).

4.1.4. Intestinal metaplasia (IM) and gastric cancer

Intestinal metaplasia is regarded as a precancerous lesion of gastric cancer. This condition develops from gastric stem cells forming epithelial cells (absorptive cells, goblet cells and Paneth cells), (Erkan et al., 2012). There are three categories of IM: types I, II and III (Eriksson et al., 2008). The type I IM has a low pathogenic effect on stomach cancer and can also be called complete IM (Leung and Sung, 2002). Further, the type I IM histologically resembles the mucosa of the small intestine and is characterized by goblet cells which produce sialomucins, Paneth and absorptive cells that are positioned below the brush border (Li et al., 2010).

On the other hand, type II and III IM increase the degree of malignancy and can also be referred to as the incomplete IM. Incomplete IM is histologically distinguished by different sizes of multiple mucus vacuoles and the absence of a brush border (Garay et al., 2004). In addition, this IM exhibits goblet cells which produce sulfomucins, however, it must be noted that Paneth cells are absent (Kang et al., 2011).

4.1.5. Helicobacter pylori infection in Gastric Cancer

The incidence of *H. pylori* infection has been investigated amongst patients with gastric adenocarcinoma in a number of studies (Watabe et al., 2005; Nam et al., 2011; Tsukanov et al., 2011). The infection has been identified in both diffuse and intestinal histological types and seen in non cardia gastric cancer (Lamb and Chen 2012). This infection plays a very important role in the progression of gastric cancer and promotes invasion and metastasis of the cancer (Parsonnet et al., Qiu et al., 2010; 1991; Kang et al., 2011). However, the eradication of *H.pylori* reduces the risk of developing stomach cancer. In a study that was carried out in different populations, *H.pylori* infection was eradicated and this was accompanied by a risk reduction rate of gastric cancer (10% in rural areas and 6% in urban locations) (Yeh et al., 2009).

4.2. Diet and its role in Gastric Cancer

Foods and beverages containing carcinogenic chemicals such as N-nitroso compounds and aromatic amines play a crucial role in tumour growth. These compounds induce alkylation and cause DNA damage (Lijinsky, 1999). Additionally, low or poor ingestion of fruit and vegetables and a high intake of salty foods increase the risk of gastric carcinogenesis. The method of food preservation and storage plays an important role in cancer development. Food smoking, as a method of preservation and less refrigeration is an important contributing factor (Yalcin, 2009). The use of refrigerators decreases the production of nitroso compounds and also sustains antioxidants in food (Park et al., 2011). Additionally, the use of salt when preserving food contributes towards gastric carcinogenesis by degrading the mucosal layer and causing inflammation in the stomach (Tsugane and Sasazuki, 2007). Nutrient deficiencies of folate, betanine, choline, zinc, vitamin C, selenium, vitamin B6 and vitamin B12 can also contribute towards gastric carcinogenesis (Su, 2012).

Numerous studies have reported the correlation between diet and gastric cancer (Ngoan et al., 2002; Kelley and Duggan, 2003; De Stafani et al., 2004; Nan et al., 2005). It was shown that pickled, salted and fatty foods significantly enhanced the risk of cancer development. This study involved 116 patients and also showed that high intake of fruit and vegetables on the other hand decreased the risk of stomach cancer (Ngoan et al., 2002). Another study with 240 subjects also discovered that a high consumption of starchy foods such as rice; white bread; potatoes and tubers played a significant role in the development of this cancer. Further, it was also concluded that diets high in vegetables and fruits reduced the risk of developing stomach cancer (De Stafani et al., 2004).

Malnutrition is an important contributing factor to the development of gastric cancer. An association between malnutrition and the risk of stomach cancer was investigated in 100 gastric cancer patients. This study showed a significant association between malnourished individuals and risk of gastric cancer (Gavazzi et al., (2011). Another dietary factor contributing towards the development of stomach cancer is heme iron. A

study carried out in 444 individuals to investigate the impact of heme iron showed that there was a 70% greater chance of developing stomach cancer in individuals with an increased heme iron in their diet (Jakszyn et al., 2011).

4.3. Smoking and alcohol intake in gastric cancer

Alcohol intake and cigarette smoking have also been implicated in the development of stomach cancer. The impact of cigarette smoking in gastric carcinogenesis was determined in a population study consisting of African Americans, Japanese Americans, Latino Americans, Native Hawaiians and Whites. Amongst these different populations, cigarette smoking increased the risk of developing this cancer in 98% of men and 78% of women (Nomura et al., 2012). In another study consisting of 391 cases, 68.79% heavy smokers were investigated and had about 59% increased risk of stomach cancer (Moy et al., 2010).

In addition, alcohol may cause abnormalities during DNA synthesis and repair and thus contribute to the development of this tumour (Zhang et al., 2007). A study cohort consisting of 444 cases showed that an alcohol consumption of greater than 60 grams per day was highly associated with developing gastric tumour. On the other hand, individuals who consumed less than 60 grams per day had a low risk of developing the cancer (Duell et al., 2011). Heavy alcohol consumers had an approximately 50% increased risk of developing the stomach cancer than individuals who consumed less (Moy et al., 2010).

5. Epigenetic and Genetic factors

5.1. Epigenetic factors

Epigenetic factors modulate gene expression and function and do not influence the DNA coding gene sequences. They include DNA methylation and histone modification (Su, 2012). Approximately 90% of epigenetic changes are responsible for the development of cancer (Compare et al., 2011).

5.1.1. DNA methylation

DNA methylation is the addition of methyl groups to the promoter region of a gene. The methyl groups are added at the 5'-carbon position by enzymes called DNA methyltranferases (DNMTs) and form 5-methyl-cytosine (Zhao and Bu, 2005). These enzymes; DNMT3A, DNMT3B and DNMT1 initiate and maintain methylation (Tomita et al., 2010). The function conducted by these enzymes is vital for embryonic development (Williams, et al., 2011). DNMT3A and DNMT3B initiate methylation, while DNMT1 maintains it (Das and Singal, 2004). Genes can either be hypermethylated in the CpG dinucleotides e.g. tumour suppressor genes or hypomethylated e.g. oncogenes (Dumitrescu, 2012).

The CpG dinucleotides are, in most cases unmethylated in normal cells (Jones and Baylin, 2002). It is when malignancy is present that methylation is observed. Gastric carcinoma is one of the malignant tumours where the CpG island is commonly methylated (Kang et al., 2003). The methylation status of some suppressor genes namely; *E-cadherin* (*CDH1*), *APC, COX-2, DAP-kinase, GSTP1, hMLH1, MGMT, p16, p14, RASSF1A, THBS1* and *TIMP3* was investigated in gastric cancer (Kang et al., 2003). The *CDH1* was one of the frequently methylated genes in 60.1% of the gastric cancer tissue samples (Kang et al., 2003).

5.2. Genetic factors

Genetic instabilities namely; germline and somatic mutations, microsatellite instability and polymorphisms contribute to the risk of malignant formation. These instabilities are commonly identified in certain genes in patients with stomach cancer (Horri et al., 1992; Richards et al., 1999; Yamada et al., 2011).

6. Cadherin Molecules

Cadherin molecules are a large family of transmembrane proteins that regulate cell-cell attachment during development. They depend on calcium ions for their function and are composed of at least three major subfamilies, namely classical cadherins, desmosomal cadherins and protocadherins, (Van Roy and Berx 2008). These calcium ions give the molecules a rigid structure and therefore protect them from proteolysis (Pokutta and Weis 2007). The name "cadherin" is used to classify cell-cell adhesion molecules in different tissues e.g. E- (Epithelial) cadherins, N- (Neural) cadherins, P-(Placental) cadherins (Van Roy and Berx 2008). All the cadherin proteins are composed of multiple extracellular cadherin repeats (EC), a single transmembrane and cytoplasmic region. These domains interact with the cytoskeleton and are therefore responsible for the cell adhesion of cadherin molecules (Nollet et al., 2000).

6.1. E-cadherin protein structure and function

The E-cadherin protein (Figure 2) plays a significant role in maintaining cell differentiation and normal architecture of epithelial tissues (Shore and Nelson, 1991; Suriano et al., 2003). It belongs to the family of classical cadherin molecules and consists of three domains i.e. 5 large extracellular repeats (EC1 to EC5) where the N-terminal region is situated. Further, it also consists of an intracellular or transmembrane and a cytoplasmic domain (Graziano et al., 2003).

The C-terminal region is located in the cytoplasmic domain which adheres to the cytoskeleton through β, α and γ-catenins (Yamamoto et al., 2011). Additionally, this molecule consists of a juxtamembrane domain where the p120-catenin binds, preventing this molecule from degradation (Nollet et al., 2000). This transmembrane glycoprotein is involved in a cell signalling pathway namely, the Wnt signalling pathway and has been implicated in the development of certain tumours (Berx et al., 1995).

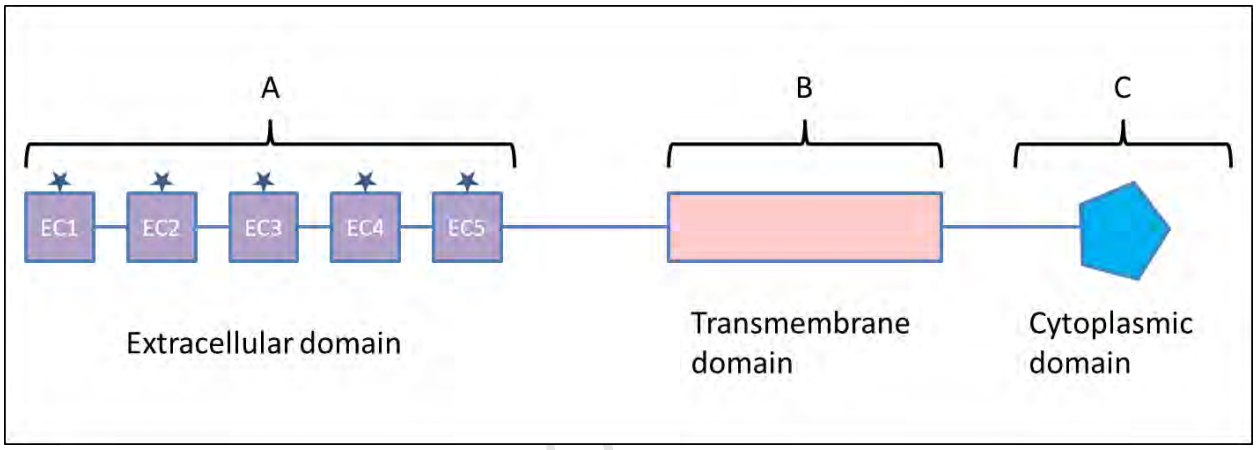


Figure 2: A diagrammatic representation of the E-cadherin protein structure showing the different domains: extracellular (A), transmembrane (B) and cytoplasmic (C) domains. The diagram was amended from Mateus et al., 2009.

7. The Wnt signalling pathway

The Wnt signalling is a cellular pathway that consists of 19 Wnt genes (Wnt1, Wnt2, Wnt2B, Wnt 3, Wnt 3A, Wnt4, Wnt 5A, Wnt 5B, Wnt6, Wnt7A, Wnt7B, Wnt 8A, Wnt 8B, Wnt 9A, Wnt 9B, Wnt 10A, Wnt 10B, Wnt 11 and Wnt16) in the human genome. It plays an important role in embryonic development (Kawakami et al., 2013; Zha et al., 2012). The first genes that were discovered in this signalling were the *int1* (*intergration1*) from mouse tumour cells and its homologous gene, the *Wingless* gene from *Drosophila melanogaster (*Klaus and Birchmeier 2008). The Wnt signalling pathway is activated via two pathways namely, the canonical or Wnt/β-catenin and non-canonical pathways. The canonical pathway consists of glycoprotein signalling

molecules such as ß-catenin, frizzled, APC, axin, GSK-3ß, TCF, c-MYC, Cyclin D1 and MMP-7 (Duacas et al., 2005).

7.1. β-catenin protein

This 92 kDa molecule is encoded by the *CTNNB1* oncogene and consists of 781 amino acids (Van Aken et al., 2001; Micu et al., 2010). It consists of a N-terminal region of about 150 amino acids, a C-terminal region with 100 residues and a central domain with 12 armadillo (arm) repeats. These regions are where the E-cadherin and catenins bind to form the E-cadherin-catenin complex. This complex enhances cell-cell adhesion. (Pokutta and Weis 2007).

7.2. Canonical or Wnt/β-catenin signalling pathway

The canonical pathway (Figure 3) is also referred to as the Wnt/ß-catenin pathway and plays a crucial role in the pathogenesis of cancer (Akaboshi et al., 2009). It is activated when the Wnt ligands bind to frizzled (Fz) and lipoprotein-related (LRP 5 & 6) receptors in the plasma membrane. This binding leads to the phosphorylation of a dishevelled protein (Polakis, 2000). The dishevelled protein then deactivates the destruction complex (Axin-GSK3β-APC complex) that is responsible for degrading ß-catenin. Since the destruction complex is deactivated, the ß-catenin protein becomes abundant in the cytoplasm and then translocates to the nucleus (Aberle et al., 1997). It then binds to T-cell factors (TCFs) and lymphoid enhancer factor (LEF) molecules (Korswagen, 2002).

After adhering to TCFs, it regulates transcription of its target molecules namely, Cyclin D1, matrix metalloproteinase-7 (MMP-7), immunoglobulin transcription factor 2 (ITF-2) and c-Myc (Doucas et al., 2005). These molecules then modulate the proliferation and differentiation of cells (Sugimura and Li, 2010). Further, this pathway also involves the interaction of the E-cadherin with catenin molecules such as beta (β), alpha (α) and

gamma (γ) catenins. The α-catenin which is also known as plaglobin, in turn interacts with the cytoskeleton for cell-cell adhesion (Van Aken et al., 2001). This complex forms a significant part of adherence junctions and regulates cell-cell adhesion. When the adhesive function of this complex is compromised, cells lose adherence to one another. This therefore allows reduced cell differentiation which is followed by the invasion and metastasis of tumour cells (Oliveira et al., 2009).

In the absence of Wnt ligands, the molecules (Axin-GSK3β-APC) within the destruction complex phosphorylate ß-catenin protein, marking it for degradation (Korswagen, 2002). The ß-catenin molecule then becomes degraded by an E3 ubiqutin ligase, β-TrCP, inhibiting the transcription of Cyclin D1, MMP-7, ITF-2 and C-myc molecules (Kawakami et al., 2012).

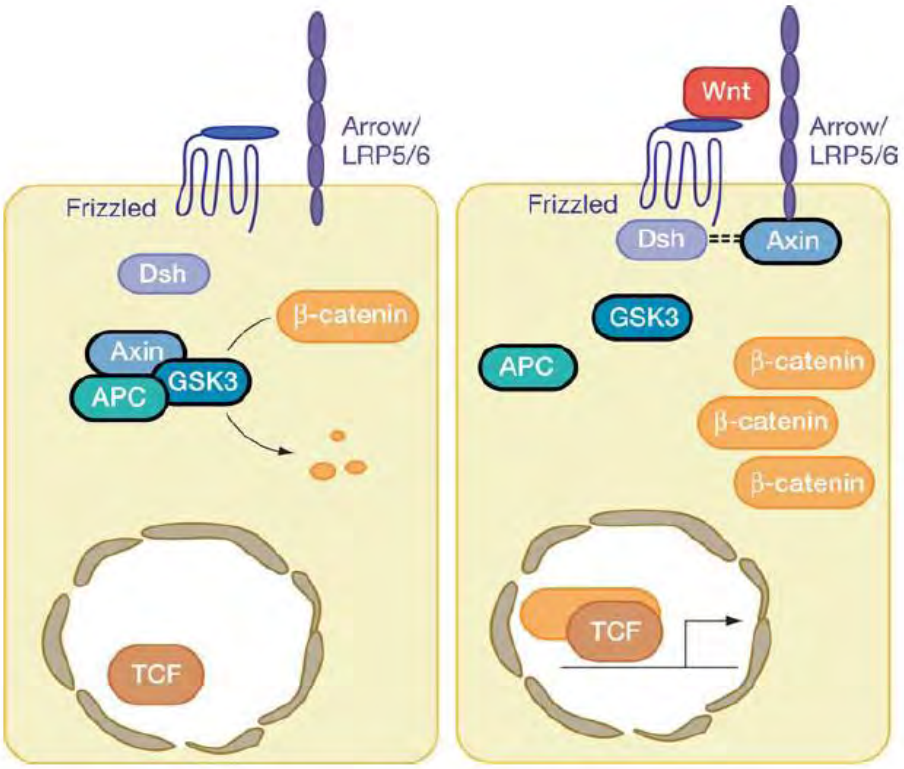


Figure 3: A diagrammatic representation of the Wnt signalling pathway. In the absence of the Wnt ligand, β-catenin binds the destruction complex (APC/Axin/GSK3), and then becomes phosphorylated and degraded. In the presence of the Wnt ligand, the destruction complex then becomes deactivated and permits the translocation of β-catenin to the nucleus (Catriona and Nusse, 2004).

7.3. The canonical signalling pathway and gastric cancer

This pathway contributes to gastric carcinogenesis. The abnormal expression of β-catenin protein has been reported to increase the development of gastric cancer (Li et al., 2008). Further, this protein is seen in the cell membrane in normal gastric mucosa, while positioned in the nucleus in gastric cancer (Clements et al., 2002). Also, the changes encountered in the E-cadherin-β-catenin complex plays a crucial role in the development of stomach cancer. A study by Joo et al., (2001) investigated the expression of the four adhesion molecules (E-cadherin, β, α, and γ-catenins) in early and advanced gastric cancer patients and found an abnormal expression of the proteins investigated in diffuse gastric cancer individuals (Joo et al., 2001).

Another study assessed the expression of the aforementioned adhesion proteins in 91 advanced gastric cancers. This study showed a reduced expression of β-catenin associated with site and depth of tumour invasion. Additionally, the E-cadherin and β-catenin aberrant expression was seen in patients with metastasis to the lymph nodes (Czyzewska et al., 2010). In 2004, it was shown among 60 patients that accumulation of β-catenin to the cytoplasm and nucleus increases the risk of gastric adenocarcinoma. This protein was also localized in the cytoplasm and nucleus in 41.7% and 48.3% patients, respectively (Song et al., 2004).

Chan and colleagues assessed the expression of the three proteins (E-cadherin, β-catenin and α-catenin) in 40 patients. Of these cases, 30% had a decreased expression of the three molecules. This study showed that the reduced expression of the E-cadherin-catenin complex enhanced the risk of gastric adenocarcinoma (Chan et al., 2003). It was also determined in another research that E-cadherin, α-catenin and β-catenin was aberrantly expressed in 82%, 85% and 88% of 65 gastric carcinomas, respectively. The abnormal expression of the proteins was associated with the depth of invasion of the tumours (Guzman et al., 2006). In another study, the abnormal expression of these molecules (E-cadherin, α-catenin, β-catenin and γ-catenin) was associated with clinicopathologic features amongst 65 patients. The abnormal expression of the four proteins was seen in 24.6% and was associated with poorly differentiated and diffuse type of tumours (Joo et al., 2000).

7.4. The Non-canonical pathway

The non-canonical pathway is also known as the planar cell polarity and or Wnt Ca^{2+} pathway (Tada et al., 2002). This pathway plays a pivotal role in development and homeostasis. It is involved in the migration of neural crest cells during development (De Calisto et al., 2005). Further, it preserves stem cells in adults and is activated via two pathways i.e. the protein kinase C (PKC) and Jun-N terminal (JNK). The PKC pathway is activated when calcium is released and when Wnt ligands adhere to Frizzled receptors (Yamanaka et al., 2002). On the other hand, the JNK pathway is triggered when the dishevelled and Wnt proteins (5a, Wnt 5 and Wnt 11) bind to Frizzled molecules (Oishi

et al., 2003). The JNK molecule thus moves to the nucleus and modulates the transcription of several proteins (Doucas et al., 2005).

This signalling pathway is also implicated in cancer development, particularly gastric malignancy, where it promotes the progression of cancer (Gencer et al., 2010). There are approximately 30% cases of gastric cancers which result from the dysfunction of the non-canonical signalling pathway (Sugimura and Li, 2010).

8. The *CDH1* gene structure

The tumour suppressor gene, *CDH1* is located on chromosome 16q22.1 (Figure 4) and has 2.6 kilobases (kb) of coding sequences and contains 16 exons. This gene encodes a 120kDa epithelial cadherin (E-cadherin) protein (Pinho et al., 2011). The gene is commonly mutated, resulting in its loss of function and therefore give rise to malignancy.

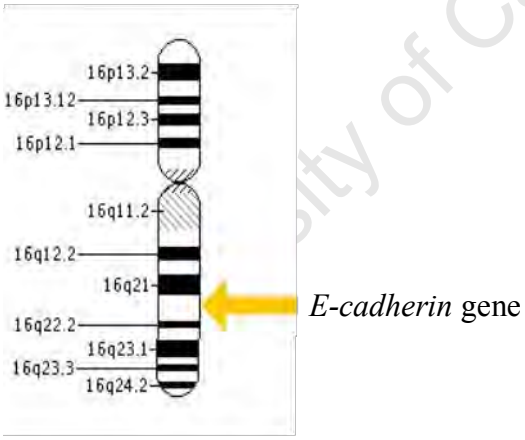


Figure 4: Diagrammatic representation of the *E-cadherin* gene location on the long arm of chromosome 16 at region 2, band 2 and sub-band 1 (16q22.1) (http://ghr.nlm.nih.gov/gene=*CDH1*).

8.1. The *CDH1* gene and Familial Gastic Cancer

Germline mutations are commonly identified in the *CDH1* gene amongst individuals with familial gastric cancer (Suriano et al., 2003; Brooks-Wilson et al., 2004; Bacani et al., 2006; More et al., 2007; Lynch et al., 2011). These alterations span the entire coding sequence of this gene and are prevalent in approximately 29% to 56% of hereditary diffuse gastric cancer (HDGC) families (Lynch et al., 2005; Bacani et al., 2006; Corso et al., 2012).

Bacani et al. (2006) identified 20 mutations in different regions of the gene, of which 12 were novel. One of these mutations included a new nucleotide deletion (nt41delT) in exon 1. This variant resulted in reduced membrane expression of the E-cadherin protein leading to loss of function. Other mutants were also identified and localized in the 5'UTR region (-117G>A; -71C>G). Alterations observed in the intronic segments were; 48(+15)C>G; 387(+26)C>T, 2295(+53)G>A, 2439(+31)G>A and 48(+15)C>G. Interestingly, the majority of mutations observed were located in exons 3, 6, 7, 8 and 15 were silent (Bacani et al., 2006).

In another study, two missense germline mutations were identified by Suriano et al., (2003) with one having a negative impact on the E-cadherin protein structure and function. The specific mutations occurred in codons 617 (1849 G>A) and 634 (1901 C>T). These produced the following amino acid substitutions; Ala to Thr and Ala to Val, respectively. The mutation in codon 617 was identified in the 5th extracellular domain, compromising the calcium binding site which impaired cell attachment, thereby promoting cell invasion (Suriano et al., 2003).

Brooks-Wilson et al., (2004), identified a splice site mutation G>A at positions (+1) IVS5 and (+5) IVS 11, which was thought to be the cause of an in-frame deletion of exon 5 at position (+1) IVS5. The entire exon 8 was deleted while only a part of exon 11 was deleted at position (+5) IVS11. Other alterations included, deletions in exons 3 (382delC), 8 (1134del8), 9 (1212delC), 10 (1476delAG), 13 (2061delTG) and 15 (2319delC). Other alterations included the following insertions 1134ins5, 1064insT and

1779insC in both exons 8 and 12. Furthermore, amino acid substitutions namely; A298T, W409R and R732Q were also identified. In addition, missense mutations identified included G892A, T1226T and G2195A in exons 7, 9 and 14, respectively. Significantly, all these genetic alterations disrupted the cell-cell adhesion and tumour suppressive role of the protein (Brooks-Wilson et al., 2004).

A similar study conducted by More et al. (2007), detected 7 novel mutations and 2 that were previously identified in 9 of 36 HDGC families. The mutations were as follows; 3(53C>G); 6 (715G>A) and 12 (1901C>T). These mutations resulted in the following amino acid substitutions; (p.Thr118>Arg; Gly239Arg and Ala634Val). Interestingly, the p.Thr118>Arg mutant generated a protein with an impaired adhesive function. Additionally, other mutations were also observed in intron 1 (49-2A>C); exon 8 (1137G>A) and exon 15 (2440-6C>G). Furthermore, base pair deletions (1bp c.1107delC and 2bp c.1391_1392delTC) which resulted in the production of stop codons Asn369LysfsX392; Leu472HisfxX481 and Gln699X were also detected. These deletions resulted in truncation of the extracellular domains and therefore increased the risk of developing the tumour (More et al., 2007).

Kaurah et al., (2007), discovered 13 mutations in 15 of 38 HDGC families. Six of these mutations were novel and were as follows; 1682-1683insA; 1913G>A; 2164+5G>A; 2245C>T; 2343A>T and 2398delC. The mutations that were previously identified included; 715G>A; 283C>T; 1137G>A; 1397-1398delTC; 1901C>T; 2064-2065delTG and 2195G>A. Two novel missense mutations namely; 2245C>T and 2343A>T were found to compromise cell-cell adhesion (Kaurah et al., 2007). The majority of these mutations resulted in a truncated protein which impaired the cell adhesion function (Bacani et al., 2006).

8.2. The *CDH1* gene and Sporadic Gastric Cancer (SGC)

Somatic alterations on the other hand have been implicated in the development of sporadic gastric cancer. The hot spot regions of these mutations are exons 7 to 10 of the *E-cadherin* gene, with exons 8 and 9 being the commonly mutated regions (Ascano et

al, 2001; Oliveira et al., 2009). However, they have also been detected in exons 2, 3, 5 and 16. Approximately 31% of stomach cancers carry *E-cadherin* (*CDH1*) somatic gene mutations (Corso et al., 2013). The majority of these genetic changes have been identified as a second hit mechanism in familial gastric cancer, with the incidence of the second hit being 75% (Oliveira et al., 2009).

Unlike germline mutations, only a few studies have investigated somatic mutations in gastric cancer (Becker et al., 1994; Becker et al., 1998; Ascano et al., 2001; Machado et al., 2001; Li et al., 2012). Becker and colleagues discovered mutations that resulted in the loss of cell-cell adhesion, in exons 8; 9 and 10. Furthermore, amino acid substitutions; aspartic to valine and valine to arpatic were identified in codons 370 and 473 in exons 8 and 10, respectively. These amino acid substitutions resulted in the elimination of calcium binding sites hence a loss of function (Becker et al., 1994).

Mutations were detected in exon 3 (129 bp deletion) which resulted in the substitution of histamine to tyrosine. In exon 4, 63 nucleotides and 20 amino acids were deleted. Further, amino acid substitutions; glutamic to glutamine and arginine to glutamine were identified in exons 10 and 12, respectively. These genetic changes were shown to have an association with the loss of protein function (Becker et al., 1998). In 2001, *CDH1* genetic mutations were identified in 7 of 25 diffuse type cases. In addition, novel mutations ag/AGT to cg/AGT; c insertion at 1208 and 5 bp insertion at 1216 g1849a) in intron 7, exon 9 and 12 were identified in 4 diffuse type cases. These mutations resulted in an aberrant protein expression (Ascano et al., 2001).

A study by Machado and colleagues identified 56.3% of somatic changes in 16 diffuse gastric cancers. Five of these mutations produced single amino acid substitutions (1742T>C; 1108G>C; 1436A>G; 1204G>A; 1000G>C). Furthermore, deletions, 1027delC and 521del10 + 1del16 were also identified. Loss of protein expression was seen in all the nine patients (Machado et al., 2001).

8.3. *CDH1* Large genomic alterations and gastric cancer

Gastric adenocarcinoma can also arise as a result of large genomic changes (Oliveira et al., 2009; Pinheiro et al., 2010; Yamada et al., 2011). These genomic changes are commonly identified in populations with low prevalence of stomach cancer amongst *CDH1* germline negative patients. Oliveira et al., (2009) detected 6.5% large genomic deletions out of 93 patients. These deletions included Del exon 1-2(193593 bp); Del exon 1-2(5671 bp); Del 5'-UTR-exon 1(150 bp); Del exon 14-16(8078 bp) and Del exon 16(828 bp) (Oliveira et al., 2009). Protein expression was not determined, however, it is believed that these genomic deletions are the cause of gastric cancer in patients negative for *CDH1* germline mutations.

Pinheiro et al., (2010) also discovered two large genomic deletions, 193.593 bp and 150 bp in positions IVS2+57.595 and TSS, respectively (Pinheiro et al., 2010). Other deletions were identified by Yamada et al., (2011) and these involved a heterozygous c.1212del and a c.1212C deletion in exon 9. The c.1212C mutation produced a premature protein (p.Asn405Ile-fxX12) in codon 415. In addition, a c.164-? _387+?del deleted exon 3 and led to a truncated protein p.Val55GlyfsX38. All alterations damaged cell-cell attachment by disrupting the extracellular and cytoplasmic domains (Yamada et al., 2011).

8.4. Microsatellite instability in gastric cancer

Microsatellites are short tandem repeats (STRs) in a gene. These repeat sequences accumulate mutations due to polymerase slippage (Ottini et al., 2004). The genetic changes are referred to as microsatellite instability (MSI) and play a pivotal role in gastric cancer development (Shokal and Sharma, 2012). Five microsatellite markers (*BAT25*, *BAT26*, *D5S346*, *D2S123* and *D17S250*) are commonly utilised to detect microsatellite status in gastric cancer. The microsatellites are considered microsatellite instability-high (MSI-H) when there is instability in two or more markers. On the other hand, microsatellites are microsatellite instability-low (MSI-L) if only one marker is unstable. They are microsatellite stable if there is no instability at any marker (An et al.,

2011). MSI is seen in 15 to 30% cases of stomach cancer and is commonly identified in intestinal tumours in the antrum, amongst the elderly people (Pedrazzani et al., 2009).

Studies have also reported MSI as a contributing factor in gastric carcinogenesis (Oki et al., 2009; Jeong et al., 2010). A study investigating microsatellite instability in 240 gastric cancer patients identified MSI-H, MSI-L and MSS in 22, 25 and 193 individuals, respectively. MSI-L and MSS were associated with older patients, tumour site (antrum) and lymph vessel involvement. No association was observed between MSI-H and clinicopathologic characteristic (Oki et al., 2009). Jeong showed that MSI was 10.2% in a study which included 49 patients. The MSI positive tumours were all intestinal types (Jeong et al., 2010). Research conducted in 2009 in 250 patients showed that MSI occurred in 25.2% of the study cohort. Of these MSI positive cases, there were 41.8% intestinal types (Pedrazzani et al., 2012).

Other studies have also investigated the impact of microsatellites in gastric cancer (Perez et al., 2004; Liu et al., 2005; Beghelli et al., 2006). One study involved a total of 510 patients and detected MSI in 83 of the cases. The instability was commonly observed in older female patients (33%) in the intestinal type. It was seen less frequently in advanced cancers and those with a lymph node involvement (Beghelli et al., 2006). Liu and colleagues investigated 36 cases in which MSI was identified in 21 (58.3%) of these cases. Further, MSI-H and MSI-L were seen in 7 and 14 of the tumours, respectively. There was a correlation between the MSI data and the clinicopathologic features. MSI-H/MSI-L was correlated with stage II and the location of the tumour in the distal stomach (Liu et al., 2005).

Perez et al., (2004), analysed 24 sporadic cases in their study and identified 5 MSI positive patients of which 3 were of the intestinal type and 2 of the diffuse type. These early stage tumours were located in the proximal stomach and had no lymph node involvement (Perez et al., 2004).

8.5. *CDH1* Polymorphisms in gastric cancer

Single Nucleotide Polymorphims (SNPs) are single base pair genetic changes in a DNA sequence (Medina-Franco et al., 2007). They can be used as genetic markers to ascertain the association between certain genes and the risk of cancer growth. Different studies have investigated these genetic markers in stomach carcinogenesis (Machado et al., 2001; Zhan et al., 2012). Various polymorphisms of the *E-cadherin* gene increase the risk of this malignancy. In one study, the effect of four polymorphisms such as rs13689, rs1801552, rs16260 and rs17690554 was investigated on the *CDH1* gene. There was an association between two SNPs (rs16260 and rs17690554) with the risk of developing diffuse gastric cancer in 30% and 22% of 57 patients, respectively (Zhan et al., 2012). Research by Al-Moundhri and colleagues investigating 4 polymorphisms in a study that involved 192 patients showed that there was an association between the polymorphism and the risk of developing gastric cancer in 60% individuals.

One of these SNPs (-160 C>A) had an association with an enhanced risk of developing the cancer in 60% of the patients (Al-Moundhri et al., 2010). Another study investigated 12 SNPs (rs16260, rs1559366, rs10431923, rs7188750, rs1801552, rs33965115, rs1801026, rs13689, rs1477407, rs12935840, rs8062856, rs2232228). Amongst these polymorphisms, rs1801026 and rs13689 had an association with the development of the cancer. However, this association with rs13689 was found not to be statistically significant (Jacobs et al., 2011).

A study by Zhang et al., (2008), involved 96 SGC patients and determined the effect of -347delA, -160C>A, -73A>C, +178T>C and +234 13N ins>del SNPs in the *CDH1* gene promoter area. Amongst these variants; -347delA, +178T>C and +234 13N ins>del cooperatively enhanced the risk of the tumour in 65.8% SGC individuals (Zhang et al., 2008). Other variants namely; rs10673765, rs9932686, rs1125557, rs9282650 and rs9931853 were also investigated amongst 134 DGC patients by Nasri et al., (2008). The rs1125557 (also referred to as 163+37235G>A) was seen in 56.7% of the patients and thus increased the risk of the disease (Nasri et al., 2008). Medina-Franco and colleagues identified an association between the -160 C/A SNP and 16 of 39 DGC

patients (Medina-Franco et al., 2007). Polymorphisms are therefore associated with the risk of developing stomach cancer.

9. Aims and Objectives

The primary aim of this study was to determine the prevalence of E-cadherin (*CDH1)* genetic mutations in gastric cancer patients seen in the Western Cape and to correlate these findings with the demographic/clinicopathologic data. The mutational data was obtained using PCR and sequencing molecular techniques. The objectives of this study were to:

- Sequence the *CDH1* gene in a cohort of gastric cancers.
- Analyse *CDH1* mutations in the gastric cancer patients.
- Correlate the molecular and demographic/clinicopathologic data.

CHAPTER 2

Materials and Methods

2.1. Ethical Approval

Ethics approval for the study was obtained from the Human Research Ethics Committee, Faculty of Health Sciences (HREC REF: 504/2009), at University of Cape Town.

2.2. Case Selection

One hundred and two gastric tumour and their corresponding normal tissue samples were obtained from the archives of the Division of Anatomical Pathology, University of Cape Town/NHLS. A statistician was consulted to assist with determining the appropriate sample size to include in the study. The number of cases selected was based on the worldwide prevalence or proportion of the *E-cadherin* gene mutations from literature reviewed. After staining the cases with haematoxylin and eosin (H & E), the slides were examined by a pathologist, Dr Michael Locketz, (Pathologist and Collaborator) from the Division of Anatomical Pathology, NHLS/UCT. The normal and tumour areas were demarcated on the slide and only cases presenting with more than 80% tumour were included in this study.

2.3. Pathologic Case Reports

The case reports were reviewed from the laboratory information system and staged according to the TNM system, 2002 AJCC Cancer staging (Manual, 6th Ed, New York, 2002, Springer), see Appendix D. In addition, information related to age, gender and *H. pylori* status was also retrieved from the reports. Furthermore, the patients' previous gastric biopsy case reports were reviewed to ascertain whether *H. pylori* were present or absent. All the biopsy specimens were examined for the presence of *H. pylori* using a modified Giemsa stain.

2.3. Histology

The corresponding archived tissue blocks for the 102 cases had been previously fixed in 10% buffered formaldehyde and processed overnight in an automatic tissue processor (Tissue Tek VIP, Sakura Finetek, Torrance, CA). They were then embedded into wax blocks using the Tissue Tek embedding machine (Sakura Finetek, Torrance, CA, USA). These wax blocks were cut at 3µm and 6µm with a rotary microtome (Accu-Cut SRM 200 Rotary Microtome, Serial No: 1429-0925). Sterile conditions were adhered to at all times. The Accu-Edge Low-Profile Blades were wiped with alcohol and gloves were worn during sectioning. The tissue sections were then picked up on glass slides and heat fixed on a hotplate at 60°C for 10 minutes. The 3µm sections were stained with haematoxylin and eosin whilst the three 6µm sections were used for the molecular studies of this project.

2.4. Haematoxylin and Eosin (H & E) Staining

The 3µm gastric tissue sections were dewaxed in xylene, passed through a series of absolute alcohol and finally washed in running tap water. Mayers haematoxylin was used to stain the slides for 6 minutes. The slides were rinsed in tap water and the nuclei blued with ammoniated water. The slides were then washed, and a 1% phloxine/eosin was used to stain the cytoplasm and surrounding tissue for 5 minutes. The cytoplasm and surrounding tissue stained various shades of pink. Finally, the slides were washed thoroughly in water, dehydrated in alcohol and cleared in xylene. Entellan was used to mount the slides onto cover slips.

2.5. DNA Extraction

The three 6µm sections were dewaxed in xylene, cleared in alcohol, thoroughly washed in running tap water, and left to air dry. After the unstained slides were dried, they were superimposed onto the H & E slides previously marked by a pathologist. The normal and tumour regions were demarcated on the unstained slides. The three sections of normal and tumour tissues were scrap ed and transferred into autoclaved 1.5ml Eppendorf tubes which were labelled accordingly. DNA isolation was performed

utilizing the QIAamp®DNA FFPE Tissue extraction kit (Whitehead Scientific) according to the manufacturer's protocol. This consisted of heating the samples in the Eppendorf tubes with 20µl PK and 180µl ATL lysis buffer at 56ºC. A heat block was used to deactivate the PK at 90 ºC for 1 hour. The tubes were then centrifuged and 200µl of AL (lysis) buffer together with 200µl of absolute ethanol was added to this and mixed by vortexing. After centrifuging, the entire lysate was transferred to a QIAampMinElute column (in a 2ml collection tube). The collection tube with the flow through was replaced with a new one and 500µl of AW1 (wash) buffer was added to the column. After centrifuging the collection tube with the flow through was discarded and replaced. 500µl of AW2 (wash) buffer was added to the column and centrifuged. The column with the flow through was once again discarded and replaced with a new one. The membrane of the column was allowed to dry completely by centrifuging. The collection tube was discarded and replaced with a 1.5ml microcentrifuge tube. 50µl of ATE (eluting) buffer was applied to the membrane of the QIAampMinElute column which was once again centrifuged. The flow through was collected and stored at -20 C for further tests. The end product being 50µl of purified DNA.

2.6. DNA quantification

The DNA samples were quantified using a Nanodrop instrument (Thermo Scientific, Wellington, DE 19810 USA) prior to insulin PCR. This instrument allows the quantification of samples such as proteins and nucleic acids (DNA and RNA). A 2µl volume of the DNA samples was used for quantification.

2.7. Insulin PCR

Insulin PCR was performed to verify the quality of DNA extracted from FFPE tissues. The insulin gene was amplified using the exon 2 region with exon-specific primers: 5'-AGG GGC AGC AAT GGG CGG TTG-3' (reverse) and 5'-ACC CAG ATC ACT GTC CTT CTG CC-3' (forward). The PCR product size for these primers was 236 base pairs (bp). A 25µl PCR reaction was carried out containing 5µl of template DNA, 5pmol of each primer, 10mM dNTPs, 10x PCR buffer, and 5u/µlTaq DNA polymerase. The PCR conditions were as follows:

a). Initial Denaturation 1 cycle: 94 Cfor 5 minutes

b). 30 cycles: 94 Cfor 1 minute

: 64 Cfor 1 minute

: 72 Cfor 2 minute

c). Final Extension cycle: 72 Cfor 10 minutes

2.8. Gel Electrophoresis

Following insulin PCR, the amplicons were separated on 1.5% Agarose/Ethidium Bromide gel in 1x TBE buffer and run at 100 V for 45 minutes and then visualized under UV-transillumination using the SynGene Image Acquisition Software (SynGene Version 7.05.02). A 100 bp molecular weight marker (Hyperladder IV) purchased from BIOLINE was used to determine the DNA product size.

2.9. Primer Design

All studies thus far investigating sequence analysis of the *E-cadherin* (*CDH1*) gene in gastric cancer, showed that the main hotspots are located on exons 8 and 9 (Berx, et al., 1995; Oliveira et al., 2009). This study also focused on investigating these exons since to date there is no published data on these investigations in South Africa. The primer sets were therefore designed for exons 8 and 9 using bioinformatics software namely, Primer3 and Oligoanalyzer. The Oligoanalyzer software was utilized to analyse primer properties such as hairpins, self and cross dimers. These are primer secondary structures produced as a result of self-complementarity within the primers. Since these secondary structures may interfere with amplification, they should be avoided during primer design. The primers (Table 1) were synthesised in the Department of Molecular Cell Biology at University of Cape Town (UCT). The PCR product size for exon 8 and 9 primers was 188 bp and 230 bp (Table 1), respectively.

Table 1: PCR Primer sequences for exons 8 and 9.

Region	Forward	Reverse
Exon 8	5'-ACTTGGTTGTGTCGATCTCT-3'	5'-CTTTGGAAACCCTCTAAGGA-3'
Exon 9	5'-TGACACATCTCTTTGCTCTG-3'	5'-CATCTTGCCAGGTACCATAC-3'

2.9.1. Exon 8 PCR

The PCR was performed in a final reaction volume of 50µl consisting of 3.5µl of template DNA (120ng); 10µM of each primer; 5u/µl of My Taq DNA polymerase and 5x My Taq PCR reaction buffer containing 5mM dNTPs and 15mM $MgCl_2$. Positive (50ng/µl high quality genomic DNA from a human blood sample) and negative controls (sterile water) were included in the reaction. These reagents were purchased from BIOLINE. The PCR was performed under the following conditions:

a). Initial Denaturation 1 cycle: 94 Cfor 5 minutes
b). 30 cycles: 94 Cfor 1 minute
: 48 Cfor 1 minute
: 72 Cfor 1 minute
c). Final Elongation cycle: 72 Cfor 7 minutes

2.9.2. Exon 9 PCR

The PCR was performed in a final reaction volume of 50µl consisting of 3.5µl of template DNA (120ng); 10µM of each primer; 5u/µl of My Taq DNA polymerase and 5x My Taq PCR reaction buffer containing 5mM dNTPs and 15mM $MgCl_2$. Positive (50ng/µl genomic DNA sample) and negative controls (sterile water) were included in the reaction. These reagents were purchased from BIOLINE. The PCR was performed under the following conditions:

a). Initial Denaturation 1 cycle: 94 Cfor 5 minutes

b). 30 cycles: 94 Cfor 1 minute

: 50 Cfor 1 minute

: 72 Cfor 1 minute).

Final Elongation cycle: 72 Cfor 7 minutes

2.10. Gel Electrophoresis

The PCR products were run on a 1.5% Agarose/Ethedium Bromide gel in 1x TBE buffer at 100 V for 45 minutes to determine the size and also verify the presence of DNA. They were visualized under UV-transillumination using the SynGene Image Acquisition Software (SynGene Version 7.05.02). A 100 bp molecular weight marker from BIOLINE was used to determine the size of PCR products.

2.13. Nested PCR

The nested PCR was carried out to improve DNA amplification since the standard sequencing PCR was unsuccessful in the majority of the samples. The unsuccessful PCR may have been due to the DNA quality affected by formalin fixing and paraffin embedding, when the tissue blocks were generated. Another factor that contributes to DNA damage is the storage time; the DNA from FFPE tissues becomes fragmented as the tissues get older. The latter, therefore makes it difficult to amplify DNA extracted from FFPE tissues, hence the DNA samples were subjected to further optimization using the antigen retrieval process. Since the nested PCR involves two sets of primers where the product from the first reaction is used as a template for the second reaction; new sets of primers which generate a bigger product size (397 bp and 484 bp for exon 8 and 9, respectively) were used.

2.13.1. Exon 8 nested PCR

Two sets of primers (Table 2 and Table 3) were used to amplify each exon. To amplify exon 8, the primer set I was used. The PCR was performed in a final reaction volume of 25µl consisting of 3.5µl of template DNA (120ng); 10µM of each primer; 5u/µl of My Taq DNA polymerase and 5x My Taq PCR reaction buffer containing 5mM dNTPs and 15mM $MgCl_2$. Positive (50ng/µl genomic DNA sample) and negative controls (sterile water) were also included in the reaction. The PCR products from this reaction were used as DNA templates for the second reaction using the primer set II. The PCR product size for exon 8 was 397 bp. The PCR was performed under the following conditions:

a). Initial Denaturation 1 cycle: 94 Cfor 5 minutes
b). 30 cycles: 94 Cfor 1 minute
: 50 Cfor 1 minute
: 72 Cfor 1 minute
c). Final Elongation cycle: 72 Cfor 7 minutes

2.13.2. Exon 9 nested PCR

Exon 9 was amplified in a final volume of 25µl consisting of 3.5µl of DNA (120ng); 10µM of each primer (primer set I); 10X PCR reaction buffer without $MgCl_2$; 5u/µl Taq DNA polymerase; 10mM of dNTPs and 5mM of $MgCl_2$. Positive (50ng/µl genomic DNA sample) and negative controls (distilled water) were also included in the reaction. The PCR products from this reaction were used as DNA templates for the second reaction using the primer set II. The PCR product size for exon 9 was 484 bp. The PCR was performed under the following conditions:

a). Initial Denaturation 1 cycle: 94 Cfor 5 minutes
b). 30 cycles: 94 Cfor 1 minute
: 53 Cfor 1 minute
: 72 Cfor 1 minute
c). Final Elongation cycle: 72 Cfor 7 minutes

Table 2: PCR Primer sequences for exon 8.

Primer set	Forward	Reverse
I	5'-TGTCAGTTAATACAGTGATGGT-3'	5'-ACTAAAAGACAACTGCCATCT-3'
II	5'-AGCTTTGCATCTAAGCTTGT-3'	5'-TCTGGAGAGAGTTTTCAAAGA-3'

Table 3: PCR Primer sequences for exon 9

Primer set	Forward	Reverse
I	5'-TTGTCACATCTTCTCCTTGA-3'	5'-TGCCTGGCTAATTTTTGTAT-3'
II	5'-TGTCAGTTAATACAGTGATGGT-3'	5'-ACTAAAAGACAACTGCCATCT-3'

2.14. DNA Sequencing of PCR products

The PCR products were diluted prior to sequencing. A volume of 60µl together with 10µM of each primer (forward and reverse) was sent to Macrogene, a company based in South Korea. This company offers sequence analysis of plasmid and PCR products using a high throughput Applied Biosystems 3730XL sequencer. The PCR primers with 188 bp and 230 bp product size were used for the sequencing reaction of exon 8 and 9, respectively. The DNA products were amplified using a BigDye Terminator V1.1/3.1 kit (Applied Biosystems, Forster City, CA 94404, U.S.A.). Bioinformatics tools; ClustalW and BioEdit were used to analyse *CDH1* genetic mutations.

2.15. Statistical Analysis

The demographic/clinicopathologic data was analysed using Stata 11.

CHAPTER 3

3. Results

3.1. Demographic and clinicopathologic characteristics of study participants

A total of 102 gastric cancer cases were enrolled in the study (Appendix C). The mean age of these participants was 57 years (range: 18-84 years). Ninety six patients had their gastric cancers classified as intestinal (n = 38), diffuse (n = 53) and mixed (n = 5), according to the Lauren classification system. Two patients were classified as mucinous adenocarcinoma and 4 were not classified. Our study consisted of 48 male patients and 51 female patients and gender was not stated in 3 cases. The tumours were grouped into different stages 0 (3 patients); IA (6 patients); IB (13 patients); II (24 patients); IIA (3 patients); III (2 patients); IIIA (19 patients); IIIB (11 patients); IIIC (1 patient) and stage IV (5 patients). Twelve of the patients were not classified for stage. Further, they were categorized into grade 1 (2 patients); grade 2 (29 patients); grade 3 (32 patients) and grade 4 (30 patients) and 7 individuals were not classified for grade. There were 54 patients who had lymph node metastasis; 25 patients were negative and 23 were not reported for lymph node metastasis. For *H. Pylori* infection, only 24 patients were *H. Pylori* positive; 76 were negative and the infection was not reported in 2 individuals (Table 4).

3.2. Descriptive statistics of demographic/clinicopathologic data

The patients were categorised into different age groups (Figure 5), <50 years, 50-69 years and >70 years.

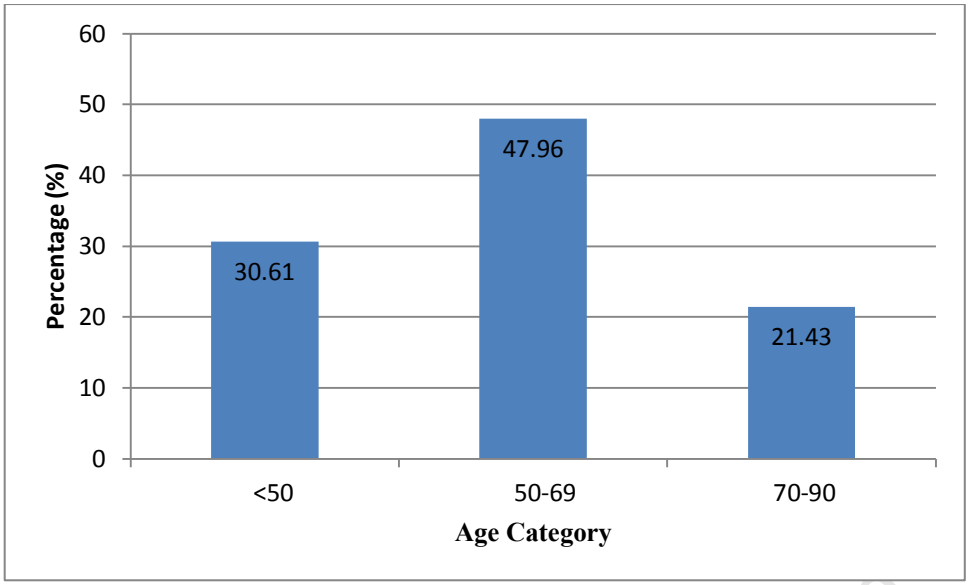


Figure 5: The distribution of different age groups in gastric cancer patients.

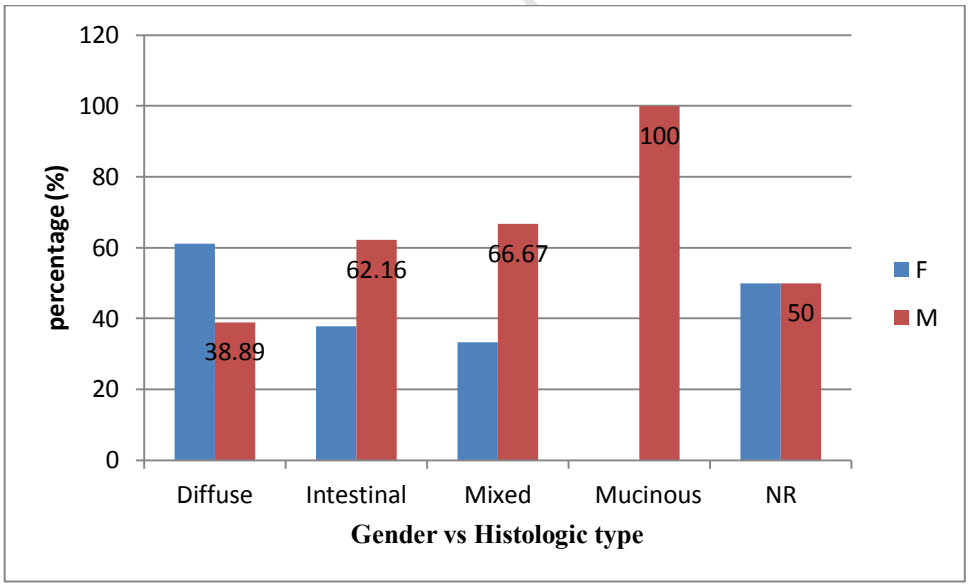


Figure 6: The relation between gender and the histologic types of gastric cancer is indicated. The diffuse type is commonly seen amongst female patients (61.11%) while the intestinal type is more frequent amongst the male patients (62.15%).

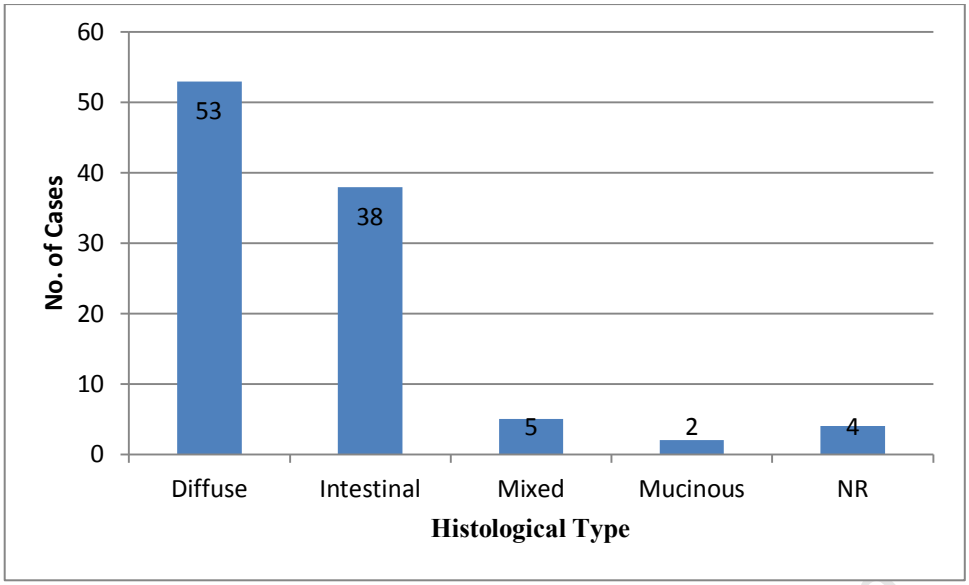


Figure 7: The distribution of histological types in gastric cancer patients.

Table 4: The demographic/clinicopathologic characteristics of the (study cohort) 102 cases.

Clinicopathologic feature	**No. of cases with valid entry/102**
Age (yrs)	98
Sex Ratio (Male:Female)	48:51
Gender not stated	2
Stage	
0	3
IA	6
IB	13
II	24
IIA	3
III	2
IIIA	19
IIIB	11
IIIC	1
IV	5
No stage assigned	12
Grade	
1	2
2	29
3	32
4	30
Not graded	7
Lymph node metastases	
No	25
Yes	54
Lymph node metastases not assigned	23
H.pylori	
No	76
Yes	24
H.pylori status not assigned	2

3.2. Microscopic examination of intestinal and diffuse gastric adenocarcinoma

All the gastric cancer cases (102) selected for this study were reviewed and examined by a pathologist (Dr M Locketz) and categorized as the diffuse and intestinal histologic types.

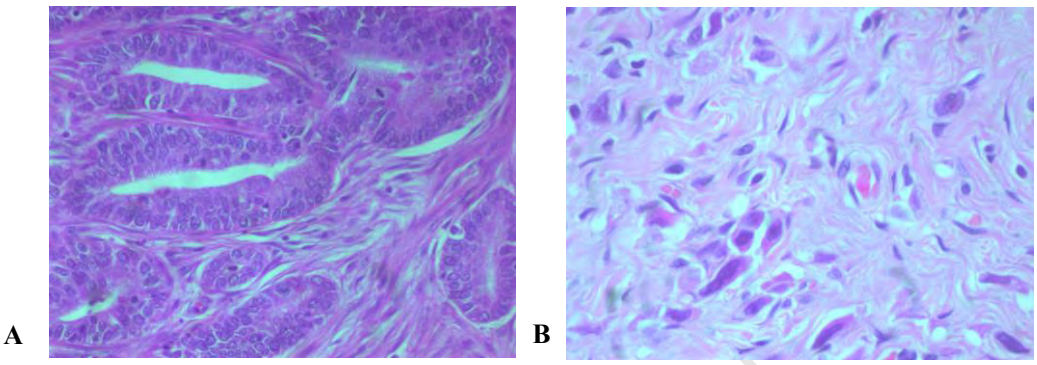


Figure 8: A diagrammatic representation of the intestinal (**A**) and diffuse (**B**) types of gastric cancer. The intestinal type is characterized by glands and the diffuse type by single line dis-cohesive cells.

3.3. Verification of DNA quality

DNA was isolated from all the 102 samples and the exon 2 region of the insulin gene was amplified using PCR to verify the quality of the DNA. However, the insulin PCR was only successful in 44 samples. This was due to the DNA not being of optimal quality due to formalin fixing and paraffin embedding. A statistician was consulted and the Epicalc2000 software was used to determine the sample size. A sample size of 20 was deemed adequate for this study. A representative 1.5% agarose gel of DNA samples is shown (Figure 9). A successful DNA isolation of the samples was indicated by the presence of a 236 bp PCR product.

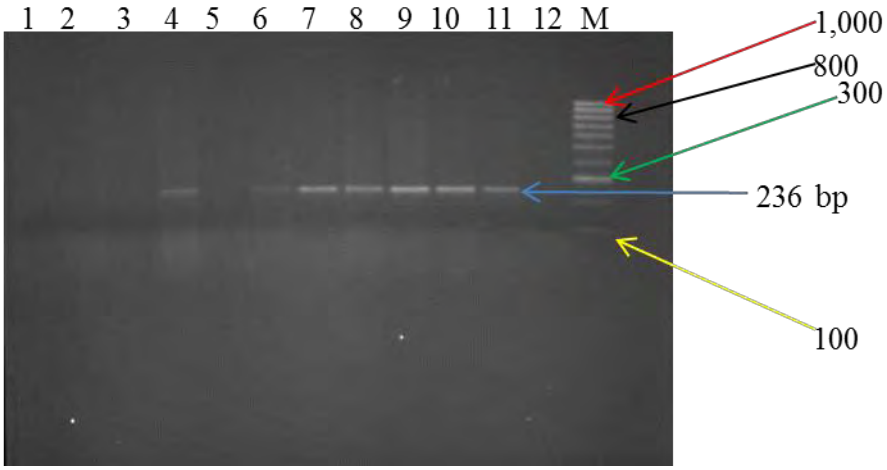


Figure 9: Representation of a 1.5% agarose gel showing amplification of exon 2 of the insulin gene which produces a PCR (236 bp) product. Lanes 1 to 10 represent the DNA samples; lane 11 positive (+) and lane 12 negative control (-); M, a 100 bp ladder.

3.4. E-cadherin (*CDH1*) gene amplification

Exon 8 and 9 of the *CDH1* gene were amplified. Of the 44 samples, only 26 (Table 5) could be amplified on exon 8 whilst only 2 could be amplified on exon 9. Representative 1.5% agarose gels of DNA samples are shown in Figure 10 and Figure 11. A successful DNA amplification was indicated by the presence of PCR products of size 397 bp and 484 bp for exon 8 and 9, respectively.

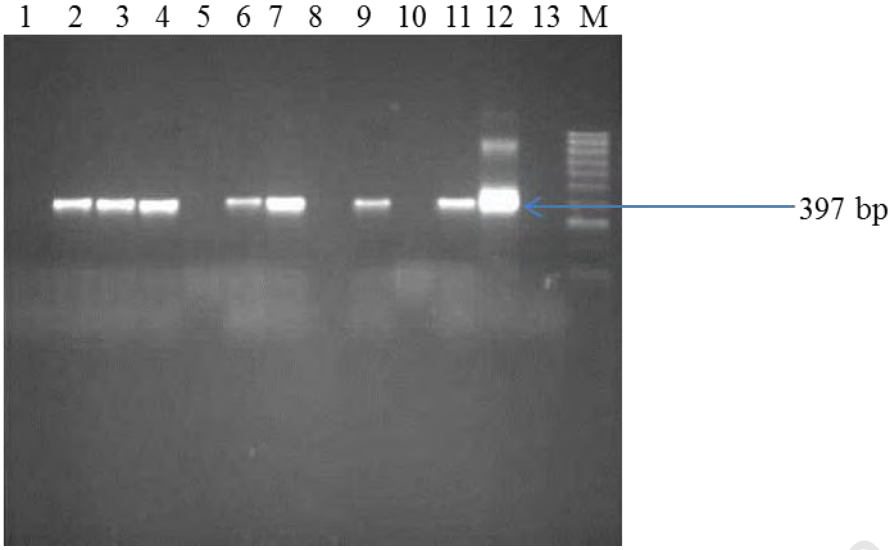


Figure 10: Representation of a 1.5% agarose gel showing amplification of exon 8 of the *CDH1* gene which produces a PCR (397 bp) product. Lanes 1 to 11 represent the DNA samples; lane 12 positive (+) and lane 13 negative controls (-); M, a 100 bp ladder.

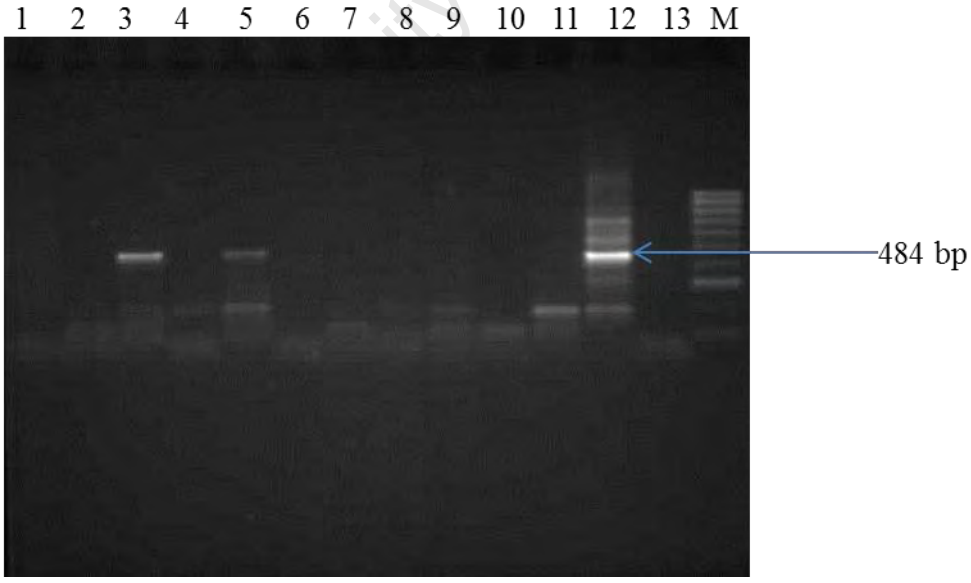


Figure 11: Representation of a 1.5% agarose gel showing amplification of exon 9 of the *CDH1* gene which produces a PCR (397 bp) product. Lanes 1 to 11 represent the DNA samples; lane 12 positive (+) and lane 13 negative controls (-); M, a 100 bp ladder.

Table 5: The demographic and clinicopathologic features of the 26 cases.

Patient ID	**Gender**	**Age**	***H.pylori***	**Node**	**Grade**	**Stage**	**Type**
T2	F	77	0	No	2	0	Intestinal
T3	M	78	0	NR	4	IB	Intestinal
T7	M	54	1	NR	1	IA	Intestinal
T17	F	74	0	Yes	4	IIIA	Diffuse
T30	M	48	0	Yes	NR	IIIB	Diffuse
T68	F	36	0	NR	4	III	Diffuse
T72	F	37	1	No	3	0	Diffuse
T79	F	64	1	Yes	4	IV	Diffuse
T82	F	37	1	No	3	0	Diffuse
T91	F	61	0	Yes	3	NR	Diffuse
T92	M	60	0	No	3	IIA	Diffuse
T95	F	42	1	No	3	IB	Diffuse
T16	M	84	1	No	2	NR	Intestinal
T81	M	64	0	Yes	3	IIIA	Diffuse
T11	F	38	0	Yes	NR	IIIA	Diffuse
T12	F	77	0	NR	3	IV	Intestinal
T13	M	60	0	No	2	NR	Intestinal
T18	M	60	0	No	2	IB	Intestinal
T31	F	70	0	Yes	2	IIB	Intestinal
T32	F	43	0	No	4	IB	Diffuse
T35	M	52	0	Yes	2	IIIB	Intestinal
T38	F	NR	0	Yes	2	IIIA	Intestinal
T39	M	60	0	NR	3	1B	Intestinal
T42	M	46	0	Yes	4	IIIA	Diffuse
T67	F	52	0	No	4	II	Diffuse
T70	F	78	1	NR	4	II	Diffuse

Key:
NR Not Recorded
0 *H. pylori* negative
1 *H. pylori* positive

3.5. Sequence analysis

After the amplification of exon 8 and 9 of the *CDH1* gene, the PCR products were subjected to a sequencing reaction. The sequencing data of each patient was then analysed.

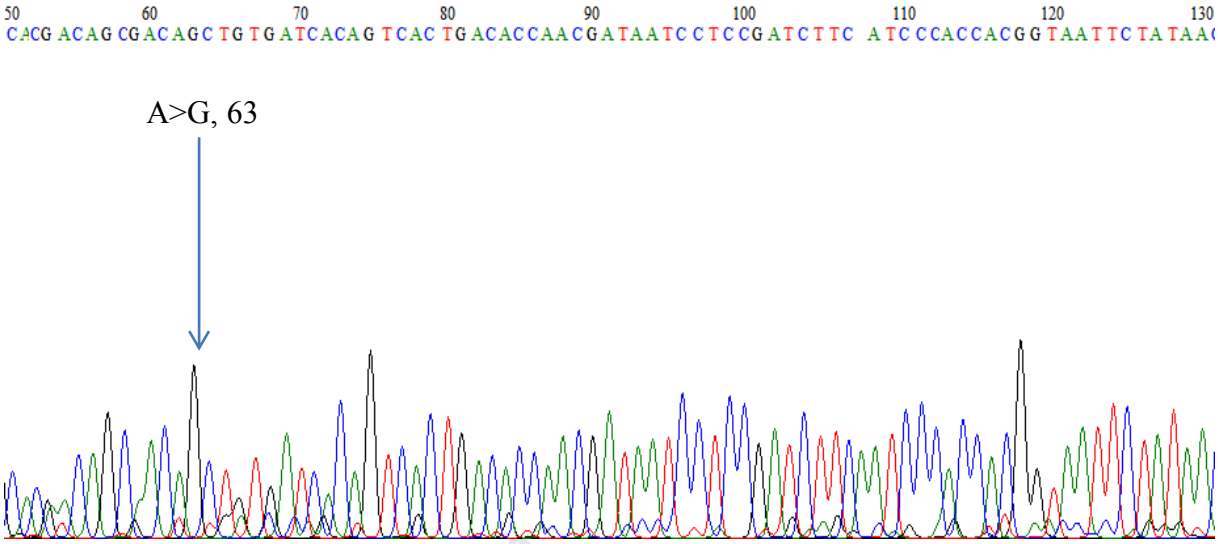


Figure 12: A representative chromatogram of sample T68 showing sequencing data as obtained from Macrogene.

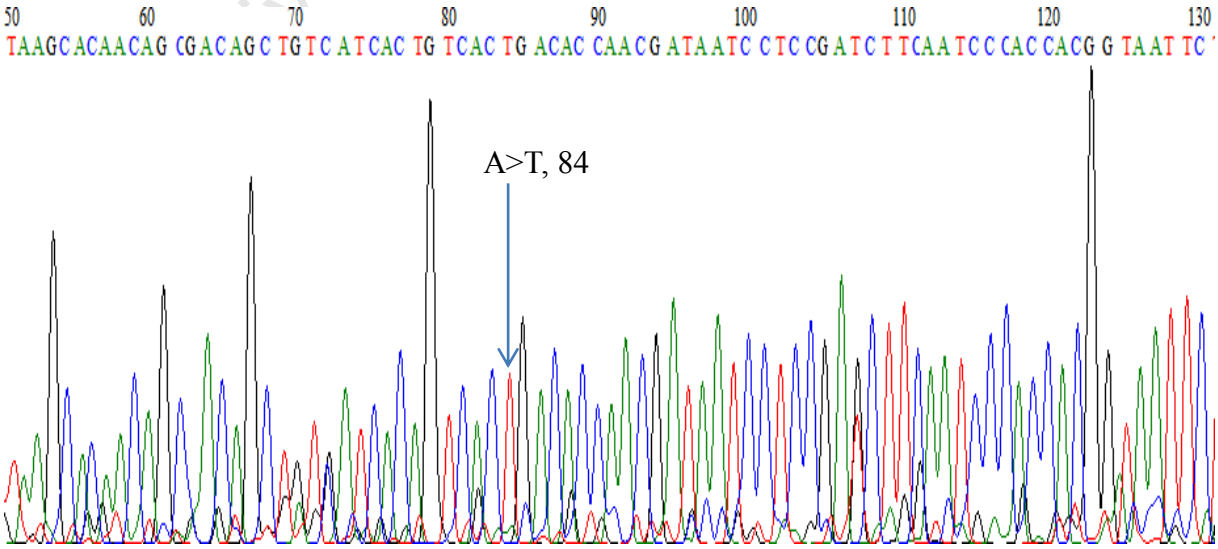


Figure 13: A representative chromatogram of sample T79 showing sequencing data as obtained from Macrogene.

Table 6: Nucleotide changes in the exon 8 of the *CDH1* gene.

Case numbers	Nucleotide change	Amino acid change	Position
T68	A>G	No	63
T79	A>T	No	84

3.6. Immunohistochemistry

Immunohistochemistry was performed in another study which was part of our comprehensive gastric research project conducted in the Division of Anatomical Pathology. The same gastric cancer cases were used. The E-cadherin immunohistochemical stains (Figure 14) for the two cases (T68 and T79) which presented with somatic mutations in this study were reviewed. The E-cadherin protein expression was analyzed by a pathologist (Dr Michael Locketz) for these two cases.

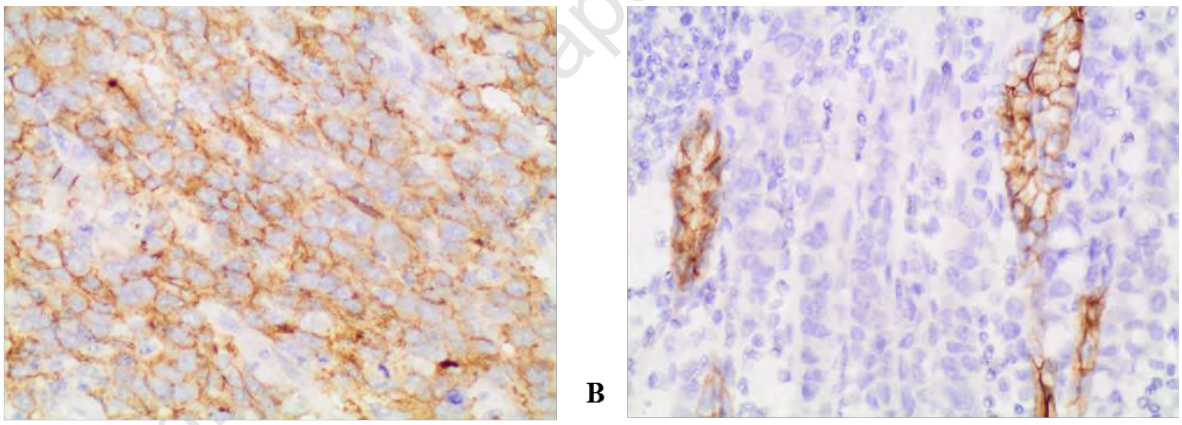


Figure 14: The immunohistochemistry stains showing the E-cadherin protein expression (with DAB chromogen). Sample T68 (**A**) - showing membrane positivity. Sample T79 (**B**) - showing absence or loss of E-cadherin.

CHAPTER 4

4. Discussion

4.1. Clinicopathologic features of tumours from gastric cancer patients

Gastric cancer is the fourth most frequent cancer having new cases estimated at 988,000 and 736,000 deaths with the majority of cases occurring in underdeveloped populations (Globacan, 2008). This disease has been rated as the second leading cause of cancer death worldwide (Parkin. 2001; Buffart et al., 2011). However, incidence and death rates vary greatly amongst different populations. The highest incidences of this cancer are seen in Africa, Eastern Asia and South America. Stomach cancer is less frequent in more developed regions such as Europe, Australia, New Zealand and North America (Globacan, 2008). The disease has also been seen as an important health problem in different populations of South Africa (Kidd eta l., 1999; Chetty et al., 2002; Dube et al., 2009).

The study cohort included the analysis of 102 cases for the different clinicopathologic features (lymph node metastasis, *Helicobacter pylori* infection, tumour grade and stage) that are commonly seen in gastric cancer (Appendix C). Further, the patients presented with different histologic types, diffuse, intestinal, mixed (diffuse/intestinal) and mucinous adenocarcinoma (Appendix C). Of the 102 patients, the majority 52.94% (54 patients) presented with lymph node metastasis, 25 patients (24.51%) did not have lymph node metastasis and 23 (21.56%) were not recorded for lymph node metastasis.

The *H.pylori* infection is an important environmental factor in the pathogenesis of stomach adenocarcinoma. With respect to *H. pylori* infection, the majority of our patients (74.51%) were negative for the infection, only 23.53% had the infection. Further, the *H. pylori* status was not recorded in 2 (1.96%) cases. Although, *H. pylori* is the major contributing factor to the development of gastric cancer (Bartchewsky et al., 2009), the frequency is low in our patient cohort. In addition, the clinical history of each patient was reviewed to check for the presence of *H. pylori* in previous biopsy reports. Only 2 of the 102 cases were positive for *H.pylori* infection.

With respect to the grade of the tumour, only 2% presented with grade 1 of the cancer. The frequency of grade 2 tumours was 28% and grade 3 tumours were seen in 31% of the cases. Grade 4 tumours were seen at a frequency of 30%. In line with our study, other studies have also shown that grade 3 and grade 4 tumours are commonly observed than grade 1 and 2 in gastric cancer patients (Suriano et al., 2003; Oliveira et al., 2009; Corso et al., 2013;).

When gender was correlated with the histologic type (Figure 6), the intestinal type was more common in males (62.16%) than females (37.84%) and the diffuse type was more frequent in females (61.11%) than males (38.89%). These results are in line with those observed by Correa Hackenson et al., (2010); Piazuelo, (2011) and Corso et al., (2012). However, our results may be biased, because the cases were collected to balance the two histologic types. When analysing the incidence of the cancer in different age groups (Figure 5), gastric cancer was seen at a frequency of 30.61% in patients younger than 50 years. The incidence was also more frequent (47.96%) in patients between the ages of 50 and 69.Additionally, the tumour was less common (21.43%) in the older age group (70 to 90 years).

4.2. Sequence analysis of the *CDH1* gene

The *CDH1* gene plays an important role in the pathogenesis of stomach cancer. Various studies have shown that this gene tends to accumulate both germline and somatic mutations in its coding sequence (Berx et al., 1998; Pharoah et al., 2001; Kaurah et al., 2007; Oliveira et al., 2009). However, the number of studies related to somatic mutations is sparse. Different tissue samples such as blood, fresh frozen material, surgical specimens and FFPE tissues have been used to isolate nucleic acids (DNA or RNA) for molecular analysis. Our study used FFPE tissues as a source of DNA. We were able to successfully amplify only 44 of the 102 samples included in the study. Of the 44 cases which had good quality DNA (Figure 9), we managed to amplify only 26 DNA samples for exon 8 of the *CDH1* gene (Figure 10). With regards to exon 9 however, only two samples were amplifiable (Figure 11). Further, these 2 samples produced faint DNA bands (low DNA concentration). It is possible that no sequencing data was obtained in these samples because of the quality of DNA. The quality of DNA

isolated from FFPE tissues could be the reason for unsuccessful amplification in exon 9. It is also well established that the formalin fixative forms cross links which compromise the DNA integrity (Thompson et al., 2005).

Of the 26 cases analysed, *CDH1* somatic mutations were seen in only 2 (7.6%) cases (Figure 12 and13) and 24 of our patients did not have nucleotide changes (Appendix A). These genetic mutations were single nucleotide changes that occurred at position 63 and 84. The nucleotide changes were A>G and A>T in positions 63 and 84, respectively (Table 6). Our results are different from those of (Becker et al., 1994; Becker et al., 1998; Ascano et al., 2001; Machado et al., 2001; Oliveira et al., 2009 and Corso et al., 2013) that reported a higher frequency (28% to 75%) of *CDH1* somatic mutations. Becker and colleagues identified *CDH1* somatic mutations amongst diffuse gastric cancers, in 13 of 26 cases (50%).

A study by Machado et al., (2001) showed a slightly higher prevalence of mutations in 16 of 23(56.3%) cases analysed. Further, these mutations were observed only in the diffuse type (Machado et al., 2001). The differences in the frequencies of *CDH1* somatic mutations could be due to a number of factors; the sample size and type of material used. Other investigations have used surgical specimens and fresh frozen tissues as a source of DNA, as opposed to our study which used FFPE tissues. DNA isolated from fresh frozen material has a better quality than that isolated from FFPE tissues.

Since the aetiology of gastric adenocarcinoma has also been shown to be associated with other environmental and genetic factors (Crew and Neugut, 2006), it is possible that dietary factors have contributed to the pathogenesis of gastric cancer in the Western Cape. In high incidence regions like Japan, the development of this malignancy is mostly associated with poor food storage and diet. The Japanese usually utilize salt to preserve their food and therefore ingestion of salty foods and smoked fish was shown to contribute to gastric malignancy. Additionally, foods containing chemicals such as nitrite compounds, low fruit and vegetable uptake is also a risk factor of the disease (Brenner et al., 2009). Alcohol consumption and smoking have also been related to the pathogenesis of this malignancy (Inoue and Tsugane, 2005). Moreover, these factors might have played a role in the development of gastric cancer amongst the patients seen in the Western Cape, however this requires further investigations.

Other genetic factors like microsatellite instabilities, polymorphisms and germline mutations also contribute to the development of the disease. However, the prevalence of *E-cadherin* germline mutations ranges from 29% to 56% (Lynch et al., 2005; Bacani et al., 2006; Corso et al., 2012). Patients who do not present with germline mutations are further analysed for large genomic deletions. These deletions have therefore been implicated as another possible contributing factor towards the pathogenesis of gastric cancer amongst the *E-cadherin* germline negative individuals (Oliveira et al., 2009; Pinheiro et al., 2010; Yamada et al. 2011). It would therefore be interesting to explore the *CDH1* large genomic deletions for future research. Another possible explanation is the presence of mutations in other genes namely, *p53*, *STK11*, *MET*, *MLH1* and *MLH2* (Yamada et al., 2011).

4.3. Association of *CDH1* mutations with Clinicopathologic data and protein expression

The 2 cases which had *CDH1* mutations (T68 and T79) were both female patients and had the diffuse histologic type (Table 5). Case T68 was 36 years old and did not have *H.pylori* infection. Additionally, the tumour in this patient had spread to the serosa (stage III) of the stomach. On the other hand, case T79 was 64 years old and had the *H. pylori* infection. The tumour had spread to the lymph nodes and also the serosa of the stomach. These results may indicate that these mutations may be associated with an advanced stage of gastric cancer (stage III and IV), which is in line with the study by Corso et al., (2013). However, due to the low number of cases with mutations in our study, we cannot conclusively make any statistical correlations. It is interesting to note that other studies have shown similar results (Ascano et al., 2001; Machado et al., 2001; Corso et al., 2013).

In sporadic gastric cancers, *CDH1* gene somatic mutations are commonly associated with loss of E-cadherin protein expression (Ascano et al., 2001; Machado et al., 2001; Brooks-Wilson et al., 2004). In our study, E-cadherin protein expression in T79 (Figure 14 **B**) was lost; patient T68 (Figure 14 **A**) on the other hand still expressed the E-cadherin protein. Although other studies have shown an association between *CDH1* gene mutations and loss of E-cadherin protein expression, our study cannot make any

definite conclusions since only two samples were analysed.

In conclusion, our study showed that *CDH1* genetic mutations were only observed in 2 of the 26 samples (7.6%). The *CDH1* genetic mutations are therefore not common in gastric cancer patients seen in the Western Cape. There was no correlation between the sequencing data and the clinicopathologic data. To our knowledge this is the first study that reports somatic mutations of the *CDH1* gene in a cohort of gastric cancers using FFPE tissue in the South African setting.

For future research, it would be interesting to use fresh tissue samples obtained from a prospective study, together with FFPE tissues as a source of DNA. Other genes (*p53*, *STK11*, *MET*, *MLH1* and *MLH2*) may also be analysed for both germline and somatic mutations, to determine if they are possible contributing factors in gastric carcinogenesis. Further, large genomic deletions in patients negative for germline mutations may be interesting to investigate as they have also been shown as contributing factors.

References

Aberle H, Bauer A, Stappert J et al (1997). β-catenin is a target for the ubiquitin-proteasome pathway. *The EMBO Journal* **16**, 3797-3804.

Akaboshi S, Watanabe S, Hino Y et al (2009). HMGA1 is induced by Wnt/ß-catenin pathway and maintains cell proliferation in gastric cancer. *The American Journal of Pathology*, 175:4.

Al-Moundhri M, Al-Khanbashi M, Al-Kindi M et al (2010). Association of *E-cadherin* (*CDH1*) gene polymorphisms and gastric cancer risk. *World J Gastroenterology***16,** 3432-3436.

An J, Kim H, Cheong J et al (2011). Microsatellite instability in sporadic gastric cancer: its prognostic role and guidance for 5-FU based chemotherapy. *International Journal of Cancer* **131**:505-511.

Apostolou P, Tsantsaridou A, Papasotiriou L et al (2011). Bacterial and fungal microflora in surgically removed lung cancer samples. *Journal of Cardiothoracic Surgery* **6,** 137.

Ascano J, Frierson H, Moskaluk C et al (2001). Inactivation of the *E-cadherin* gene in sporadic gastric cancer. *Mod Pathol* **14,** 942-949.

Bacani J, Soares M, Zwingerman R et al (2006). *CDH1/E-cadherin* germline mutations in early-onset gastric cancer. *Journal of Medical Genetics* **43,** 867-872.

Bartchewsky W, Martini M, Masiero M et al (2009). Effect of Helicobacter pylori infection on IL-8, IL-1β and COX-2 expression in patients with chronic gastritis and gastric cancer. *Scandinavian Journal of Gastroenterology* **44,** 153-161.

Bartek J, Lukas J and Bartkova J (1999). Perspective:defects in cell cycle control and cancer. *Journal of Pathology* **187,** 95-99.

Becker K, Atkison M, Reich U et al (1994). *E-cadherin* gene mutations provide clues to diffuse gastric carcinoma. *Cancer Research* **54,** 3845-3852.

Becker K, Reich U, Schott C et al (1998). Identification of eleven novel tumour-associated E-cadherin mutations. *Human Mutation* **215,** 1-5.

Beghelli S, Manzoni G, Barbi S et al (2006). Microsatellite instability in gastric cancer is associated with better prognosis in only stage II cancers. *Surgery*, 347-356.

Berx G, Becker K, Hofler H et al (1998). Mutations of the E-cadherin (*CDH1*) gene. *Human Mutation* **12,** 226-237.

Berx G, Staes K, van Hengel Y et al (1995). Cloning and characterization of the human invasion suppressor gene *E-cadherin* (*CDH1*). *Genomics* **26,** 281-289.

Burne R and Chen Y (2000). Bacterial ureases in infectious diseases. *Microbes and Infection* **2,** 533-542.

Brooks-Wilson A, Kaurah P, Suriano G et al (2004). Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. *J Med Genet* **41,** 508-517.

Carnero A (2002). Targeting the cell cycle for cancer therapy. *British Journal of Cancer* **87,** 129-133.

Chan A, Wong B, Lan H et al (2003). Deregulation of E-cadherin-catenin complex in precancerous lesions of gastric adenocarcinoma. *Journal of Gastroenterology and Hepatology* **18,** 534-539.

Chang W, Yang H, Cheng H et al (2011). Increased gastric osteopontin expression by *Helicobacter pylori* infection can correlate with more severe gastric inflammation and intestinal metaplasia. *Helicobacter* **16,** 217-224.
Chen M, Kingham K, Ford J et al (2011). A prospective study of total gastrectomy for CDH1-positive hereditary diffuse gastric cancer. *Ann Surg Oncol* **18,** 2594-25298.
Chetty R, Naidoo R, Tarin M and Sitti C (2002). Chromosome 2p, 3p, 5q and 18q status in sporadic gastric cancer. Pathology **34**, 275-281.
Choi Y, Choi H, Jeon W et al (2012). Optimal number of endoscopic biopsies in diagnosis of advanced gastric and colorectal cancer. *Journal of Korean Medical Science* **27,** 36-39.
Clements W, Wang J, Sarnaik A et al (2002). β-catenin mutation is a frequent cause of Wnt pathway activation in gastric cancer. *Cancer Research* **62,** 3503-3506.
Collins K, Jacks T and Pavletich N (1997). The cell cycle and cancer. *Proc Natl Acad Sci USA* **94,** 2776-2778.
Compare D, Rocco A, Ligouri E et al (2011). Global DNA hypomethylation is an early event in *Helicobacter pylori*-related gastric carcinogenesis. *J Clin Pathol* **64,** 677-682.
Coup A, Cross S, Culligan D et al. General and systemic pathology.
Corso G, Carvalho J, Marreli D et al (2013). Somatic mutations and deletions of the E-cadherin gene predict poor survival of patients with gastric cancer. *Journal of Clinical Oncology* **31,** 868-875.
Corso G, Seruca R and Roviello F (2012). Gastric cancer carcinogenesis and tumour progression. *Annali Italiani di Chirurgia* **83,** 172-176.
Correa P (2003). Bacterial infections as a cause of cancer. *Journal of the National Cancer Institute* **95**.
Correa P and Piazuelo M (2011). Helicobacter pylori infection and gastric adenocarcinoma. *US Gastroenterol Hepatol Rev* **7,** 59-64.
Czyzewska J, Guzinska-Ustymowicz K, Ustymowicz M et al (2010). The expression of E-cadherin-catenin complex in patients with advanced gastric cancer: role in formation of mestastasis. *Folia Histochem Cytobiol* **48,** 37-45.
Das P and Singal R (2004). DNA methylation and cancer. *Journal of Clinical Oncology* **22,** 4632-4642.
Doucas H, Garcea G, Neal CP et al (2005). Changes in the Wnt signalling pathway in gastrointestinal cancers and their prognostic significance. *European Journal of Cancer* **41,** 365-379.
De Calisto J, Araya C, Marchant L et al (2003). Essential role of non-canonical Wnt signalling in neural crest migration. *Development* **132,** 2587-2597.
De Stafani E, Correa P, Boffetta P et al (2004). Dietary patterns and risk of gastric cancer: a case-control study in Uruguay. *Gastric Cancer* **7,** 211-220.
Dube C, Nkosi T, Clarke A, Mkwetshana N, Green E and Ndip R (2009). Helicobacter pylori antigenemia in an asymptomatic population of Eastern Cape province, South Africa: Public health implications. *Reviews on environmental health* **24,** 249-255.
Duell E, Travier N, Lujan-Barroso L et al (2011). Alcohol consumption and gastric cancer risk in the European prospective investigation into cancer and nutrition (EPIC) cohort. *Am J Clin Nutr* **94,** 1266-1275.
Dumitrescu R (2011). DNA methylation and histone modifications in breast cancer.

Erkan G, Gonul I, Kandilci U et al (2012). Evaluation of apoptosis along with BCL-2 and Ki-67 expression in patients with intestinal metaplasia. *Pathology – Research and Practice* **208,** 89-93.
Eriksson N, Karkkainen P, Farkkila M and Arkkila P (2008). Prevalence and distribution of gastric intestinal metaplasia and it's subtypes. *Digestive and liver disease* **40,** 355-360.
Finn K, Lowndes N and Grenon M (2012). Eukaryotic DNA damage checkpoint activation in response to double-strand breaks. *Cell Mol Life Sci* **69,** 1447-1473.
Forx J and Wang T (2007). Inflammation, atrophy and gastric cancer. *The Journal of Clinical Investigation* **117,** 60-69.
Garay J, Bravo J, Correa P et al (2004). Infrequency of microsatellite instability in complete and incomplete gastric intestinal metaplasia. *Human Pathology* **35,** 102-106.
Gavazzi C, Colatruglio S, Sironi A et al (2011). Importance of early nutritional screening in patients with gastric cancer. *British Journal of Nutrition* **106,** 1773-1778.
Gencer S, Sen G, Dogusoy G (2010). β-catenin-independent non-canonical Wnt pathway might be induced in gastric cancer. *Turk J Gastroenterol* **21,** 224-230.
Grander D (1998). How do mutated oncogenes and tumour suppressor genes cause cancer? *Medical Oncology* **15,** 20-26.
Graziano F, Arduini A, Ruzzo A et al (2004). Combined analysis of E-cadherin gene (*CDH1*) promoter hypermethylation and E-cadherin protein expression in patients with gastric cancer: implications for treatment with demethylating drugs. *Annals of Oncology* **15,** 489-492.
Graziano F, Humar B and Guilford P (2003). The role of the E-cadherin gene (*CDH1*) in diffuse gastric cancer susceptibility:from the laboratory to the clinical practice. *Annals of Oncology* **14,** 1705-1713.
Greene F, Page D, Balch C et al (2002). AJCC cancer staging handbook: from the AJCC cancer staging manual (6^{th} edition).
Guo W, Dong Z, Guo Y et al (2009). Detection of the promoter hypermethylation of the CpG islands of E-cadherin in gastric cardiac gastric adenocarcinoma. *Eur J Med Res* **14,** 453-458.
Guzman P, Araya J, Villaseca M et al (2006). Immunohistochemical expression of the E-cadherin-catenin complex in gastric cancer. *Revista Medica De Chile* **134,** 1002-1009.
Hackenson D, Edelman D, McGuire T et al (2010). Prophylactic laparoscopic gastrectomy for hereditary diffuse gastric cancer: a case series in a single family. *Journal of the Society of Laparoendoscopic Surgeons* **14,** 348-352.
Hino R, Uozaki H, Murakami N et al (2009). Activation of DNA methyltransferase 1 by EBV latent membrane protein 2A leads to promoter hypermethylation of *PTEN* gene in gastric carcinoma. *Cancer Res* **69,** 2766-2774.
Hoeijmakers J (2001). Genome maintenance mechanisms for preventing cancer. *Nature* **411,** 366-374.
Horri A, Nakatsuru S, Miyoshi Y et al (1992). The APC gene responsible for familial adenomatous polypopsis, is mutated in gastric cancer. *Cancer Res* **52,** 3231-3233.
Hottenrott C (2012). Histology classification challenges for the endoscopic treatment of early gastric cancer. *Surg Endosc* **26,** 2092-2093.
Huff J (1993). Chemicals and cancer in humans: first evidence in experimental

evidence. *Environmental Health Perspectives* **100,** 201-210.
Jacobs G, Hellmig S, Huse K et al (2011). Polymorphisms in the 3'-untranslated region of the *CDH1* gene are a risk factor for primary gastric diffuse large B-cell lymphoma. *Haematological* **96,** 987-995.
Jakszyn P, Agudo A, Lujan-Borroso L et al (2011). Dietary intake of heme iron and risk of gastric cancer in the European prospective investigation into cancer and nutrition study. *International Journal of Cancer* **130,** 2654-2663.
Jeong C, Lee J, Sohn S et al (2010). Mitochobdrial microsatellite instability in gastric cancer and gastric epithelial dysplasia as a precancerous lesion. *Cancer Epidemiology* **34,** 323-327.
Jones P and Baylin S (2002). The fundamental role of epigenetic events in cancer. *Nature Reviews Genetics* **3,** 415-427.
Joo Y, Park C, Kim H e al (2000). Prognostic significance of E-cadherin /catenin complex expression in gastric cancer. *J Korean Med Sci* **15,** 655-666.
Joo Y, Rew J, Kim H et al (2001). Changes in the E-cadherin-catenin complex expression in early and advanced gastric cancers. *Digestion* **64,** 111-119.
Kang G, Lee H, Hwang K et al (2003). Aberrant CpG Island hypermethylation of chronic gastritis, in relation to aging, gender, intestinal metaplasia and chronic inflammation. *American Journal of Pathology* **163,** 1551-1555.
Kang J, Lee B, Kim N et al (2011). CDX1 and CDX2 expression in intestinal metaplasia, dysplasia and gastric cancer. *J Korean Med Sci* **26,** 647-653.
Kang S, Han J, Ahn M et al (2011). Helicobacter pylori infection as an independent prognostic factor for locally advanced gastric cancer patients treated with adjuvant chemotherapy after curative resection. *International Journal of Cancer* **130,** 948-958.
Kaurah P, MacMillan A, Boyd N et al (2007). Founder and recurrent *CDH1* mutations in families with hereditary diffuse gastric cancer. *Journal of American Medical Association* **297,** 2360-2373.
Kawabata T, Mizuno T, Okazaki S et al (2008). Optical diagnosis of gastric cancer using near-infrared multichannel Raman spectroscopy with a 1064-nm excitation wavelength. *Journal of Gastroenterology* **43,** 283-290.
Kawakami T, Ren S and Duffield J (2013). Wnt signalling in kidney diseases: dual roles in renal injury and repair. *Journal of Pathology* **229,** 221-223.
Kelly J and Duggan J (2003). Gastric cancer epidemiology and risk factors. *Journal of Clinical Epidemiology* 56, 1-9.
Key T, Allen E, Spencer E et al (2002). The effect of diet on risk of cancer. *Lancet* **360,** 861-868.
Kidd M, Lastovica A, Atherton J and Louw J (1999). Heterogeneity in the *Helicobater pylori vacA* and *cagA* genes: association with gastroduodenal disease in South Africa (1999). *Gut* **45,** 499-502.
Kim S, Ruiz V, Carroll J et al (2011). *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Letters* **305,** 228-238.
Kim Y, Chung J, Lee S et al (2010). Progression from chronic atrophic gastritis to gastric cancer; tangle, toggle, tackle with Korea Red Ginseng. *J Clin Biochem. Nutr* **46,** 195-204.
Kluijt I, Sijmons R, Hoogerbrugge N et al (2012). Familial gastric cancer: guidelines for diagnosis, treatment and period surveillance. *Familial Cancer* **11,** 363-369.
Kristensen L, Wojdacz T, Thestrup B et al (2009). Quality assessment of DNA derived from up to 30 years old formalin fixed paraffin embedded (FFPE)

tissues for PCR-based methylation analysis using SMART-MSP and MS-HRM. *BMC Cancer* **9,** 453.
Kubota K, Kuroda J, Yoshida M et al (2012). Medical image analysis: computer-aided diagnosis of gastric cancer invasion on endoscopic images. *Surgical Endoscopy* **26,** 1485-1489.
Kudo Y, Moroshashi S, Takasugi K et al (2011). Histopathological phenotypes of early gastric cancer and its background mucosa. *Biomedical Research* **32,** 127-134.
Kusters J, van Vliet A and Kuipers E (2006). Pathogenesis of *Helicobacter pylori* infection. *Clin Microbiol* **19,** 449-490.
Lai D, Visser-grieve S and Yang X (2012). Tumour suppressor genes in chemotherapeutic drug response. *Biosci Rep* **32,** 361-374.
Langauer C, Kinzler K and Volgestein B (1998). Genetic instabilities in human cancers. *Nature* **396,** 643-649.
Leal M, Lima E, Silva P et al (2007). Promoter hypermethylation of *CDH1*, *FHIT*, *MTAP* and *PLAGL1* in gastric adenocarcinoma in individuals from Northern Brazil. *World J Gastroenterol* **13,** 2568-2574.
Leibeling D, Laspe P and Emmert S (2006). Nucleotide excision repair and cancer. *J Mol Hist* **37,** 225-238.
Leung W and Sung J (2002). Review article: intestinal metaplasi and gastric carcinogenesis. *Aliment Pharmacol Ther* **16,** 1209:1216.
Li C and Li Y (2010). Endomicroscopy of intestinal metaplasia and gastric cancer. *Gasroenterology Clinics of North America* **39,** 785-796.
Li Y, Zhang K, Sun Y et al (2008). Frequent S100A4 expression with unique splicing pattern in gastric cancer: a hypomethylation event paralleled with E-cadherin reduction and Wnt activation. *Translational oncology* **1,** 165-176.
Li Y, Tian Z, Zhang J and Fu B (2012). *CDH1* promoter polymorphism and stomach cancer susceptibility. *Mol Biol Rep* **39,** 1283-1286.
Lijinsky W (1999). N-nitroso compounds in the diet. *Mutation Research* **443,** 129-138.
Lindgren A, Yun C, Sjoling A et al (2011). Impaired IFN-γ production after stimulation with bacterial components by natural killer cells from gastric cancer patients. *Experimental Cell Research* **317,** 849-858.
Lowe S and Lin A (2000). Apoptosis in cancer. *Carcinogenesis* **21,** 485-495.
Ludyga N, Grunwald B, Azimzadeh O et al (2012). Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analysis. *Virchows* **460,** 131-140.
Lynch H, Aldoss I and Lynch J (2011). The identification and management of hereditary diffuse gastric cancer in a large Jordanian family. *Familial Cancer* **10,** 667-672.
Lynch H, Grady W, Suriano G and Huntsman G (2005). Gastric cancer: new genetic developments. *Journal of Surgical Oncology* **90,** 114-133.
Machado J, Oliveira C, Carvalho R et al (2001). *Ecadherin* gene (*CDH1*) promoter methylation as the second hit in sporadic diffuse gastric carcinoma. **Oncogene 20** 1525-1528.
Micu G, Staniceanu F, Zurac S et al (2010). E-cadherin and β-catenin expression in gastric neoplastic and non-neoplastic lesions-correlations with *H.pylori* infection. *Rom J Intern Med* **48,** 271-280.
Macleod K (2000). Tumour suppressor genes. *Current Opinion in Genetics and Development* **10,** 81-93.

Matsunari O, Shiota S, Suzuki R et al (2012). Association between *Helicobacter pylori* virulence and gastroduodenul diseases in Okinawa, Japan. *Journal of Clinical Microbiology* **50,** 876-883.
Mclaughlin-Drubin M and Munger K (2008). Viruses associated with human cancer. *Biochemica et Biophysica Acta* **1782,** 127-150.
Mohammed S, Springfiled S and Das R (2012). Role of epigenetics in cancer health disparities. *Cancer Epienetics* **863,** 395-410.
Moy K, Fan Y, Wang R et al (2010). Alcohol and tobacco use in relation to gastric cancer: a prospective study of men in Shanghai, China. *Cancer Epidemiol Biomarkers Prev* **19,** 2287-2297.
More H, Humar B, Weber W et al (2007). Identification of seven Novel germline mutations in the human E-cadherin (*CDH1*) gene. *Human mutation* **944,** 1-9.
Muretto P, Ruzzo A, Pizzagalli F et al (2008). Endogastric capsule for E-cadherin gene (*CDH1*) promoter hypermethylation assessment in DNA from gastric juiceof diffuse gastric cancer patients. *Annals of Oncology* **19,** 516-519.
Nadauld L and Ford J (2013). Molecular profiling of gastric cancer: toward personalized cancer medicine. *Journal of Clinical Oncology* **31,** 1-3.
Nam J, Choi I, Co S et al (2012). Associations of the interval between endoscopies with gastric cancer stage at diagnosis in a region of high prevalence. *Cancer* 4953-4960.
Nam J, Choi I, Co S et al (2011). Helicobacter pylori infection and histological changes in siblings of young gastric cancer patients. *Journal of Gastroenterology and Hepatolog* **26,** 1157-1163.
Nan H, Song Y, Yun H et al (2005). Effects of dietary intake and genetic factors on hypermethylation of the hMLH1 gene promoter in gastric cancer. *World J Gastroenterol* **11,** 3834-3841.
Nasri S, More H, Graziano F et al (2008). A novel diffuse gastric cancer susceptibility variant in E-cadherin (CDH1) intron 2: a case control study in an Italian population. *BMC Cancer* **8,** 138.
Ngoan L, Mizoue T, Fujino Y et al (2002). Dietary factors and stomach cancer mortality. *British Journal of Cancer* **87,** 37-42.
Nollet F, Kools P and Van Roy F (2000). Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J. Mol. Biol* **299,** 551-572.
Nomura A, Wilkens L, Henderson B et al (2012). The association of cigarette smoking with gastric cancer: the multiethnic cohort study. *Cancer Causes Control* **23,** 51-58.
Oishi I, Suzuki H, Onishi N et al (2003). The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway. *Genes to cells* **8,** 645-654.
Okayasu R (2011). Repair of DNA damage induced by accelerated heavy ions – a mini review. *International Journal of Cancer* **130,** 991-1000.
O'Leary D, Wang J, Cotter T et al (2011). Less stress, more success? Oncological implications of surgery induced oxidative stress. *Gut* 1-10.
Oliveira M, Costa A, Costa A et al (2009). CagA associates with c-Met, E-cadherin, and p120-catenin in a multiproteic complex that suppresses *Helicobacter pylori*-induced cell invasive phenotype. *The Journal of Infectious Diseases* **200,** 745-755.
Osawa H, Yamamoto H, Miura Y et al (2012). Diagnosis of depressed-type early gastric cancer using smaller-caliber endoscopy with flexible spectral imaging

color enhancement. *Digestive endoscopy* **24,** 231-236.
Ottini L, Falchetti M, Saieva C et al (2004). MRE11 expression is important in gastric cancer with microsatellite instability. *Carcinogenesis* **25,** 2337-2343.
Oue N, Motoshita J, Yokozaki H et al (2002). Distinct promoter hypermethylation of $p16^{INK4a}$, *CDH1* and RAR-beta in intestinal, diffuse-adherent, and diffuse-scattered type gastric carcinomas. *Journal of Pathology* **198,** 55-59.
Padmavathy F, Siddaraju N and Sistla S (2011). Distinction of intestinal and diffuse types of gastric adenocarcinoma on brush cytology. *Acta Cytology* **55,** 187-192.
Park B, Shin A, Park S et al (2011). Ecological study for refrigerator use, salt, vegetable and fruit intakes and gastric cancer. *Cancer causes control* **22,** 1497-1502.
Pedrazzani C, Corso G, Velho S et al (2009). Evidence of tumour microsatellite instability in gastric cancerwith familial aggregation. *Familial Cancer* **8,** 215-220.
Perez R, Jacob C, L'ofreddo F et al (2004). Microsatellite instability in solitary and sporadic gastric cancers. *Rev Hos Clin Fac Med S Paulo* **59,** 279-285.
Pharoah P, Guilford P, Caldas C et al (2001). Incidence of gastric cancer and breast cancer in *CDH1* (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology* **121,** 1348-1353.
Pinho S, Seruca R, Gartner F et al (2011). Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cell Mol Life Sci* **68,** 1011-1020.
Pinheiro H, Bordeira-Carrico R, Seixas S et al (2010). Allele-specific CDH1 downregulation and hereditary diffuse gastric cancer. *Human Molecular Genetics* **19,** 943-952.
Pokutta S and Weis W (2007). Structure and mechanisms of cadherins and catenins in cell-cell contacts. *Annu Rev Cell Vev Biol* **23,** 237-261.
Polakis P 2000 (2000). Wnt signalling and cancer. *Genes and development* **14,** 1837-1851.
Qiu H, Zhang L, Keshari R et al (2010). Relationship between *H.pylori* infection and clinicopathological features and prognosis of gastric cancer. *MBC Cancer* **10,** 374-382.
Renehan A, Booth C and Potten C (2001). What is apoptosis and why is it important? *BMJ* **322,** 1536-1538.
Richards F, McKee S, Rajpar H et al (1999). Germline E-cadherin mutations predispose to familial gastric cancer and colorectal cancer. *Human Molecular Genetics* **8,** 607-610.
Serrano C, Diaz M, Valdivia A et al (2007). Relationship between Helicobacter pylori virulence factors and regulatory cytokines as predictors of clinical outcome. *Microbes and infection* **9,** 428-434.
Shokal U and Sharma P (2012). Implication of microsatellite instability in human gastric cancers. *Indian J Med Res* **135,** 599-613.
Shore E and Nelson W (1991). Biosynthesis of the cell adhesion molecule Uvomorulin (E-cadherin) in Madin-Darby Canine kidney epithelial cells. *The Journal Of Medical Chemistry* **266,** 19672-19680.
Song B, Park Y, Kim H et al (2004). Expression of beta-catenin and E-cadherin in early gastric cancer:correlation with clinicopathologic parameters. *Korean Journal of Gastroenterology* **43,** 1598-9992.
Stingl K, Altendorf K and Bakker E (2002). Acid survival of Helicobacter

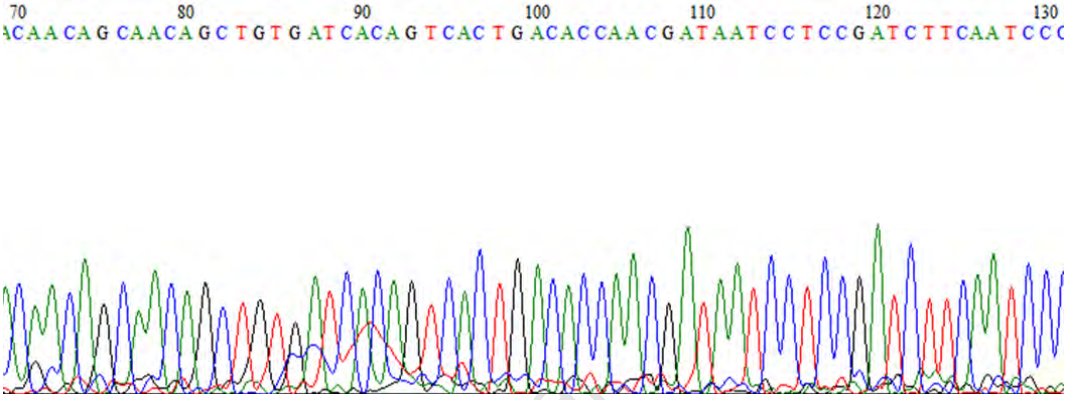
pylori: how does urease activity trigger cytoplasmic ph Homeostasis? *Trends in Microbiology* **10,** 70-74.
Stockfleth E, Nindl I, Sterry W et al (2004). Human papillomaviruses in transplant-associated skin cancers. *Dermatol Surg* **30,** 604-609.
Sugimura R and Li L (2010). Non-canonical Wnt signalling in vertebrate development, stem cell and disease. *Birth Defects Research* **90,** 243-256.
Suriano G, Mulholland D, de Wever O et al (2003). The intracellular E-cadherin germline mutation V832M lacks the ability to mediate cell-cell adhesion and to suppress invasion. *Oncogene* **22,** 5716-5719.
Tada M, Concha M and Heisenberg C (2002). Non-canonical signalling and regulation of gastrulation movements. *Cell and Developmental Biology* **13,** 251-260.
Tamura G, Sakata K, Nishizuka S et al (1996). Inactivation of the E-cadherin gene in primary gastric carcinomas and gastric carcinoma cell lines. *Cancer Res* **87,** 1153-1159.
Thompson E, Hebert S, Forrest S et al (2005). Whole genome SNP arrays ising DNA derived from Formalin-Fixed, Paraffin-Embedded ovarian tumour tissues. *Human mutation* **26,** 384-389.
Tomita H, Hirata A, Yamada Y et al (2010). Suppressive effect of global DNA hypomethylation on gastric carcinogenesis. *Carcinogenesis* **31,** 1627-1633.
Tsugane S and Sasazuki S (2007). Diet and the risk of gastric cancer: review of epidemiological evidence. *Gastric Cancer* **10,** 75-83.
Tsukanov V, Butorin N, Maady A et al (2011). Helicobacter pylori infection, intestinal metaplasia and gastric cancer risk in Eastern Siberia. *Helicobacter* **16,** 107-112.
Van Aken E, De Weaver O et al (2001). Defective E-cadherin/catenin complexes in human cancer. *Virchows Arch* **439,** 725-751.
Van Roy F and Berx G (2008). The cell-cell adhesion molecule E-cadherin. *Cell. Mol .Life Sci* **65,** 3756-3788.
Verma M (2012). Epigenetic Biomarkers in Cancer Epidemiology. *Cancer Epigenetics* **863,** 467-480.
Vermeulen K, Bockstaele D and Bernemen Z (2003). The cell cycle:a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif* **36,** 131-149.
Wang C, Chen J, Yang Z et al (2003). Correlation of E-cadherin, beta-catenin and gamma-catenin gene expression with the invasion and metastasis of gastric carcinoma. *Di Yi Jun Yi Da Xue Xue Bao* **23,** 1198-1201.
Watabe H, Mitsushima T, Yamaji Y et al (2005). Predicting the development of gastric cancer from combining *Helicobacter pylori* antibodies and serum pepsinogen status: a prospective endoscopic cohort study. *Gut* **54,** 764-768.
Wen S and Moss S (2009). Helicobacter pylori virulence factors in gastric carcinogenesis. *Cancer Letters* **282,** 1-8.
Witkowska and Smolewski (2013). *Helicobacter pylori* infection, chronic inflammation, and genomic transformations in gastric MALT lymphoma. *Mediators of inflammation*.
Wu P, Zhang Z, Wang J et al (2012). Germline promoter methylation of tumour suppressor genes in gastric cancer. *World J Gastroenterol* **18,** 70-78.
Yamada H, Shinmura K, Ito H et al (2011). Germline alterations in the *CDH1* gene in familial gastric cancer in the Japanese population. *Cancer Sci* **102,** 1782-1788.

Yamamoto E, Suzuki H, Takamaru H et al (2011). Role of DNA methylation in the development of diffuse-type gastric cancers. *Digestion* **83,** 241-249.
Yamanaka H, Moriguchi T, Masuyama N et al (2002). JNK functions in the non-canonical Wnt pathways to regulate convergent extension movements in vertebrates. *European Molecular Biology Organization* **31,** 69-75.
Yamazato T, Oyama T, Yoshida T et al (2012). Two years' intensive training in endoscopic diagnosis facilitates detection of early gastric cancer. *Internal Medicine* **51,** 1461-1465.
Yeh J, Goldie S, Kuntz K and Ezzati M (2009). Effects of Helicobacter pylori infection and smoking on gastric cancer incidence in China: a population-level analysis of trends and projections. *Cancer Causes Control* **20,** 2021-2029.
Yoon H, Kim N, Lee H et al (2011). *Helicobacter pylori*-negative gastric cancer in South Korea: Incidence and clinicopathologic characteristics. *Helicobacter* **16,** 382-388.
Zazula M, Ferreira A, Czopek J et al (2006). *CDH1* gene promoter hypermethylation in gastric cancer. *Diagn Mol Pathol* **15,** 24-29.
Zha L, Zhang J, Tang W et al (2012). HMGA2 elicits EMT by activating the Wnt/β-catenin Pathway in gastric cancer. *Dig Dis Sci*.
Zhan Z, Wu J, Zhang J et al (2012). *CDH1* gene polymorphisms, plasma CDH1 levels and risk of gastric cancer in a Chinese population. *Mol Biol Rep* **39,** 8107-8113.
Zhang B, Pan K, Liu Z et al (2008). Genetic polymorphism of the E-cadherin gene promoter and risk of sporadic gastric carcinoma in Chinese populations. *Cancer Epidemiology Biomarkers Prev* **17,** 2402-2408.
Zhang F, Hou L, Terry M et al (2007). Genetic polymorphisms in alcohol metabolism, alcohol intake and the risk of stomach cancer in Warsaw, Poland. *Int J Cancer* **121,** 2060-2064.
Zhang J, Guo S and Duan Z (2011). Application of magnifying narrow-band imaging endoscopy for diagnosis of early gastric cancer and precancerous lesion. *BioMed Central Gastroenterology* **11,** 135.
Zhao B and Bu X(2005). Promoter methylation of tumour-related genes in gastric cancer.

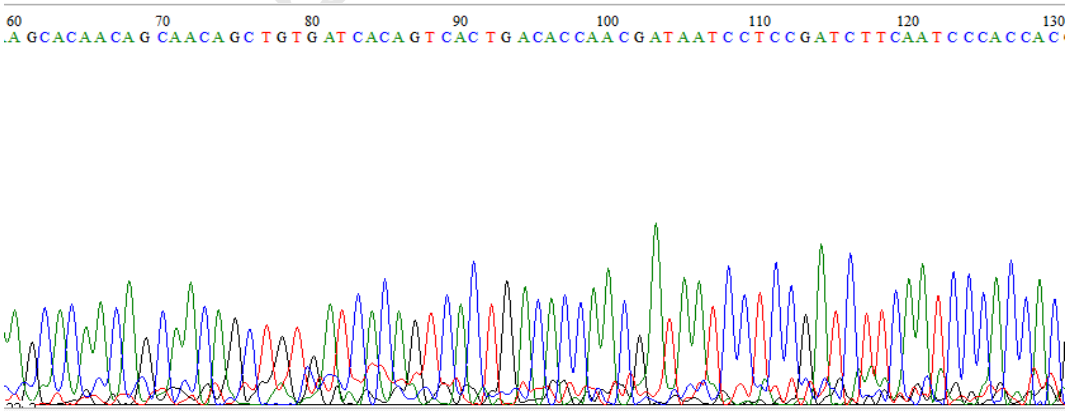
Appendices

Appendix A: The chromatograms of the 26 cases

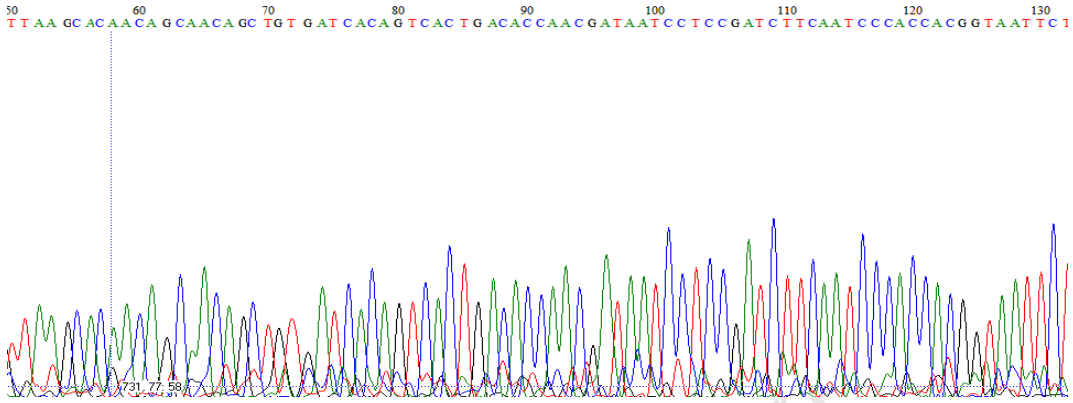
T2



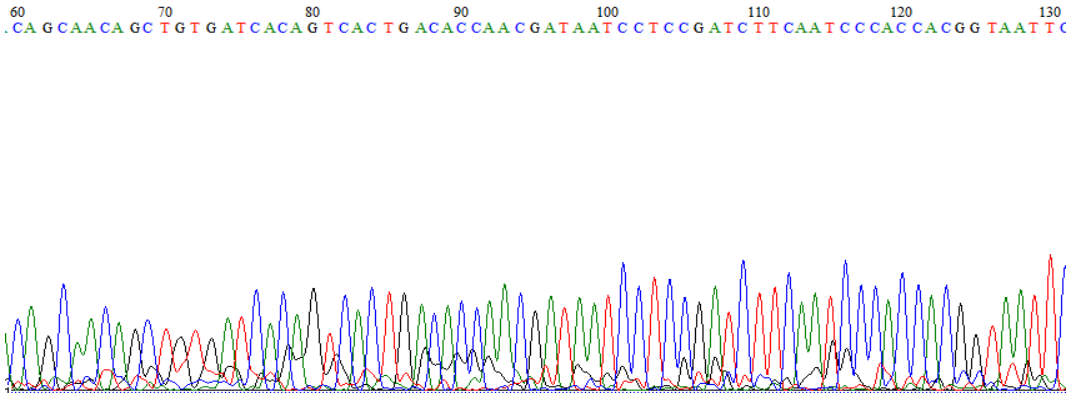
T3



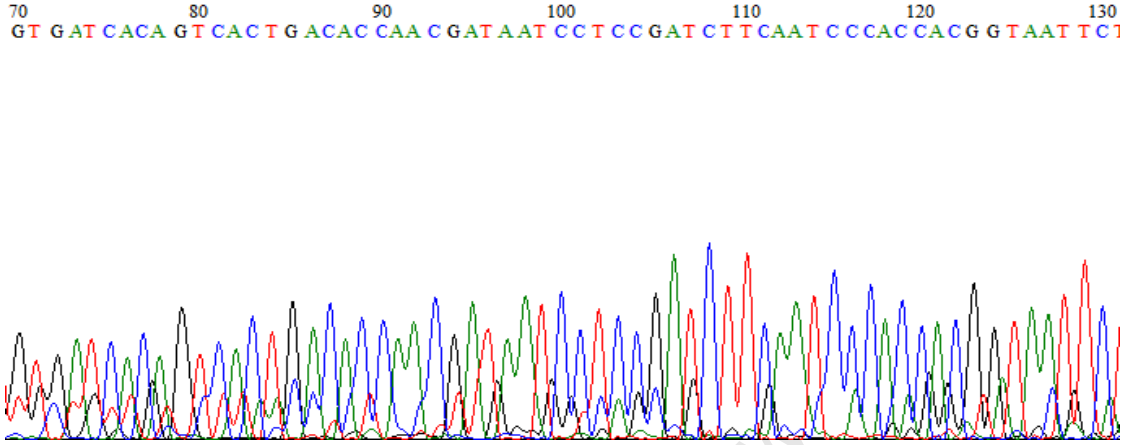
T7



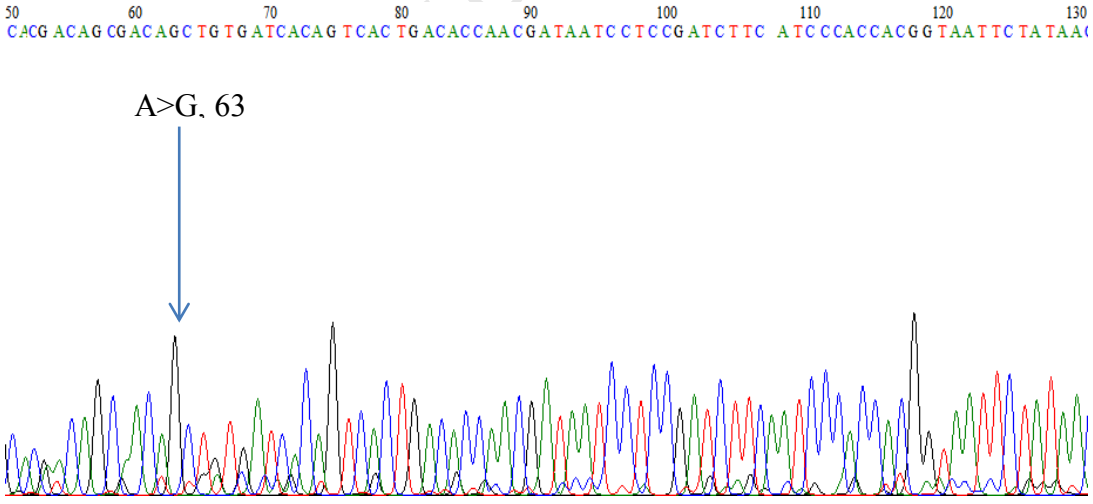
T17



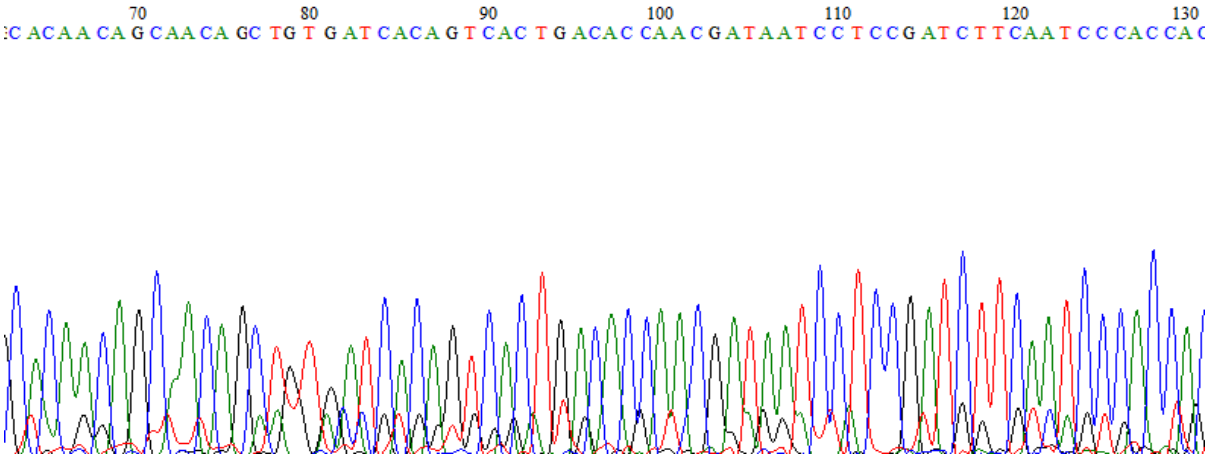
T30


70 80 90 100 110 120 130
GT GAT CACA GT CAC T G ACAC CAA C GAT AAT CC T CC GAT C T T CAAT C CCAC CAC G G T AAT T C T

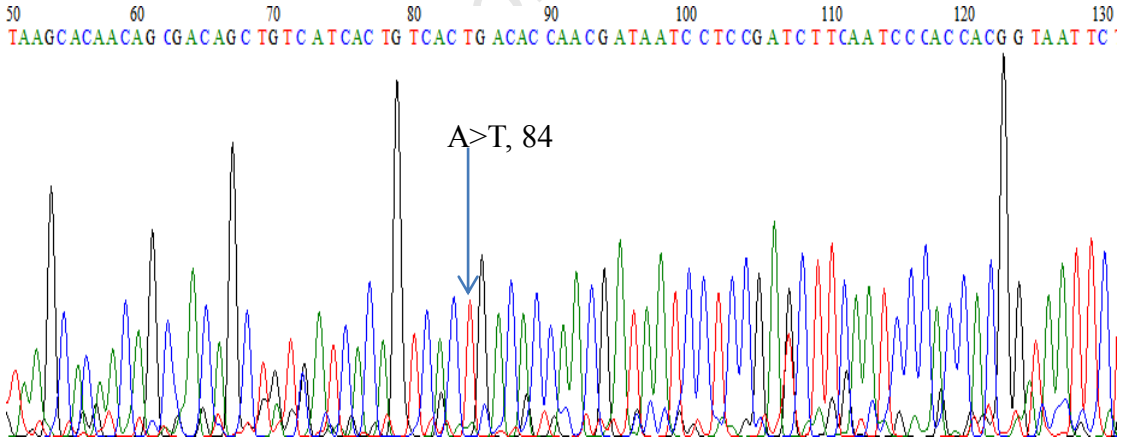
T68


50 60 70 80 90 100 110 120 130
CACG ACAG CG AC AG C T G T G AT CACAG T CAC T G ACACCAA C G AT AAT C C T CC G AT C T T C ATCCCACCACGG T AAT T C T AT AA
A>G, 63

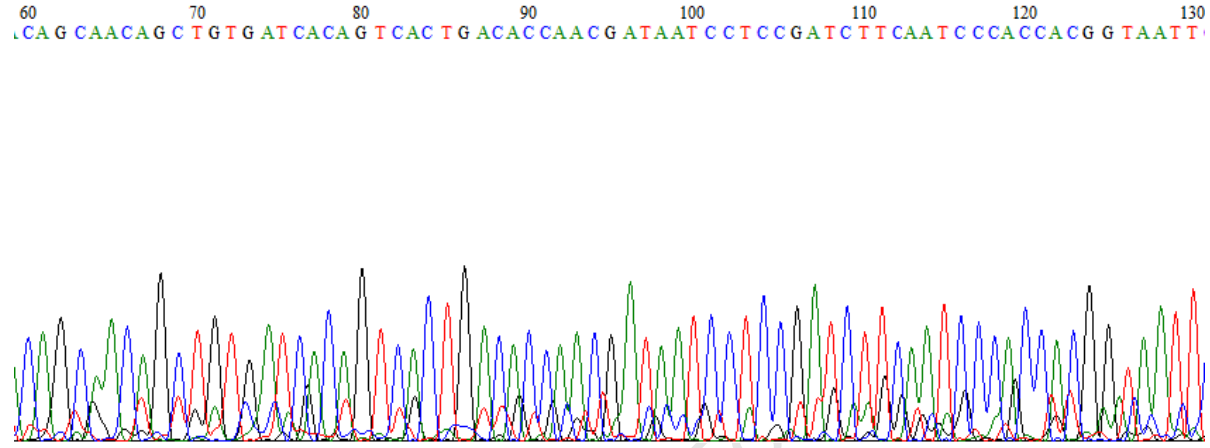
T72



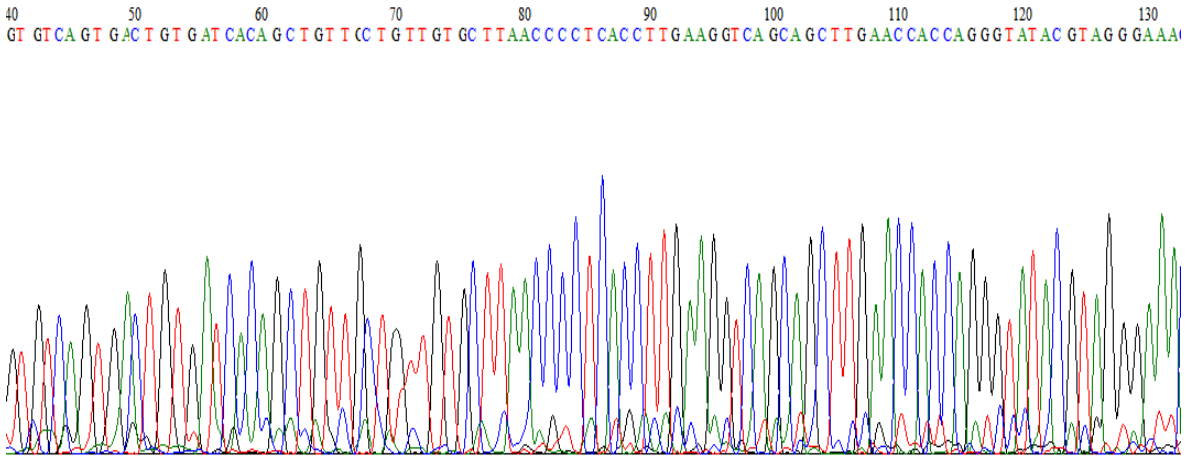
T79



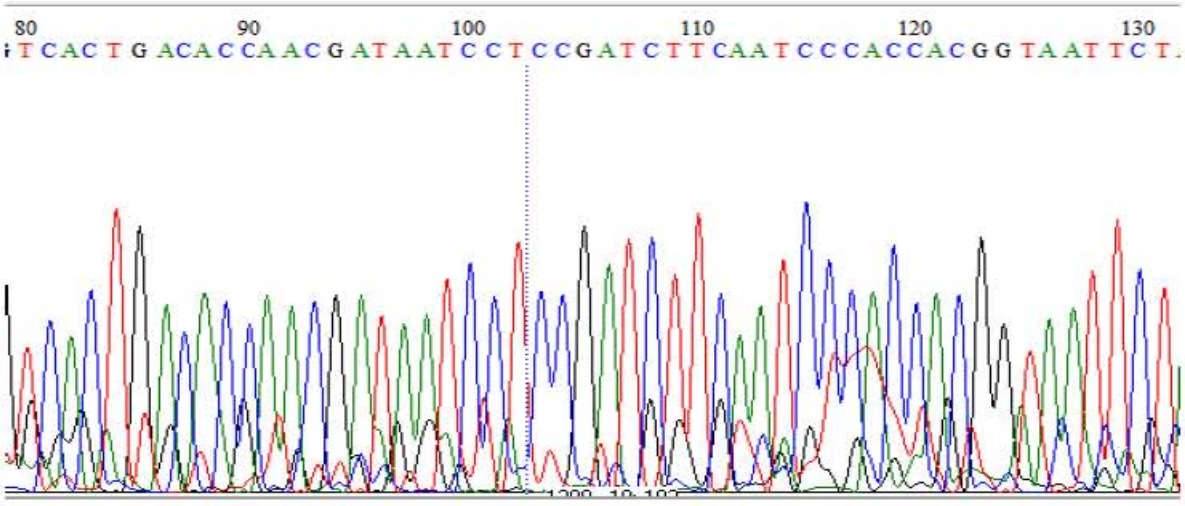
T82



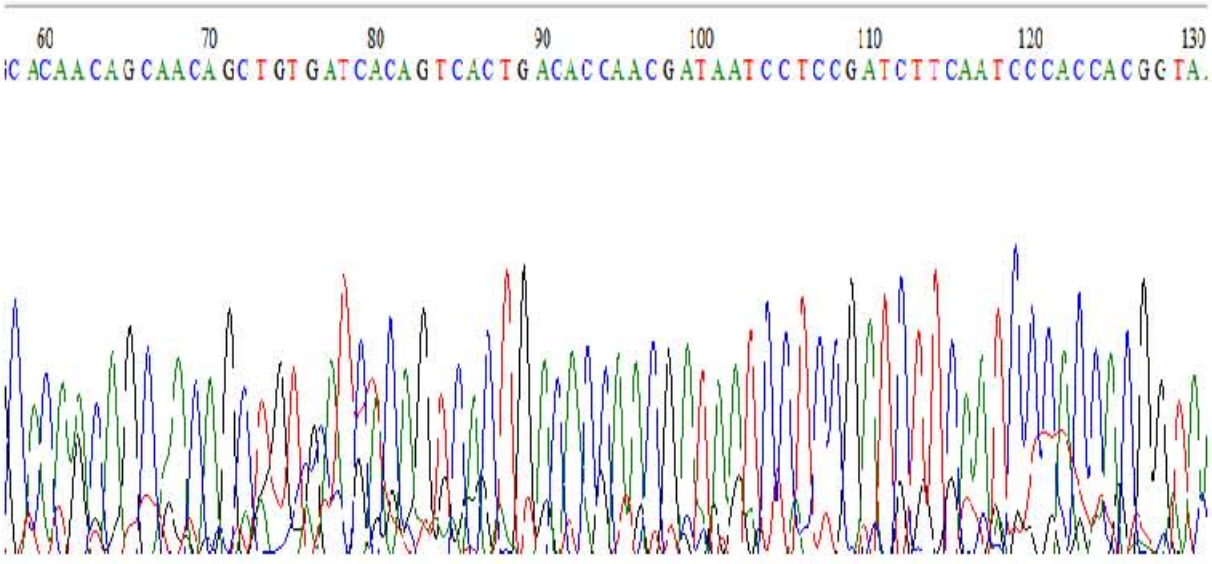
T91



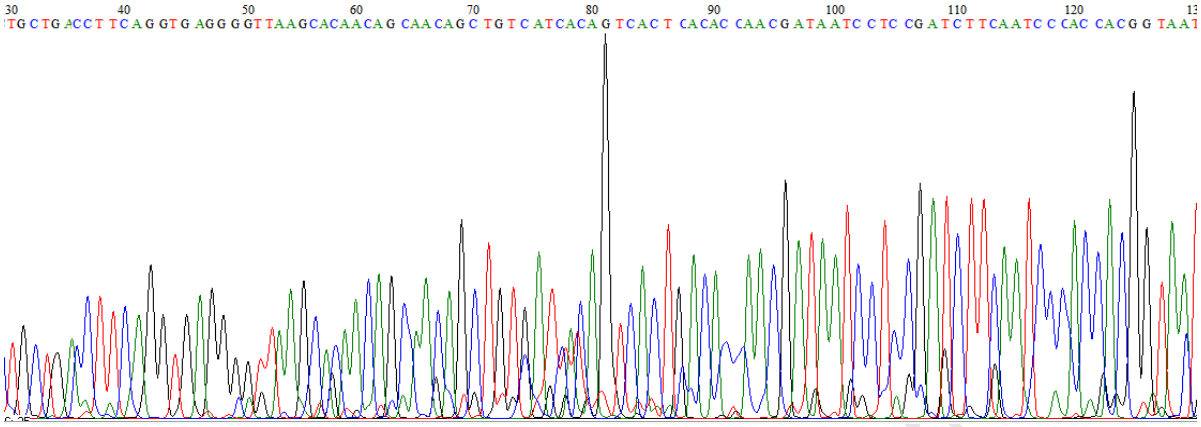
T92


80
90
100
110
120
130
TCACTGACACCAACGATAATCCTCCGATCTTCAATCCCACCACGGTAATTCT.

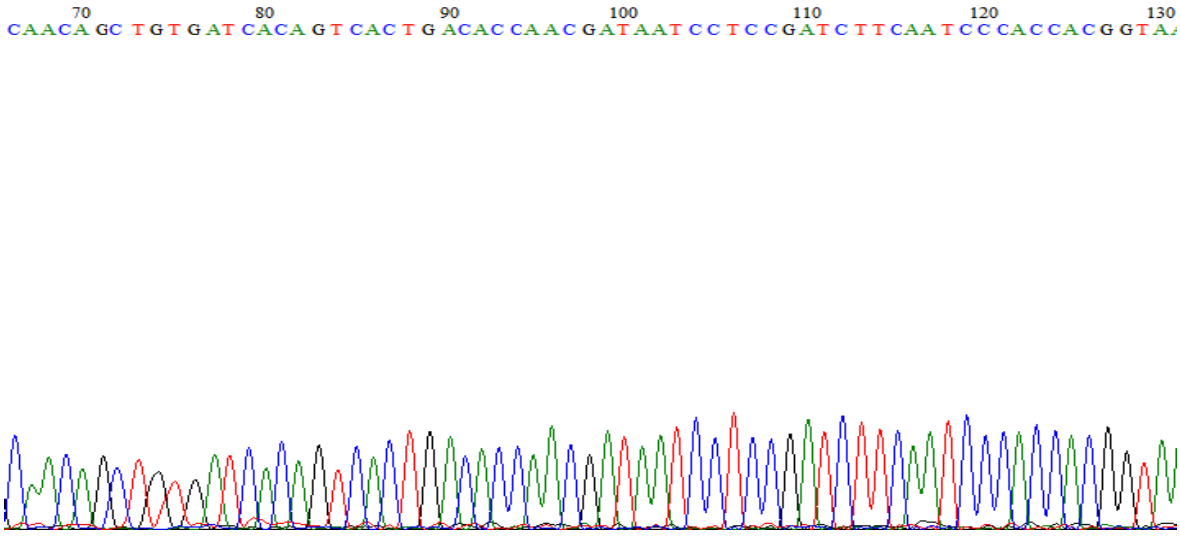
T95


60
70
80
90
100
110
120
130
CACAACAGCAACAGCTGTGATCACAGTCACTGACACCAACGATAATCCTCCGATCTTCAATCCCACCACGGTA.

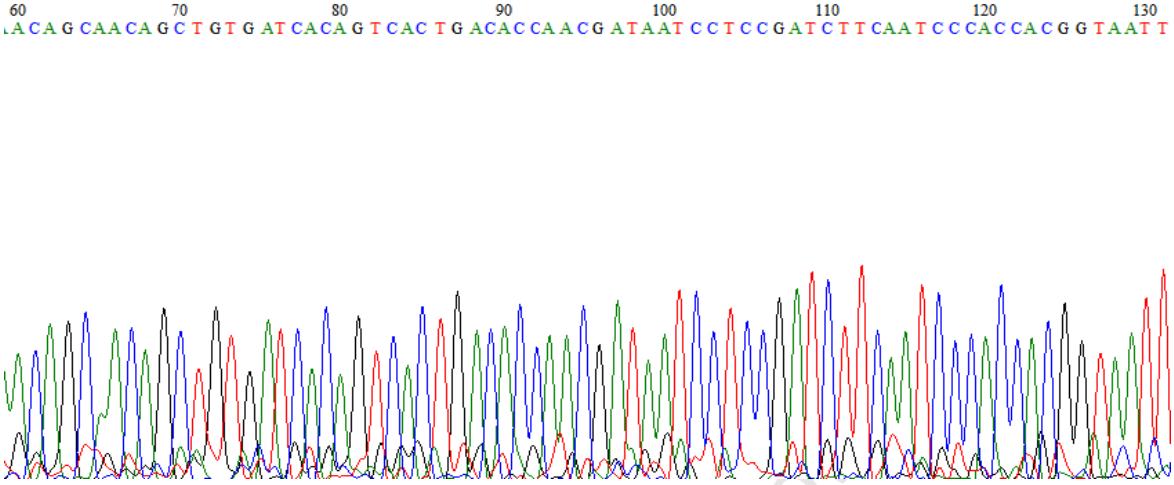
T16



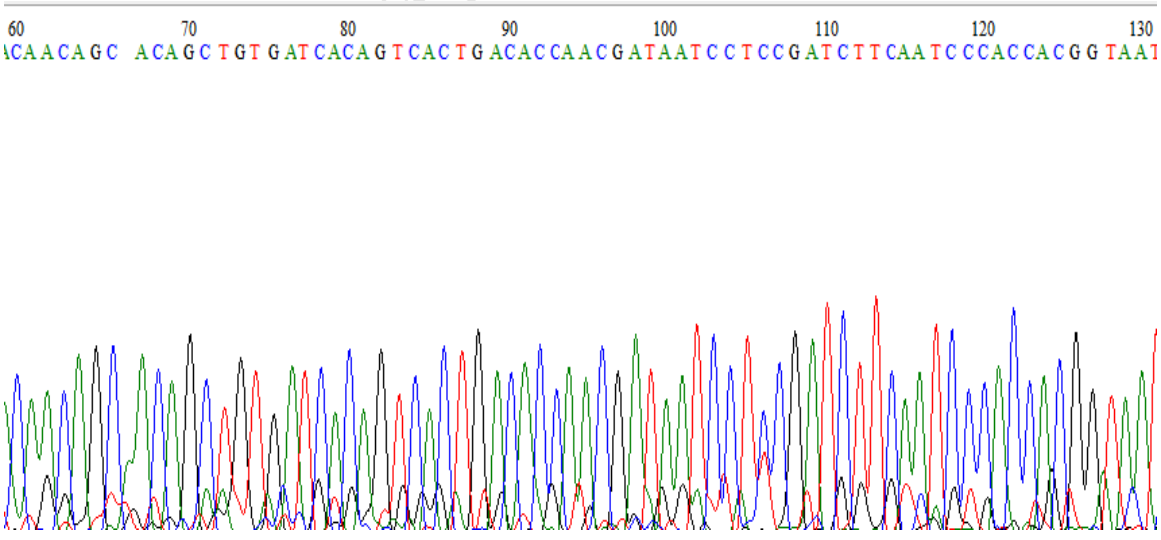
T81



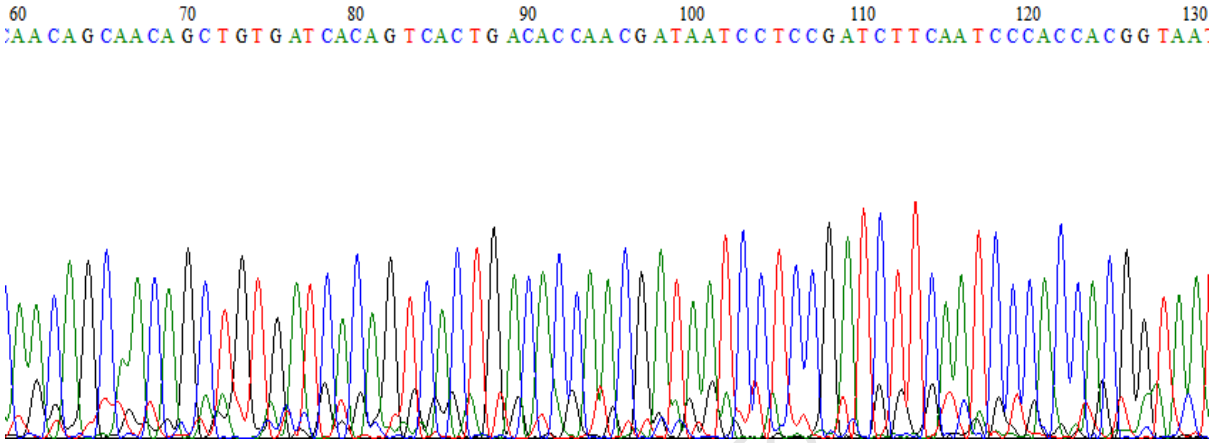
T11



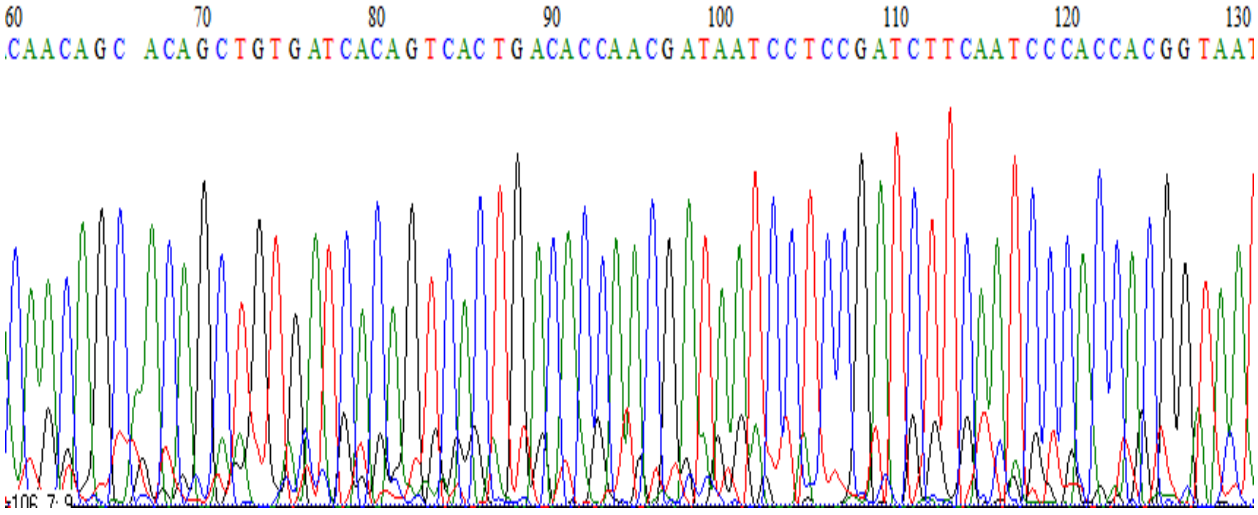
T12



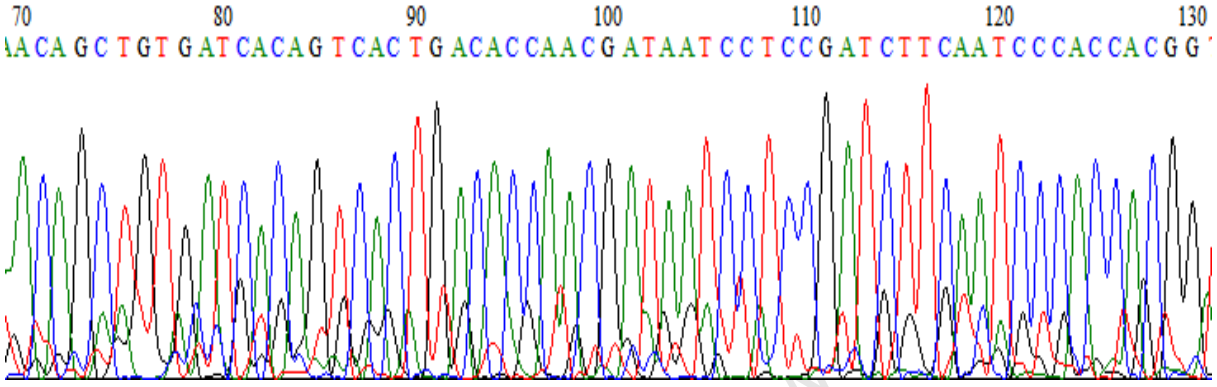
T13



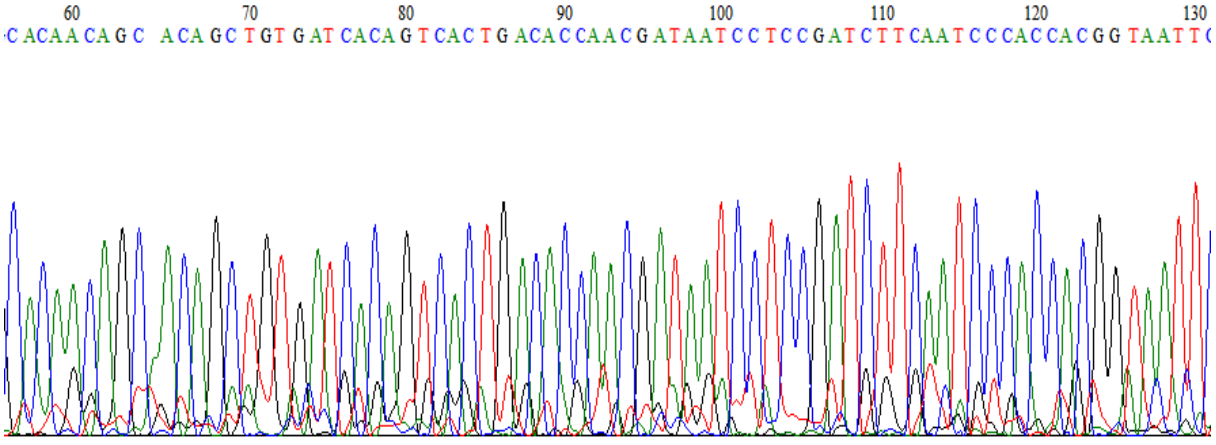
T18



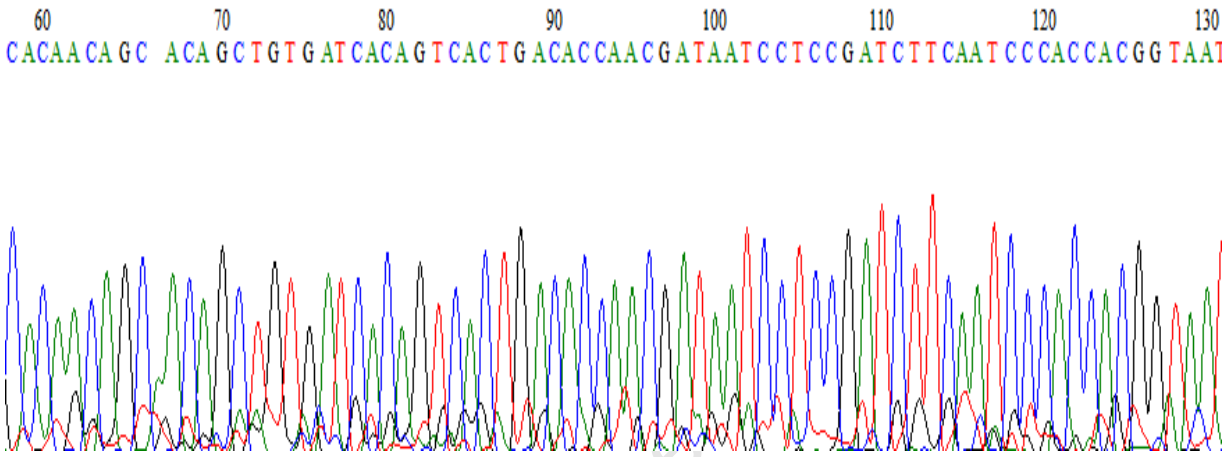
T31



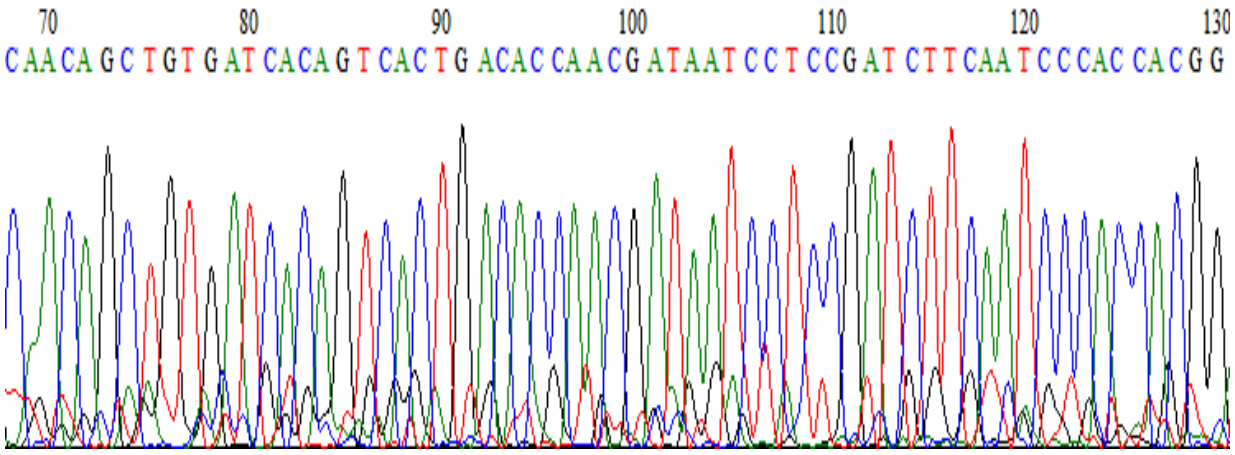
T32



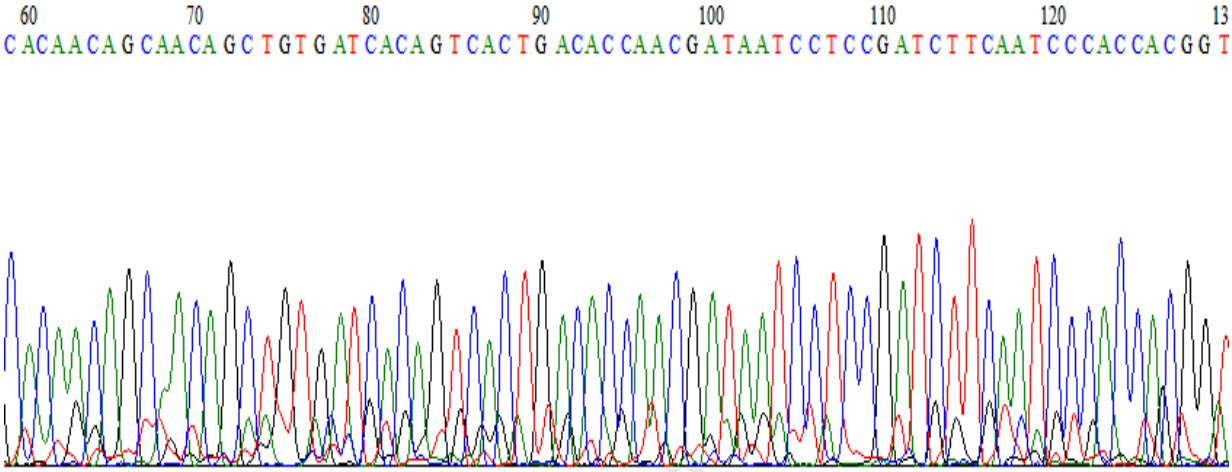
T35



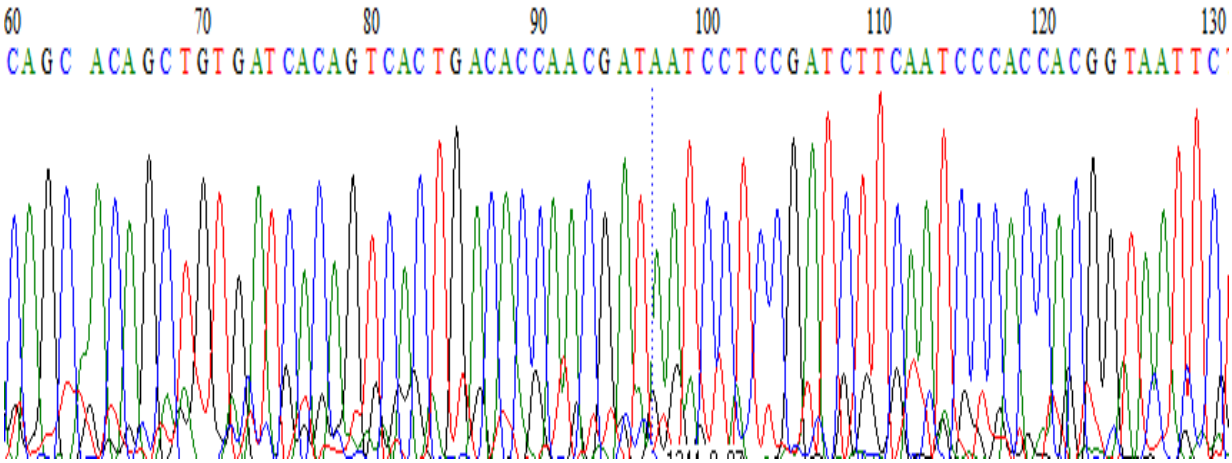
T38



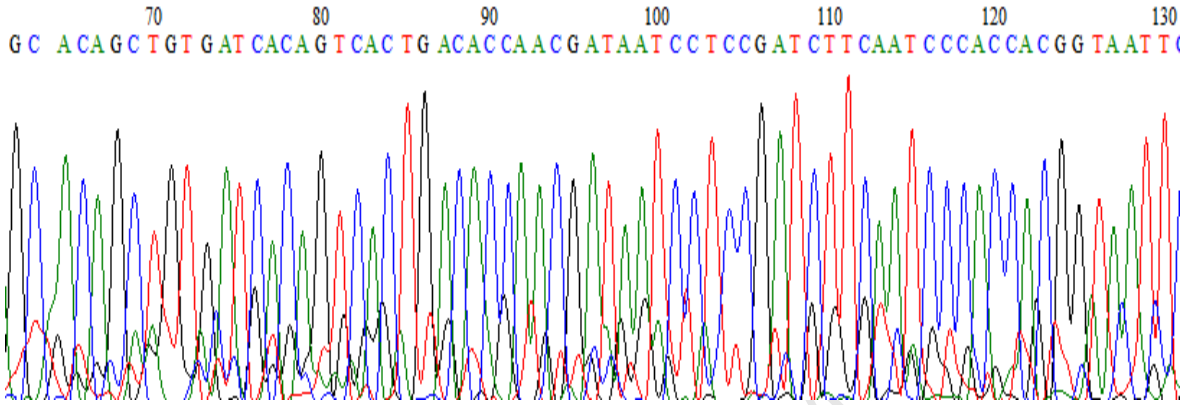
T39



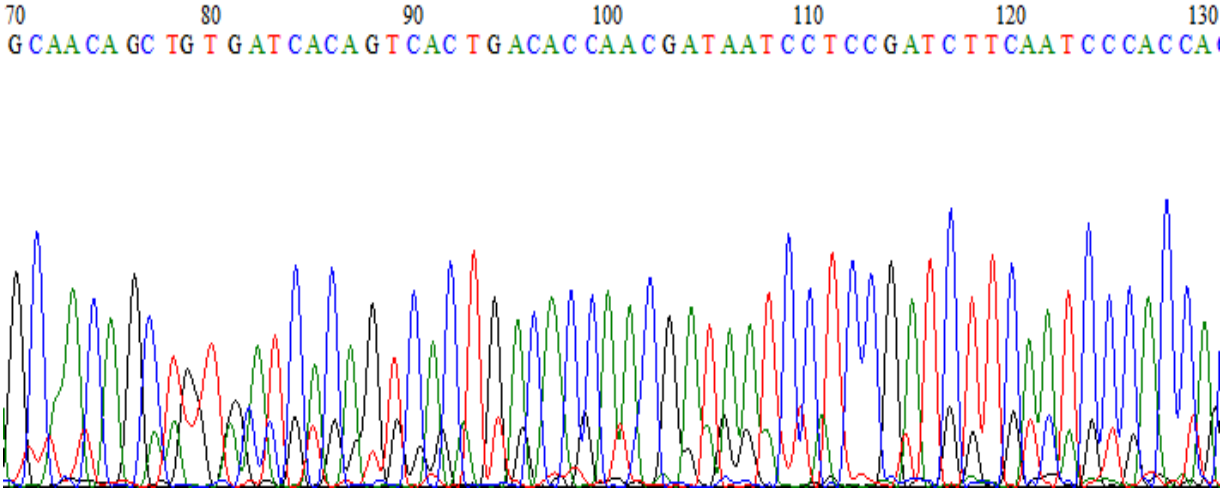
T42



T67



T70



Appendix B: Preparation of reagents for Agarose Gel Electrophoresis

A) Tris-borate/EDTA (TBE) (10X)

- Tris base (54 g), boric acid (27.5 g) and EDTA (4.65 g) was dissolved in 400ml of sterile water and then brought up to 500ml.
- The solution was dissolved and stored at room temperature (approximately 25 C).

B) Tris-borate/EDTA (TBE) (1X)

- TBE (1X) was prepared using (10X) TBE in a final volume of 100ml
- 10ml of 10x TBE was mixed with 90ml of sterile water.

C) Agarose gel (1.5%)

- Agarose (1.8 g) was dissolved in 120 ml of 1X TBE.
- The solution was heated up in a microwave until agarose granules completely dissolved.
- The gel was allowed to cool at room temperature, and then 10µl of ethidium bromide was added.
- The mixture was poured into a casting tray containing a comb that was placed in one end of the tray.
- The gel was allowed to set for 45 minutes at room temperature prior to use.

Appendix C: The demographic/clinicopathologic data of the 102 cases

Patient ID	Gender	Age	*H.pylori*	Node	Grade	Stage	Type
1	M	63	No	No	2	IB	Intestinal
2	F	77	No	No	2	0	Intestinal
3	M	78	No	NR	4	IB	Intestinal
4	M	38	Yes	No	3	IIA	Diffuse
5	F	58	Yes	Yes	4	IIA	Diffuse
6	M	41	No	Yes	2	II	Intestinal
7	M	54	Yes	NR	1	IA	Intestinal
8	F	18	No	NR	3	IIIA	Mixed
9	M	53	No	Yes	3	IIIB	Diffuse
10	M	40	No	No	4	IA	Diffuse
11	F	38	No	Yes	NR	IIIA	Diffuse
12	F	77	No	NR	3	IV	Intestinal
13	M	60	No	No	2	NR	Intestinal
14	M	54	Yes	Yes	3	II	Intestinal
15	F	51	Yes	No	2	IIIA	Intestinal
16	M	84	Yes	No	2	NR	Intestinal
17	F	74	No	Yes	4	IIIA	Diffuse
18	M	60	No	No	2	IB	Intestinal
19	F	59	Yes	Yes	4	II	Intestinal
20	M	59	No	NR	2	IB	Intestinal
21	M	69	No	No	2	IA	Intestinal
22	F	50	Yes	Yes	3	II	Intestinal
23	M	44	No	Yes	4	IIIB	Diffuse
24	F	64	Yes	Yes	3	IIIA	NR
25	M	53	Yes	Yes	2	IIIB	Intestinal
26	F	43	No	Yes	2	NR	Mixed
27	M	78	No	Yes	2	II	Intestinal
28	M	36	No	No	3	NR	Diffuse
29	M	79	No	No	3	II	NR
30	M	48	No	Yes	NR	IIIB	Diffuse
31	F	70	No	Yes	2	IIIB	Intestinal
32	F	43	No	No	4	IB	Diffuse
33	M	46	No	NR	3	II	Mixed

Patient ID	Gender	Age	*H.pylori*	Node	Grade	Stage	Type
35	M	52	No	Yes	2	IIIB	Intestinal
36	M	49	No	Yes	1	IV	Intestinal
37	F	41	No	Yes	3	NR	Intestinal
38	F	NR	No	Yes	2	IIIA	Intestinal
39	M	60	No	NR	3	IB	Intestinal
41	F	72	No	Yes	3	IIIA	Intestinal
42	M	46	No	Yes	4	IIIA	Diffuse
43	F	54	No	Yes	3	IIIA	Diffuse
44	F	72	No	Yes	2	1b	Intestinal
45	M	81	No	NR	2	1b	Intestinal
46	F	NR	No	NR	2	II	Intestinal
47	M	82	No	Yes	2	II	Intestinal
48	M	36	No	NR	NR	IV	Diffuse
49	F	46	No	Yes	NR	IIIA	Mixed
50	M	66	Yes	Yes	3	II	Intestinal
51	M	43	No	Yes	2	II	Intestinal
52	M	72	No	Yes	2	IV	Intestinal
53	M	44	No	Yes	4	IIIB	Diffuse
54	F	64	Yes	NR	2	IIIA	Intestinal
55	M	69	No	Yes	2	II	Intestinal
56	F	73	No	Yes	2	II	Mucinous
57	M	52	No	Yes	2	IB	Intestinal
58	M	57	Yes	No	3	NR	Mixed
59	M	61	No	NR	3	IA	Diffuse
60	F	76	No	Yes	2	IB	Intestinal
61	F	73	No	Yes	4	II	Diffuse
62	M	57	No	No	3	II	Intestinal
63	F	54	No	Yes	NR	IA	Diffuse
64	F	58	Yes	Yes	2	II	Intestinal
65	F	53	No	Yes	NR	IIIA	Diffuse
66	M	42	No	NR	2	IIIA	Mucinous
67	F	52	No	No	4	II	Diffuse
68	F	36	No	NR	4	III	Diffuse
69	F	33	No	No	3	IB	Diffuse
70	F	78	Yes	NR	4	II	Diffuse
71	M	60	No	No	3	IB	Diffuse
72	F	37	Yes	No	3	0	Diffuse
73	F	62	No	Yes	4	II	Diffuse
74	M	63	No	NR	4	IIIA	Diffuse

Patient ID	Gender	Age	*H.pylori*	Node	Grade	Stage	Type
76	F	41	No	Yes	4	II	Diffuse
77	F	62	Yes	Yes	4	IIIA	Diffuse
78	F	73	No	Yes	4	IIIB	Diffuse
79	F	64	Yes	Yes	4	IV	Diffuse
81	M	64	No	Yes	3	IIIA	Diffuse
82	M	32	No	No	4	III	Diffuse
83	F	60	Yes	NR	4	II	Diffuse
84	M	55	No	Yes	4	IIIB	Diffuse
85	F	65	No	Yes	4	NR	Diffuse
86	F	44	No	NR	4	II	Diffuse
87	M	44	No	No	3	IA	Diffuse
88	F	75	Yes	Yes	NR	IIIA	Diffuse
89	F	69	Yes	Yes	3	NR	Diffuse
90	NR	NR	NR	NR	3	NR	NR
91	F	61	No	Yes	3	NR	Diffuse
92	M	60	No	No	3	IIA	Diffuse
93	F	67	No	No	4	IIIA	Diffuse
94	M	40	No	No	3	NR	Diffuse
95	F	42	Yes	No	3	IB	Diffuse
96	M	57	No	No	NR	NR	Diffuse
97	M	44	No	Yes	4	IV	Diffuse
98	M	82	Yes	No	4	0	Diffuse
99	NR	NR	NR	NR	NR	NR	NR
100	M	56	Yes	Yes	3	NR	Diffuse
101	F	58	No	Yes	3	IIIC	Diffuse
102	F	54	No	Yes	3	IIIA	Diffuse

Appendix D: Staging

Stomach

Primary Tumour (T)

TX Primary tumour cannot be assessed

TO No evidence of primary tumour

Tis Carcinoma in situ: intraepithelial tumour without invasion of the lamina propria

T1 Tumour invades lamina propria or submucosa

T2 Tumour invades muscularis propria or subseroa

T2a Tumour invades muscularis propria

T2b Tumour invades subserosa

T3 Tumour invades serosa (visceral peritoneum) without invasion of adjacent structures

T4 Tumour invades adjacent structures

**Note:* A tumour may penetrate the muscularis propria with extension into the gastrocolic or gastrohepatic ligaments, or into the greater or lesser omentum, without perforation of the visceral peritoneum covering these structures. In this case, the tumour is classified T2. If there is perforation of the visceral peritoneum covering the gastric ligaments or the omentum, the tumour should be called T3.

***Note:* The adjacent structures of the stomach include the spleen, transverse colon, liver, diaphragm, pancreas, abdominal wall, adrenal gland, kidney, small intestine and retroperitoneum.

****Note:* Intramural extension to the duodenum or esophagus is classified by the depth of the greatest invasion in any of these sites, including the stomach.

Regional lymph nodes (N)

NX Regional lymph node(s) cannot be assessed

NO No regional lymph node metastasis*

N1 Metastasis in 1 to 6 regional lymph nodes

N2 Metastasis in 7 to 15 regional lymph nodes

N3 Metastasis in more than 15 regional lymph nodes

**Note:* A designation of pNO should be used if all examined lymph nodes are negative, regardless of the total number removed and examined.

Distant metastasis (M)

MX Distant metastasis cannot be assessed

MO No distant metastasis

M1 Distant metastasis

Stage grouping

Stage			
Stage 0	pTis	N0	M0
Stage IA	pT1	N0	M0
Stage IB	pT1	N1	M0
	pT2a/b	N1	M0
Stage II	pT1	N2	M0
	pT2a/b	N1	M0
	pT3	N0	M0
Stage IIIA	pT2a/b	N2	M0
	pT3	N1	M0
	pT4	N0	M0
Stage IIIB	pT3	N2	M0
Stage IV	pT4	N1-3	M0
	pT1-3	N3	M0
	Any pT	Any N	M1