CHARACTERIZATION OF HEREDITARY CANCER SYNDROMES

by

Kasmintan Alexandra Schrader

M.B.B.S., The University of British Columbia, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Pathology and Laboratory Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2013

Abstract

Hereditary cancer syndromes predispose to early-onset or multiple cancers in a person or family, follow Mendelian inheritance patterns and demonstrate stereotyped patterns of tumor development. Genotype-phenotype correlations direct clinical genetic testing and provide guidance for hereditary cancer management. This thesis began by examining the association of lobular breast cancer with germline mutations in *CDH1*, the gene encoding the epithelial cell-cell adhesion molecule E-cadherin and tested whether *CDH1* represented a high-frequency breast cancer susceptibility gene, apart from its association with hereditary diffuse gastric cancer. In addition it examined other genotype-phenotype correlations including the association between granular cell tumors and a multiple congenital anomaly syndrome, the specific correlation between a recurrent somatic mutation in a transcription factor and adult-type granulosa cell tumors and the strong association of germline *BRCA1* and *BRCA2* mutations with high-grade serous epithelial ovarian cancer. These candidate gene analyses were performed using low-throughput molecular technologies. With the advent of cheaper DNA-sequencing capabilities, the application of these new technologies to novel Mendelian disease gene discovery and hereditary cancer management became the subsequent focus of the thesis. Objectives: To determine the frequency of germline *CDH1* mutations in women with lobular breast cancer unselected for familial gastric cancer; to define the associations between several alternative genotype-phenotype correlations; and, to apply next-generation sequencing to determine the basis of a Mendelian disorder, in order to determine its utility as a potential familial cancer gene discovery and clinical tool. Selected Methods: Single amplicon mutation screening and sequence analysis of a large cohort of women with early-onset or familial lobular breast cancer. Next-generation sequencing analysis of a family with multiple individuals cosegregating spondyloepiphyseal dysplasia and retinitis pigmentosa. Results: Without a selective history of diffuse gastric cancer, potentially pathogenic germline mutations in *CDH1* occured in women with early-onset or hereditary lobular breast cancer in less than two percent of individuals. Diagnosis of a Mucolipidosis type III gamma was possible using new sequencing technologies. Conclusion: There is utility in understanding genotype-phenotype correlations in order to direct genetic testing and novel gene identification. Next-generation sequencing technologies can succesfully be applied to Mendelian disorders with clear phenotypes for gene discovery.

Preface

The introductory chapter 1 integrates excerpts from the book chapter Hereditary diffuse gastric cancer, written by me, with planning and critical editing by Dr David Huntsman.

K. Schrader, and D. Huntsman, 'Hereditary Diffuse Gastric Cancer', **Cancer Treat Res**, 155 (2010), 33-63.

The second chapter integrates the article authored by me: the experimental work I did is listed below. S Masciari and I contributed equally to the work.

K. A. Schrader*, S. Masciari*, N. Boyd, C. Salamanca, J. Senz, D. N. Saunders, E. Yorida, S. Maines-Bandiera, P. Kaurah, N. Tung, M. E. Robson, P. D. Ryan, O. I. Olopade, S. M. Domchek, J. Ford, C. Isaacs, P. Brown, J. Balmana, A. R. Razzak, P. Miron, K. Coffey, M. B. Terry, E. M. John, I. L. Andrulis, J. A. Knight, F. P. O'Malley, M. Daly, P. Bender, kConFab, R. Moore, M. C. Southey, J. L. Hopper, J. E. Garber, and D. G. Huntsman, 'Germline Mutations in CDH1 Are Infrequent in Women with Early-Onset or Familial Lobular Breast Cancers', **J Med Genet**, 48 (2011), 64-8.*equal contributions. I wrote the manuscript and coordinated the edits with the co-authors. I performed 85% of the experimental work. I performed mutation screening and sequencing using DHPLC and Sanger sequencing on half of the samples. The remainder of samples were sequenced by the BC Cancer Agency Genome Sciences Centre. I analysed all of the sequence data. I performed site-directed mutatgenesis of the E-cadherin WT construct for each of the missense variants studied. I transiently transfected the cells and performed the immunohistochemistry. Collection of the samples was by S Masciari and coordination of samples from the Breast CFR by N Boyd. Analysis of the results was undertaken jointly between me and S Masciari, JE Garber and DG Huntsman. DG Huntsman and JE Garber conceived the study.

Approval to study the association between lobular breast cancer and germline mutations in E-cadherin was sought through the University of British Columbia research ethics board. Research ethics approval number: (H05-60120) R05-0120.

The first chapter also integrates excerpts from the review article Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer, for which I wrote the first draft, and S Masciari and J Garber contributed the section on management. SM, JG, and DH also critically edited the manuscript.

K. A. Schrader, S. Masciari, N. Boyd, S. Wiyrick, P. Kaurah, J. Senz, W. Burke, H. T. Lynch, J. E. Garber, and D. G. Huntsman, 'Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer', **Fam Cancer**, 7 (2008), 73-82.

The third chapter integrates three articles written by me and the work I did in each manuscript is listed below. B Gorbatcheva and I contributed equally to the work on the first manuscript.

K. A. Schrader*, B. Gorbatcheva*, J. Senz, A. Heravi-Moussavi, N. Melnyk, C. Salamanca, S. Maines-Bandiera, S. L. Cooke, P. Leung, J. D. Brenton, C. B. Gilks, J. Monahan, and D. G. Huntsman, 'The Specificity of the *FOXL2* C.402c>G Somatic Mutation: A Survey of Solid Tumors', **PLoS One**, 4 (2009), e7988.*equal contributions. I wrote the publication and coordinated the edits with the co-authors. (I contributed 50% to the team effort) Experimental work performed by B Gorbatcheva, J Senz and N Melnyk. Analysis of results was joint between B Gorbatcheva, J Monahan, DG Huntsman and me. DG Huntsman and J Monahan conceived the study.

K. A. Schrader, T. N. Nelson, A. De Luca, D. G. Huntsman, and B. C. McGillivray, 'Multiple Granular Cell Tumors Are an Associated Feature of Leopard Syndrome Caused by Mutation in *PTPN11*', **Clin Genet**, 75 (2009), 185-9. I wrote the publication and performed the experimental work. I performed the germline mutation testing of *PTPN11* and identified the p.T468M *PTPN11* mutation in the patient. I extracted DNA from paraffin tumor blocks from the patient's multiple granular cell tumors. I performed Sanger sequencing of the paraffin DNA to look for LOH. I performed flourecense *in situ* hybridization (FISH) on tumor sections for to examine for a large deletion of the *PTEN* locus. (I contributed 85% to the team effort). Analysis of results was jointly between the coauthors. BC McGillivray and I conceived the study. In a follow-up analysis, I extracted DNA from another 10 parrafin

granular cell tumor blocks and genotyped them for the recurrent p.T468M *PTPN11* variant associated with LEOPARD syndrome.

K. A. Schrader, J. Hurlburt, S. E. Kalloger, S. Hansford, S. Young, D. G. Huntsman, C. B. Gilks, and J. N. McAlpine, 'Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy', **Obstet Gynecol**, 120 (2012), 235-40. I wrote the publication and performed the analytic work. I reviewed the genetic counselor-ascertained family histories, the age of onset of ovarian cancer and the *BRCA1* and *BRCA2* mutation reports. I performed the statistical analysis using tests advised by the statistical editor of Obstetrics and Gynecology. (I contributed 75% to the team effort) DGH, CBG, JM and I conceived the study and critically edited the manuscript.

Approval was obtained from the University of British Columbia Ethics Board to approach women from the Vancouver General Hospital and British Columbia Cancer Agency in Vancouver, British Columbia, Canada, who had been fully staged and were undergoing debulking surgery (primary or delayed) for cancers of ovarian/peritoneal/fallopian tube origin for informed consent for the banking of tumor tissue. Research ethics approval number: (H05-60119) R05-0119

The fourth chapter integrates an article authored by myself. A Heravi-Moussavi and I contributed equally to the work.

K. A. Schrader*, A. Heravi-Moussavi*, P. J. Waters, J. Senz, J. Whelan, G. Ha, P. Eydoux, T. Nielsen, B. Gallagher, A. Oloumi, N. Boyd, B. A. Fernandez, T. L. Young, S. J. Jones, M. Hirst, S. P. Shah, M. A. Marra, J. Green, and D. G. Huntsman, 'Using Next-Generation Sequencing for the Diagnosis of Rare Disorders: A Family with Retinitis Pigmentosa and Skeletal Abnormalities', **J Pathol**, 225 (2011), 12-8. *equal contributions. I wrote the publication and coordinated the edits with the co-authors. (I contributed 75% to the team effort). Analysis of results was joint between me and A Heravi-Moussavi. I created the study protocol for exome sequencing: H09-00971. I coordinated the project from the Vancouver site and reviewed the clinical data from the Newfoundland site. I directed the downstream analysis performed by A Heravi-Moussavi following variant calling. This included designing the strategy to filter out all known variants in public databases and also variants from within

in-house exomes to eliminate systematic platform-dependent artifacts. I advised the Mendelian filtering strategy and I reviewed all candidate genes with germline variants. I performed orthogonal validation of the candidate variant and performed segregation analysis in the family. I coordinated biological validation of the disease causing mutation through multi-institutional collaborations. DG Huntsman, J Green and I conceived the study.

Approval to study the genetic basis of this family's bone and eye disease was sought through the University of British Columbia and Memorial University Research Ethics Boards. Family members had been asked to participate in research to determine the genetic basis of the eye and bone disease. Informed written consent was obtained from all participants or their proxy prior to investigation, review of their medical records and collection of DNA. Research ethics approval number: H09-00971 and Memorial University approval number: HIC-06.15.

Table of Contents

List Of Tables

List Of Figures

List of Abbreviations

A-GCTs	adult-type granulosa-cell tumors
ADNP	activity-dependent neuroprotector homeobox
ALK	anaplastic lymphoma receptor tyrosine kinase
APC	adenomatosis polyposis coli
ARFGEF2	ADP-ribosylation factor guanine nucleotide-exchange factor 2
ATM	ataxia telangiectasia mutated
BAM files	a binary version of a tab-delimited text file that containing sequence alignment data
BBS	Bardet–Biedl syndrome
BC	breast cancer
BCC	basal-cell carcinoma
BCCA	British Columbia Cancer Agency
BCRF	Breast Cancer Research Foundation
bp	Base-pair
BPES	blepharophimosis–ptosis–epicanthus–inversus syndrome
BRAF	v-raf murine sarcoma viral oncogene homolog B1
BRCA1	breast cancer 1, early onset
BRCA2	breast cancer 2, early onset
Breast CFR	Breast Cancer Family Registry
BRIP1	BRCA1 interacting protein C-terminal helicase 1
BUB1B	BUB1 mitotic checkpoint serine/threonine kinase B
cagA	cytotoxin-associated gene A
CDH1	cadherin 1, type 1, E-cadherin
CDKN2A	cyclin-dependent kinase inhibitor 2A
CDKN2B	cyclin-dependent kinase inhibitor 2B
CFC	cardio-facio-cutaneous syndrome
cGMP	cyclic guanosine monophosphate
CHEK2	checkpoint kinase 2
COL2A1	collagen, type II, alpha 1
COL9A1	collagen, type IX, alpha 1
COL9A2	collagen, type IX, alpha 2
COMP	cartilage oligomeric matrix protein
COX2	cytochrome c oxidase subunit II
CS	Costello syndrome
CTAG	Centre for Translational and Applied Genomics
CTNNB1	catenin (cadherin-associated protein), beta 1
DAPI	4',6-diamidino-2-phenylindole
DDB2	damage-specific DNA binding protein 2
DDX27	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27
dEF1/ZEB-1	protein endcoded by zinc finger E-box binding homeobox 1
DGC	diffuse gastric cancer
DHODH	dihydroorotate dehydrogenase (quinone)

DHPLC	Denaturing high-performance liquid chromatography
DNA	Deoxyribonucleic acid
DPM1	dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EKG	electrocardiogram
EPIC-EURGAST	European Prospective Investigation into Cancer and Nutrition
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2
ERCC2	excision repair cross-complementing rodent repair deficiency, complementation group 2
ERCC3	excision repair cross-complementing rodent repair deficiency, complementation group 3
ERCC4	excision repair cross-complementing rodent repair deficiency, complementation group 4
ERCC5	excision repair cross-complementing rodent repair deficiency, complementation group 5
ERG	electroretinogram
Erk	Extra-cellular signal-regulated kinase 1
ESP	Exome Sequencing Project
FAP	familial adenomatous polyposis
FFPE	Formalin-Fixed, Paraffin-Embedded
FHIT	fragile histidine triad
FISH	fluorescence *in situ* hybridization
FOXL2	forkhead box L2
GAB1	GRB2-associated binding protein 1
GAPPS	Gastric adenocarcinoma and proximal polyposis of the stomach
GC	gastric cancer
GCT	granular cell tumors
GI	gastrointestinal tract
GNPTAB	N-acetylglucosamine-1-phosphate transferase, alpha and beta subunits
GNPTG	N-acetylglucosamine-1-phosphate transferase, gamma subunit
GSK3b	glycogen synthase kinase 3 beta
GTP	guanosine triphosphate
H. pylori	*Helicobacter pylori*
H&E	Hematoxylin-eosin stain
HBOC	Hereditary Breast and Ovarian Cancer
HDGC	hereditary diffuse gastric cancer
HER2	Human Epidermal Growth Factor Receptor 2
HGSOC	high-grade serous ovarian cancer
HRM	High Resolution Melt Curve Analysis
IBD	identity-by-descent
IDC	invasive ductal carcinoma
IFNgR1	interferon gamma receptor 1
IGC	intestinal gastric cancer
IGCLC	International Gastric Cancer Linkage Consortium
IL-10	interleukin 10

IL-1b	interleukin 1, beta
IL-1RN	interleukin 1 receptor antagonist
IOSE-80pc	immortalized ovarian surface epithelium
JMML	juvenile myelomonocytic leukemia
kConFab	The Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer
KGN	human ovarian granulosa-like tumour cell line
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LBC	lobular breast cancer
LCIS	Lobular carcinoma in situ
LEF1/ Tcf	lymphoid enhancer-binding factor 1
LFS	Li Fraumeni Syndrome
LOH	Loss of heterozygosity
LS	LEOPARD syndrome
MAP	Mitogen-activated protein kinase
MAQ	mapping assembly with quality
MATN3	matrilin 3
Mb	megabase
MCA	Multiple Congenital Anomalies
MED	multiple epiphyseal dysplasia
MET	met proto-oncogene
MGCT	multiple granular cell tumors
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2
MLIIIg	mucolipidosis type III gamma
MLIIIα/β	mucolipidosis type III alpha/beta
MLPA	Multiplex ligation-dependent probe amplification
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MSH2	mutS homolog 2, colon cancer, nonpolyposis type 1
MSH6	mutS homolog 6
MSI	Microsatellite instability
MTAP	methylthioadenosine phosphorylase
MUTYH	mutY homolog (E. coli)
MYH3	myosin, heavy chain 3, skeletal muscle, embryonic
NCBI	National Center for Biotechnology Information
NF1	neurofibromin 1
NGS	next-generation sequencing
NHLBI	National Heart, Lung, and Blood Institute
NS	Noonan syndrome
OMIM	Online Mendelian Inheritance in Man
OvCaRe	Ovarian Cancer Research
PALB2	partner and localizer of BRCA2
PARD6B	par-6 partitioning defective 6 homolog beta
PARP	polyADP-ribose polymerase

PAS-D	Periodic Acid-Schiff Reaction with diastase digestion
PBRM1	polybromo 1
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDGFRB	platelet-derived growth factor receptor, beta polypeptide
PET	Positron emission tomography
PFA	paraformaldehyde
PHTS	PTEN hamartoma tumor syndrome
PI3K	Phosphatidylinositol 3-kinase
PIPK1g	Ig phosphatidylinositol phosphate kinase
PJS	Peutz-Jeghers syndrome
PML	promyelocytic leukemia
PMS1	PMS1 postmeiotic segregation increased 1
PMS2	PMS2 postmeiotic segregation increased 2
POLH	polymerase (DNA directed), eta
PPI	proton pump inhibitors
PSCA	prostate stem cell antigen
PTCH	patched 1
PTEN	phosphatase and tensin homolog
PTG	Prophylactic total gastrectomy
PTPN11	protein tyrosine phosphatase, non-receptor type 11
RAF1	v-raf-1 murine leukemia viral oncogene homolog 1
RAF1	v-raf-1 murine leukemia viral oncogene homolog 1
RARA	retinoic acid receptor, alpha
RB1	retinoblastoma 1
RET	ret proto-oncogene
Rho	rhodopsin
RP	retinitis pigmentosa
RPL3L	ribosomal protein L3-like
RT-PCR	Reverse transcription polymerase chain reaction
RUNX3	runt-related transcription factor 3
SAM	Sequence Alignment/Map
SED	spondyloepiphyseal dysplasia
SEGA	subependymal giant cell astrocytoma
SIFT	Sorting Tolerant From Intolerant
Sip-1/ZEB-2	protein endcoded by zinc finger E-box binding homeobox 2
SLC26A2	solute carrier family 26 (sulfate transporter), member 2
SMAD4	SMAD family member 4
SMO	smoothened, frizzled family receptor
SNP	single-nucleotide polymorphism
SNV	single nucleotide variant
SRC	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
SSCP	Single-strand conformation polymorphism

STK11	serine/threonine kinase 11
SV40	Simian virus 40
SYK	spleen tyrosine kinase
E12/E47	protein encoded by transcription factor 3 (TCF3)
TCGA	The Cancer Genome Atlas
TGFbRII	transforming growth factor, beta receptor II
TMEFF2	transmembrane protein with EGF-like and two follistatin-like domains 2
TMEM189-UBE2V	TMEM189-UBE2V1 readthrough
TNF-a	tumor necrosis factor
TP53	tumor protein p53
TRAPPC2	trafficking protein particle complex 2
Twist	protein endcoded by *TWIST1* twist basic helix-loop-helix transcription factor 1
UCSC	University of California, Santa Cruz
UTR	untranslated region
WEC	whole exome capture
WISP3	WNT1 inducible signaling pathway protein 3
WRN	Werner syndrome, RecQ helicase-like
WT	wild-type
XP	Xeroderma pigmentosum
ZFAS1	ZNFX1 antisense RNA 1

Dedication
To my family

Chapter 1: Introduction

1.1 Distinct Tumor Phenotypes In Hereditary Cancer Syndromes

Hereditary cancer syndromes are a group of Mendelian conditions that can predispose the individual to early-onset or multiple cancers. The tissue-specific predilections in many of the hereditary cancer syndromes are much like those seen in multiple congenital anomaly (MCA) syndromes. Many genes involved in cell growth, differentiation, and proliferation are found mutated in sporadic cancers, and these same genes can also be mutated in various MCA syndromes. Therefore, it is not unexpected that there are MCA syndromes that have abnormal cell proliferation, be it benign or malignant, as an associated feature. In attempting to dissect the underlying cause of the cancer susceptibility it is the underlying tumor pathology that provides major insight into the biologic mechanisms that have been compromised to allow the transformation to neoplasia. An example of germline mutations giving rise to a very specific tumor type is that of *CDH1*, the only known gene to be associated with hereditary diffuse gastric cancer. This particular example is poignant as it exemplifies a syndrome, whereby, genetic testing is currently based upon the histological diagnosis in the affected individual's cancer.

1.1.1 Hereditary Diffuse Gastric Cancer

Gastric cancer is one of the world's leading causes of cancer mortality where a small percentage can be attributed to heritable mutations in highly penetrant cancer susceptibility genes. Until 10 years ago, individuals from affected families lived with the uncertainty of developing lethal gastric cancer. Today, Hereditary Diffuse Gastric Cancer (HDGC) families can be identified, tested for causative mutations in *CDH1*, and for those families where a pathogenic mutation can be identified, prophylactic total gastrectomy (PTG) can be implemented in asymptomatic mutation carriers who elect to virtually eliminate their risk of developing this lethal disease.

1.1.1.1 Gastric Cancer Pathology, Epidemiology, And Molecular Genetics

1.1.1.1.1   Pathological Classification Of Gastric Cancer

Adenocarcinomas comprise the vast majority of primary gastric cancer (GC). Multiple histological classification systems for adenocarcinomas have been developed to better predict their prognosis, however, for the purpose of defining genetic risk, the most useful system is the classification of Lauren [1]. This system classifies the majority of adenocarcinomas into two main types; the intestinal-type and the diffuse-type, with the remainder forming an indeterminate category [1]. Tumors with components of both types are classified as mixed [1].

The more common, intestinal-type of gastric cancer (IGC) [2] is composed of glandular structures resembling intestinal epithelium. IGC arises from its precursor lesion, intestinal metaplasia [3], to form an exophytic tumor that ulcerates the stomach lining. Due to its localized presentation and distinctive appearance, IGC tends to be amenable to detection by endoscopic surveillance.

In contrast, DGC shows scattered, disorganized growth without distinctive architecture. Malignant cells infiltrate the wall of the stomach, gradually thickening it so that it takes on a leather bottle appearance, otherwise known as linitus plastica. The neoplastic cells have a distinctive signet ring appearance caused by an accumulation of intracellular mucin that pushes the nucleus to one side. This is demonstrated in Figure 1 where Figure 1a shows the signet ring appearance with a regular hematoxylin and eosin (H&E) stain. These cells can be confused with small blood vessels that have been sectioned transversely; however staining with PAS-D easily highlights the mucin-containing signet ring cells (Figure 1c). Unlike IGC, DGC has no defined premalignant lesion, although as noted, analysis of almost all reported PTG specimens, have demonstrated multiple microscopic foci of invasive DGC [4-17]. The DGC lesions associated with HDGC are usually very small and intramucosal, in situ or with pagetoid spread of signet ring cells [18]. These very small (<3mm) superficial clusters of invasive cancer appear to follow a more indolent course [19]. What causes these small invasive cancers to become clinically significant is not fully understood however the phenotype of the DGC cells that spread beyond the mucosa is that of poor differentiation, and activation of a known epithelial-mesenchymal transition inducer, Src kinase (SRC) [19].

Figure 1 These pictures show a small invasive focus of a DGC from a prophylactic gastrectomy specimen.

(a) H&E stain; (b) E-cadherin stain showing down-regulated expression in the invasive signet ring cells in comparison to the normal E-cadherin-positive epithelium: (c) PAS-D stain for mucin showing the presence of intracellular mucin in the cytoplasm of signet ring cells. Photographs taken by Dr Martin Köbel. Reproduced from [Hereditary Diffuse Gastric Cancer'., Schrader KA, Huntsman D, 155, 33-63, 2010] with permission from Springer.
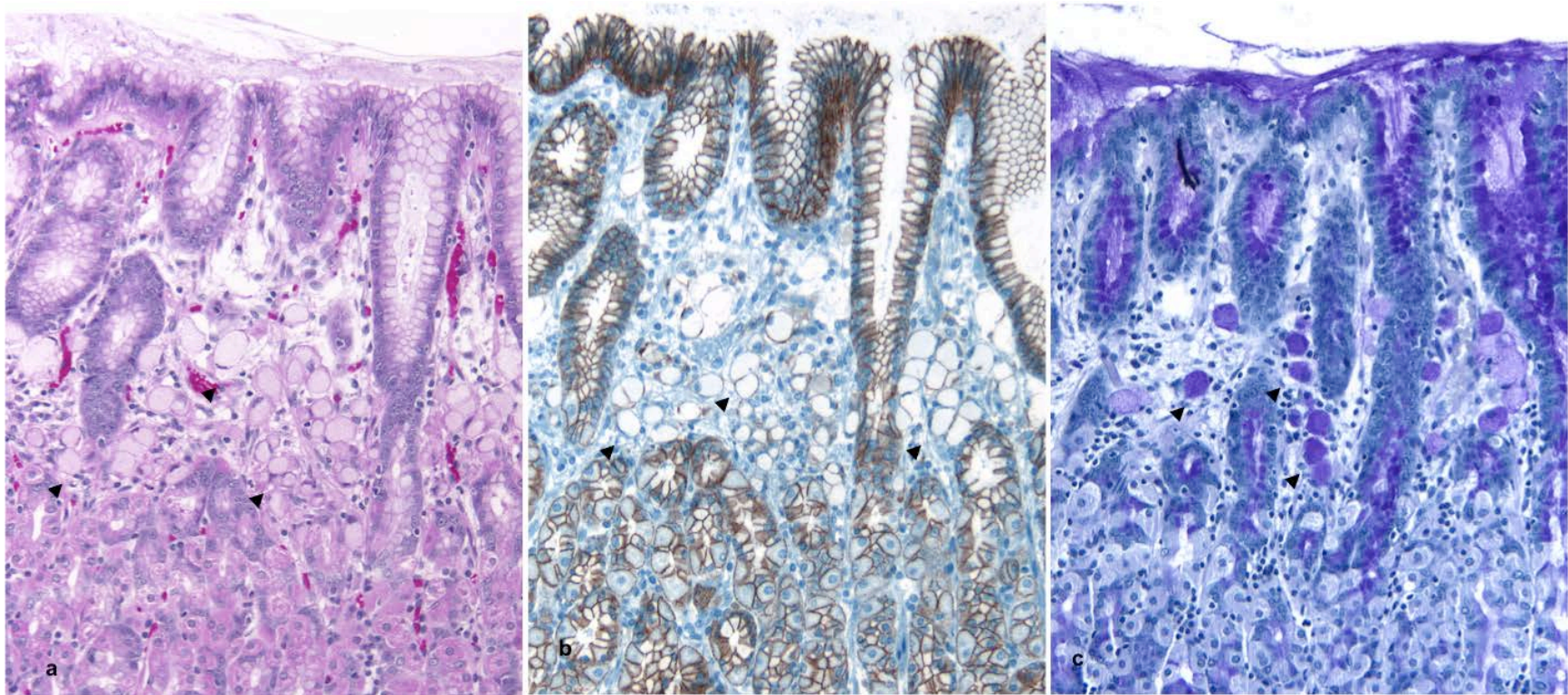
The Lauren classification of DGC is analogous to the Carneiro classification system's isolated cell type; as it is to the World Health Organization's classification of signet ring cell type [20]. Pathology reports indicating undifferentiated, mucinous adenocarcinoma or poorly differentiated adenocarcinomas also raise the index of suspicion for DGC. DGC typically exhibits decreased or absent immunohistochemical staining for E-cadherin, consistent with its disorganized architecture (Figure 1b). Recognition of families with an autosomal dominant predisposition towards DGC led to the discovery of causative germline mutations in *CDH1* [21]. To date, *CDH1* remains the only gene associated with HDGC; likewise, germline aberrations in *CDH1* are exclusive to the syndrome, emphasizing the importance of the pathologic classification of these tumors.

1.1.1.1.2   Epidemiology Of The Two Types Of Gastric Cancer

The differences between the two types of GC extend beyond their morphologic appearances to their risk factors and patient demographics. As compared with DGC, the incidence of IGC increases more with age and affects males more than females. Worldwide there is marked variation in the incidence of GC and the proportion of the two subtypes. The highest rates of GC are found in Japan, China, Eastern Europe, and South America and the lowest in North America, Northern Europe, Southeastern Asia, and Northern and Western Africa [20, 22]. IGC comprises the majority of GC diagnoses in higher incidence countries, while DGC forms a higher proportion of GC cases in lower incidence as compared to higher incidence countries [20].

Environmental factors contributing more than genetic background to the development of IGC are thought to be responsible for these disparities. Chronic gastric mucosal infection with *Helicobacter pylori* leading to a chronic atrophic gastritis [23, 24] is the most well-recognized environmental risk factor for GC, with a relative risk of 5.9 for non-cardia GC[25]. Compared to the vast global rates of *H. pylori* infection, only a relatively small proportion of infected individuals go on to develop GC. This reflects the influence of genetic factors in the bacteria and the host. For example, strains of *H. pylori* containing the virulence factor cytotoxin-associated gene A (cagA) are carcinogenic [26]. CagA is a secreted bacterial oncoprotein introduced into gastric epithelial cells by bacterial

secretion machinery [27]. When phosphorylated by Src or Abl kinase, it deregulates the tyrosine phosphatase Src homology related protein (SHP-2) that acts upstream of the oncogenic Ras MAP kinase pathway [28]. Genetic variations in the host, such as particular polymorphisms in genes for the inflammatory mediators; IL-1β, IL-1 receptor antagonist, TNF-α, IL-10, and IFNγR1 [29-31] dictate the type of immune and inflammatory response triggered by *H. pylori* infection. These bacterial and host genetic factors contribute to the progression of gastritis to chronic atrophic gastritis, to intestinal metaplasia, and finally GC. Additionally, other environmental factors such as smoking contribute to GC risk [32]. Furthermore, diets high in salt, nitrites or smoked foods, pickled vegetables, and low in fruit and vegetable intake [3, 26, 33], are also thought to increase GC risk.

The influence of environmental factors on the genesis of GC is evident by the diminution of GC risk with migration from a higher incidence to lower incidence area [34]. Over the past several decades there has been a decline in the incidence of the IGC in the United States [35]. This echoes the worldwide decline in the overall incidence of GC that has been attributed to alterations in diet, improved food storage and preservation, and decreased infection and colonization by *H. pylori*. The increased intake of fruits and vegetables combined with the advent of refrigeration has alleviated the need for food preservation by salt and other methods. Decreased crowding and improved living conditions are also felt to have reduced *H. pylori* exposure and as a result early colonization [36].

In contrast to the global GC incidence, that of DGC, in particular the signet ring cell type, is not decreasing. Indeed, in North America, it may even be rising [35, 37]. The underlying cause for this increased incidence is not understood. *H. pylori* infection poses a similar risk for DGC as it does for IGC[38], although DGC is not linked to a precursor lesion. A prospective study examining baseline surrogate markers of *H. pylori* infection and chronic atrophic gastritis in patients who developed IGC or DGC showed an association between low titers of antibodies against *H. pylori* surface antigen in those that developed IGC and increased titers of antibodies in those that developed DGC. *H. pylori* only infects normal gastric mucosa, therefore these findings were consistent with expectations of decreased rates of *H. pylori* colonization in chronic atrophic gastritis, a known precursor to IGC[39]. There is evidence to support epigenetic effects of *H. pylori* infection,

where promoter hypermethylation of *CDH1* in normal infected gastric mucosa was reversible with antibiotic treatment of the bacteria [40]. Furthermore, methylation of *CDH1*, among other tumor suppressors, has been demonstrated in normal gastric mucosa of patients with GC, independent of the epigenetic modifications associated with normal aging [41]. In the context of particular *H. pylori* strains, individuals with a family history of GC had an increased risk of GC, however due to the relatively small number of cases, there were no conclusions based on histological classifications [42]. Although there is no evidence of increased rates of *H. pylori* infection associated with the microscopic foci of DGC in the prophylactic gastrectomy specimens of *CDH1* mutation carriers, in light of its known role in GC carcinogenesis and in particular with regard to its ability to induce promoter hypermethylation of *CDH1*, *H. pylori* infection should be ruled out or treated in all *CDH1* mutation carriers.

1.1.1.1.3 Clinical Features Of Gastric Cancer

Despite its low incidence in North America (~10 per 100 000 men and women per year), GC still remains a major health burden. According to the National Cancer Institute's Surveillance Epidemiology and End Results database, the overall 5-year relative survival rate for invasive GC from 1996 to 2004 was 24.7% (http://seer.cancer.gov/). For the most part, the poor survival rates are indicative of the delay in diagnoses. Early GC is usually clinically silent. Occasionally, it can present with gastrointestinal symptoms such as epigastric pain, dyspepsia, a sensation of gastric fullness, or frank symptoms of gastric obstruction. More often, GC is only detected following constitutional symptoms such as loss of weight. By then, the GC has usually progressed to stage III, or locally invasive cancer. In countries where the incidences of GC are very high, nationwide screening programs utilize upper endoscopy as a means of detecting asymptomatic early stage GC amenable to treatment by endoscopic resection. In Japan, this type of screening has proven effective at reducing GC-mortality rates [43]. However, in low incidence countries, such as the United States, population based endoscopic screening has not been

implemented because the incidence is too low to justify such an invasive screening program [44].

1.1.1.1.4   Overview Of The Molecular Genetics Of Gastric Cancer

Global genome analysis of GC by array comparative genomic hybridization has revealed recurrent regions of somatic copy number aberrations (CNAs). Frequent gains have been detected at 20q13, 8q24, and 7p[45-48] and frequent losses at 18q21, 3p14, 17p [45, 46, 48]. By correlating CNA with expression data, Tsukamoto *et al*. identified 114 genes significantly over-expressed in 14 amplified regions and 11 genes down-regulated in 5 deleted regions [45]. This data correlated over-expression of *DDX27*, *ARFGEF2*, *C20orf199*, *Kua-UEV*, *PTPN1*, *PARD6B*, *ADNP*, and *DPM1* with 20q13 amplification that was present in 97% of the cases [45]. Deletion of 3p correlated with decreased expression of the putative tumor suppressor, *FHIT* [45], where abnormal sequence transcripts have been detected in a GC cell line [49] and decreased protein expression of FHIT has been found to correlate with undifferentiated tumors, diffuse histology and poor prognosis [50]. Over-expression of genes occurs at many other amplified regions in particular *ERBB2* at 17q21, and *EGFR* at 7p11. *ERBB2* over-expression has been correlated with IGC and has been found to be significantly increased in metastatic disease [51] and to correlate with poor prognosis [52]. *EGFR* expression has also been associated with IGC where expression in the primary GC was shown to independently predict poor prognosis regardless of the expression level in the metastatic lesion [51]. Deleted regions were also concordant with down regulation of candidate tumor suppressors; *SMAD4* at 18q21 and *CDKN2B* at 9p21 [45]. Normal gastric mucosa, intestinal and diffuse GC have been shown to have distinct cytogenetic profiles [53]. A consistent gain at 12q was reported in laser micro-dissected DGC (n=14) and laser micro-dissected signet ring cell GC (n=7) [45, 48].

1.1.1.1.4.1   The Tumor Suppressor p53

Somatic mutations in *TP53* encode the cell cycle control protein, p53, and are common to many cancers. Over 950 different *TP53* mutations have been reported in stomach cancer (http://www-p53.iarc.fr/, R13, November 2008) [54]. The majority of mutations cause missense changes and occur between exons 5 and 8 that encode the DNA binding domain of the protein [55]. Mutations in *TP53* are preferentially associated with IGC rather than DGC. In a series of 62 GC, 17 out of 50 (34%) IGC had associated *TP53* mutations as compared with 0 out of 12 cases of DGC [56]. Incidentally, both IGC and DGC can occur in association with germline *TP53* mutations that give rise to the familial cancer syndrome, Li Fraumeni (LFS). Individuals with LFS are predisposed to a range of primary cancers. The genetic risks of the non-synonymous arginine/proline polymorphism at residue 72 of *TP53* have also been examined. The proline allele confers a reduced apoptotic ability and increased risk of cancer to the individual [57]. Additionally, in individuals with advanced GC, the proline genotype was associated with a lower response rate to chemotherapy [58].

1.1.1.1.4.2   Mismatch Repair Genes

Approximately 15% sporadic GCs exhibit microsatellite instability (MSI) [59]. This is due to genetic or epigenetic perturbations of the mismatch repair genes, *MLH1* or *MSH2* [60]. MSI probably functions in tumor progression rather than tumor initiation. This is supported by the finding of decreased MLH1 protein expression and *MLH1* promoter hypermethylation in sporadic GC lesions with high MSI, but not in adjacent precursor lesions [61]. GCs with high MSI tend to mainly occur in the antrum, be of the intestinal type, exhibit a predominantly lymphocytic infiltrate, occur in the elderly, and have better survival rates with low metastatic rates [60-62]. Particular genes are frequently mutated in association with the defect in mismatch repair. There is high frequency of frameshift mutations found in the poly(A) tract of *TGFBRII*, the gene encoding a receptor for

transforming growth factor b [62]. There is no apparent correlation with *TP53* mutations [56]. A recent comparison of the expression profiles of GCs with MSI and GCs without MSI revealed differential expression of genes involved in immune response, apoptotic pathways and DNA repair pathways [56, 59]. This study and previous studies provide supportive evidence suggesting that the heightened immune response contributes to the longer survival rates.

Lynch syndrome [OMIM #120435] is associated with germline mutations in the mismatch repair genes and leads to the development of colorectal and other cancers with MSI [63]. GC risk in the context of Lynch syndrome will be discussed below.

1.1.1.1.4.3   E-Cadherin

Decreased E-cadherin expression is a feature of many poorly differentiated epithelial cancers [64-67]. In particular, E-cadherin expression is down-regulated in sporadic DGC [64]. As highlighted above, molecular genetic differences exist between IGC and DGC, however overall, loss of E-cadherin expression remains the major discriminator between the two subtypes.

1.1.1.2 The Molecular Biology Of CDH1 And The Putative Role Of E-Cadherin In Cancer

1.1.1.2.1   Structure And Function Of E-Cadherin

E-cadherin belongs to a large family of transmembrane glycoproteins and is the primary mediator of epithelial cell-cell adhesion [68]. It has multiple roles in morphogenesis, cell polarization, structural organization of tissues [69], and cell migration [70] and is essential for normal development. Mouse embryos deficient in the protein fail to form a

trophectodermal epithelium or a blastocyst [71]. *CDH1* [OMIM *192090] is located on chromosome 16q22.1. The genomic sequence of *CDH1* spans almost 100 kb and encodes 16 exons [72]. These 16 exons are transcribed and translated into the precursor protein that is cleaved prior to the delivery of molecules to the cell membrane as mature E-cadherin [73]. The mature E-cadherin protein contains three major domains, the extracellular domain encoded by exons 4-13, the transmembrane domain encoded by part of exon 13 and part of exon 14, and the highly conserved cytoplasmic domain encoded by the remainder of exon 14 to exon 16 [74]. E-cadherin is located at the baso-lateral surfaces of the epithelial cell where it forms dimers [75]. There, the large extracellular domain of E-cadherin, comprised of five cadherin repeats, homodimerizes with E-cadherin expressed on neighboring epithelial cells in a Ca^{2+}-dependent manner, mediating cell-cell adhesion at the zonula adherans junctions. The cytosolic, carboxy-terminus of E-cadherin binds to β−catenin and α-catenin that in turn binds to the F-actin microfilaments of the cytoskeleton via α-catenin [68].

Several molecules have been implicated in the regulation of membrane trafficking of E-cadherin. p120-catenin, located at the juxtamembrane domain, not only strengthens the adhesion between cells, but also plays a role in maintenance of E-cadherin at the membrane and degradation of the adhesion molecule [76, 77]. The members of the Rho family of GTPases contribute to epithelial morphogenesis, maintenance, adhesion and cell migration in part through the regulation of E-cadherin and their downstream effects on the organization of the actin cytoskeleton [78-80].

The expression of E-cadherin is subject to positive and negative transcriptional regulation. Transcriptional repressors, such as Snail, Slug, dEF1/ZEB-1, Sip1/ZEB-2, Twist and E12/E47, bind to the E-box motifs at the *CDH1* promoter [81, 82]. Other regulatory regions outside of the promoter have also been identified such as the enhancer element in intron 2 [83]. In *CDH1*, intron 2 accounts for the majority of non-coding intronic sequence (~60 kb) and contains conserved cis-regulatory elements. The importance of intron 2 for normal expression of the gene has been underlined by a study of murine embryonic development following deletion of the intron in early mouse embryogenesis, which showed reduced transcriptional activation of the gene [81].

1.1.1.2.2   Variations In *CDH1* And The Association With Cancer

A *CDH1* promoter polymorphism at -160 C/A has been shown *in vitro* to have a role in transcriptional regulation, where the A allele was shown to have decreased transcriptional efficiency and weaker transcription factor binding affinity [84]. Analysis of eight *CDH1* haplotype-tagging polymorphisms, within the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST) study, failed to demonstrate an elevated risk for GC for seven of the individual SNPs, including the -160 C/A polymorphism, or their associated haplotypes [85]. Likewise, no association was seen between the promoter polymorphism and GC risk in a recent Italian study [86]. However, meta-analysis ethnically stratifying cases and controls, revealed the -160 A allele to be a risk factor for GC in Europeans but not Asians [87]. As separate disease haplotypes in different populations could account for these discrepancies, it has been proposed that the positive associations could potentially be clinically relevant to the populations in which they were studied [88].

Recently another polymorphism in intron 2 was also associated with sporadic DGC in an Italian population [89]. This result will require replication in further studies.

The HDGC-associated germline *CDH1* mutations are dispersed across the gene [90] (Figure 2). These mutations interfere with normal E-cadherin function in a variety of ways from alterations to conserved amino acid residues with predicted effects on protein structure, to deletions of critical domains, to protein truncation and haploinsufficiency due to nonsense-mediated mRNA decay. Recently our group reported large deletions as another genetic aberration of *CDH1* in 3.8% of HDGC families [91].

Figure 2 DGC and LBC associated CDH1 germline mutations.

Colors denote type of mutation (light blue: insertion/deletion; brown: splice site; grey: truncating; dark blue: missense). Mutations below *CDH1* occur in families with LBC history. Mutations marked with (*) indicates breast cancer history but not LBC. Reproduced from [Hereditary Diffuse Gastric Cancer'., Schrader KA, Huntsman D, 155, 33-63, 2010] with permission from Springer.
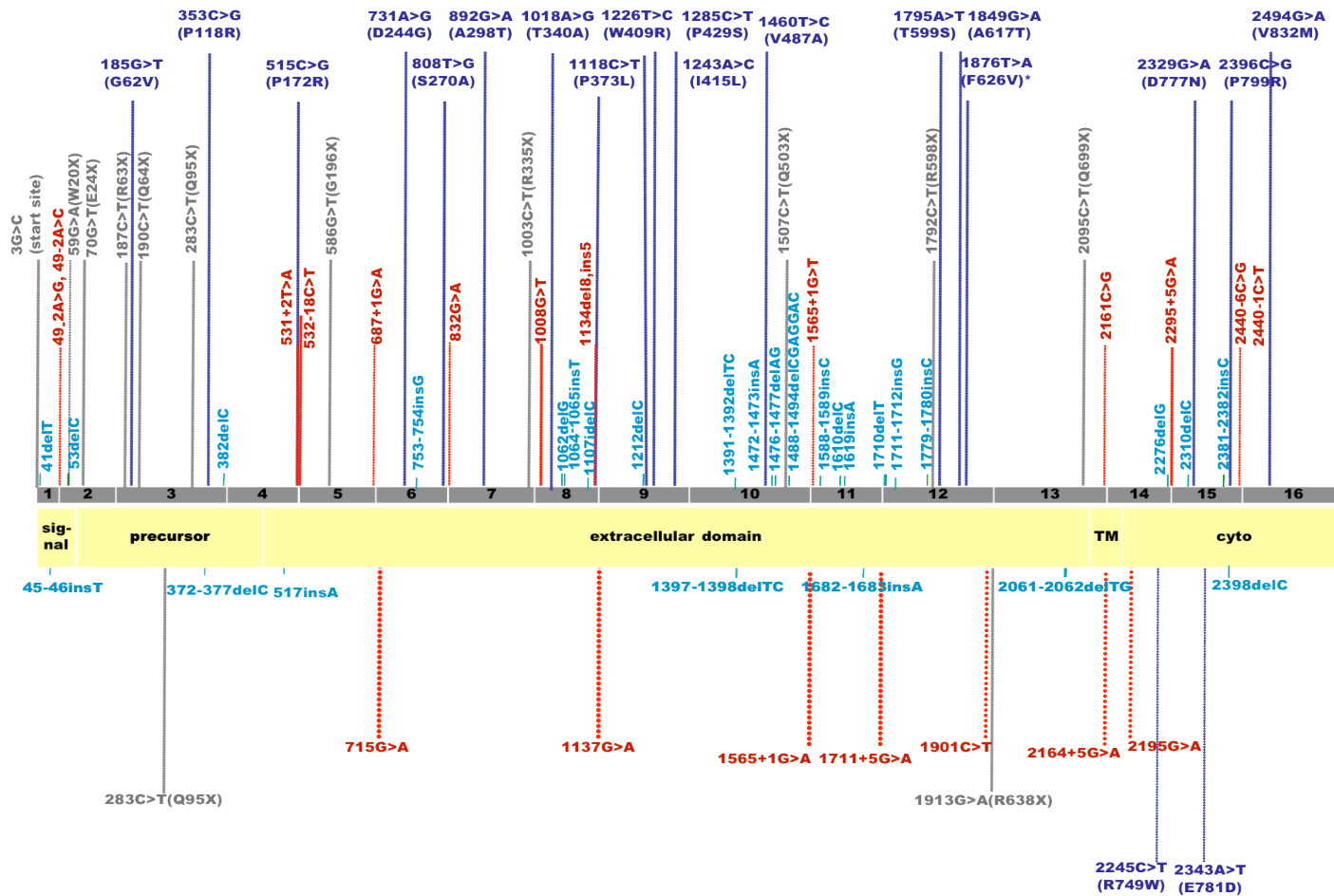
Haploinsufficiency for E-cadherin is sufficient for normal development. However, there have been two families reported in which the inheritance of splicing mutations in regions encoding the extracellular domain of E-cadherin (intron 4 splicing donor site; c.531+2 T>A and exon 8; c.1137G>A), have been associated with cleft lip with or without cleft palate [92]. Both mutations led to aberrant splicing that created in-frame deletions predicted to escape nonsense-mediated mRNA decay. Nonsense-mediated mRNA decay is the degradation of mRNA molecules containing a premature stop codon greater than 50 nucleotides prior to the last splice junction [93]. The abnormal splicing created by each of these mutations would result in a protein lacking parts of its extracellular cadherin binding domains. As E-cadherin is expressed in the frontonasal prominence, and the lateral and medial nasal prominences during the critical stages of lip and palate development [92], the authors postulated that the aberrant E-cadherin proteins might exert a dominant-negative effect over the WT E-cadherin protein by abnormal homodimerization. This association with cleft lip +/- cleft palate, however, was not seen in two other families with the c.1137G>A mutation [94], suggesting that the previous observation could have been due to a gene-environment interaction.

1.1.1.2.3   Loss Of E-Cadherin And Cancer

The role of *CDH1* in cancer is believed to be related to the promotion of invasiveness caused by the loss of E-cadherin expression [95]. Cells deficient in E-cadherin lose the ability to adhere to each other and therefore become more invasive and metastasize [96]. The silencing of E-cadherin expression requires inactivation of both *CDH1* alleles either at the genetic or epigenetic level. Intriguingly, re-expression of E-cadherin has been observed in the tumor cells at the metastatic site [97]. In sporadic DGC, the inactivation of the first allele is typically by mutations clustering in exons 8 and 9 resulting in exon-skipping and in-frame deletions of the extracellular domain [68]. Mutations and deletions in this critical area have been shown to have functional consequences [98]. Mutations in *CDH1* can be found in 50% of GC tumor specimens [99], where the inactivation of the remaining normal allele is often by hypermethylation of the *CDH1* promoter [100].

Loss of E-cadherin expression has been shown to be an early event as depicted by the in situ DGC lesion from a PTG specimen of a *CDH1* mutation carrier shown in Figure 3. Figure 3a is the H&E stain of the lesion and Figure 3b shows the loss of membrane E-cadherin staining in the in situ signet ring cells indicating that the loss of E-cadherin is an early event that precedes invasion. Additionally, in sporadic LBC, *in situ* cancers, situated beside their invasive counterparts also stain negatively for the cell adhesion molecule [101]. Moreover, neighbouring invasive and *in situ* LBC both share the same mutations in *CDH1* and harbor loss of heterozygosity (LOH) of 16q [102], indicating that loss of E-cadherin is an early initiating event. The mechanism by which loss of E-cadherin protein expression occurs varies. E-cadherin expression can be heterogeneous depending on the part of the tumor that is being tested. In addition to interpatient heterogeneity of the mechanisms that cause loss of expression of the normal allele of *CDH1*, there is also intrapatient heterogeneity whereby different silencing mechanisms can be seen across and within patients' tumors [103]. Decreased expression of E-cadherin can also be a transient event, facilitating invasion and metastasis [104], with subsequent re-expression of E-cadherin in the metastatic cells [97]. Recently LOH was more frequently seen as the second hit in metastatic tumors [103].

Figure 3 These pictures show a small in situ focus of a DGC from the same prophylactic gastrectomy specimen as shown in Figure 1.

(a) H&E stain. Note the similarity between the signet ring cells within the duct and the cross-section of a mucosal blood vessel; (b) E-cadherin stain showing down-regulated expression in the signet ring cells of the in situ focus of DGC in comparison to the normal E-cadherin-positive epithelium. This picture implies that loss of E-cadherin expression is an early event in tumorigenesis. Photographs taken by Dr. Martin Köbel. Reproduced from [Hereditary Diffuse Gastric Cancer'., Schrader KA, Huntsman D, 155, 33-63, 2010] with permission from Springer.
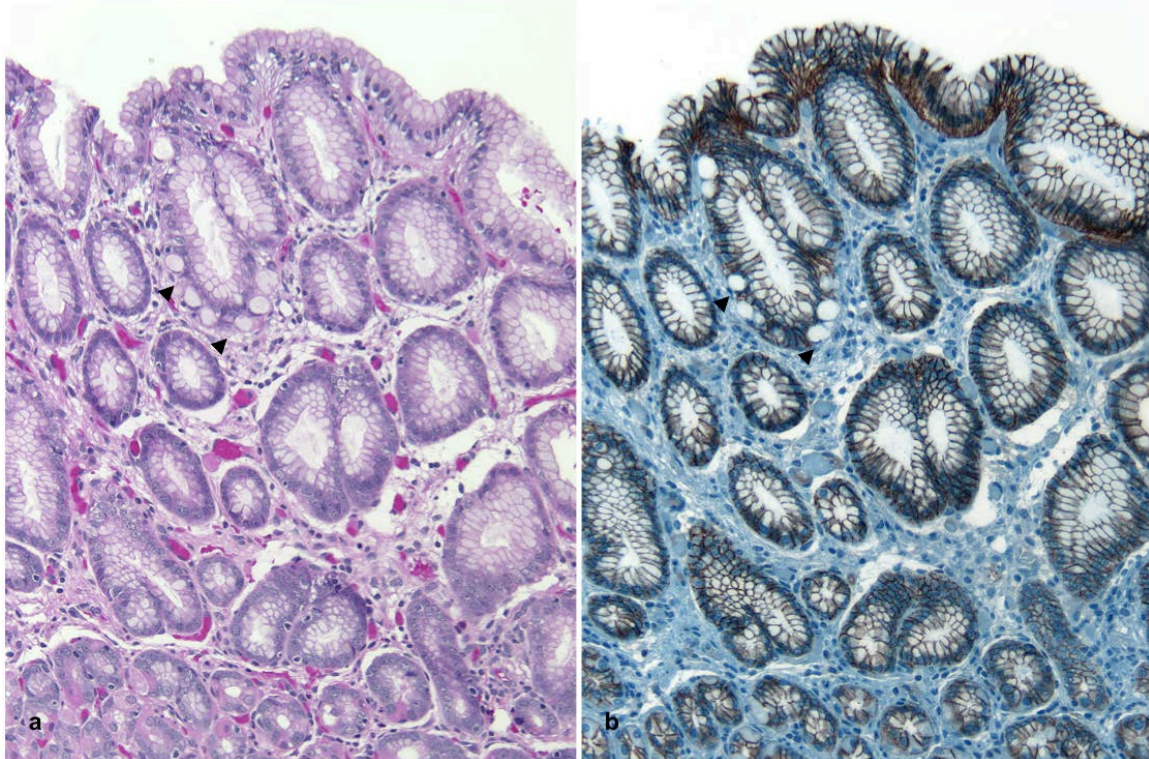
The tumor suppressor function of E-cadherin [105-107] is supported by evidence of loss of expression of the other *CDH1* allele [103, 108, 109]. HDGC-associated GC exhibit a lack of expression of E-cadherin from the normal allele of *CDH1* that is achieved by epigenetic suppression of transcription, or by mutation or LOH [103, 108, 109]. LOH is a common phenomenon seen in association with loss of expression of tumor suppressor genes [110]. It refers to the somatic loss of the WT allele usually due to deletion of the gene or loss of a whole chromosome arm. It is detected by comparing microsatellite markers linked to the gene of interest in germline and tumor DNA. The markers in germline DNA are heterozygous, therefore the appearance of homozygosity in the markers of somatic tumor cells infer that there has been a loss of the WT allele [93].

The tumor suppressor role of E-cadherin is thought to be in part due to its association with β−catenin a key player in the canonical Wnt signaling pathway [111]. The Wnt signaling pathway is implicated in familial adenomatous polyposis (FAP) where germline mutations in the *APC* gene [112] cause the autosomal dominant predisposition to gastrointestinal polyposis. Both β−catenin and *APC* are phosphorylated by the kinase, GSK3β resulting in ubiquitination and degradation of β−catenin. Activation of the Wnt signaling cascade inhibits the activity of GSK3β. This causes an increase in the free cytoplasmic β−catenin molecule, which then translocates to the nucleus and binds to the transcription factor, lymphocyte enhancer factor/T cell factor (LEF1/ Tcf). This results in transcription of Wnt responsive genes such as the oncogene, c-Myc [113, 114]. In addition to this role in regulating gene transcription, β−catenin also functions in epithelial cell adhesion through its association with E-cadherin and β−catenin. This association is thought to sequestrate β−catenin at the plasma membrane, thus preventing it from entering the nucleus. The proposed existence of different forms of β−catenin with distinct binding properties has shed light on how the roles of β−catenin in cell adhesion and nuclear signaling might be regulated [115]. Thus, further elucidation of E-cadherin's relationship with this canonical oncogenic pathway is awaited.

1.1.1.3 Hereditary Gastric Cancer

1.1.1.3.1   The Crucial Role Of Family History

Five to ten percent of GCs demonstrate familial clustering [116]. Shared environmental factors, such as diet and *H. pylori* infection, account for the majority of familial clustering of the intestinal type, although approximately 5% of the total GC burden is thought to be due to germline mutations in genes causing highly penetrant, autosomal dominant predispositions to cancer such as Lynch syndrome, Peutz-Jehgers syndrome (PJS), Li-Fraumeni syndrome (LFS), FAP, and HDGC [63].

Lynch syndrome or hereditary nonpolyposis colorectal cancer is caused by germline mutations in the mismatch repair genes; *MSH2*, *MLH1*, *MSH6*, *PMS1*, *PMS2*. The syndrome is mainly characterized by susceptibility to colorectal cancer. However, after endometrial cancer, GC is the third most common cancer in these patients (in countries of low GC incidence). In a case series from the United Kingdom, GC accounted for 5% of cancers in families harboring *MLH1* or *MSH2* mutations [117].

IGC is the predominant subtype in Lynch syndrome [63]. The original Lynch syndrome family initially presented with a susceptibility to gastric and uterine cancer. However over the years, the incidence of GC within this large pedigree has become insignificant compared to the incidence of cancer of the colon and endometrium [118]. This decrease in the incidence of GC in germline mutation-carrying families largely echoes the overall decline in GC incidence in the general population although, in countries with higher incidences of GC, it is the second most common tumor associated with Lynch syndrome [119, 120].

With a relative risk of 213, GC is considered an integral tumor of the PJS caused by mutations in *STK11* [121].

LFS due to mutations in *TP53* or *CHEK2*, is associated with both IGC and DGC [44, 122, 123].

18

FAP is caused by germline mutations in *APC*. GC occurs in 0.6% of patients [124]. A greater number of reports of GC (in particular IGC) associated with FAP have been reported in individuals from Japan, consistent with the overall higher incidences of sporadic GC in that population [125].

Increased risks of GC have also been found to be associated with *BRCA1* [126] and *BRCA2* mutation carriers [127, 128]. Reports of other genetic syndromes associated with GC exist although due to their paucity, it is hard to establish true associations.

Genome-wide association studies have uncovered low to moderate risk susceptibility genes for GC, although currently, the clinical significance of these results is hard to interpret. Recently an intronic SNP in *PSCA*, encoding prostate stem cell antigen, was identified in Japanese and Korean subjects, as having a significant association with DGC with an allele specific odds ratio = 1.62, 95%CI=1.39-1.89[129]. Although the exact function of *PSCA* is not known, the protein is expressed in the normal gastric epithelium and lost in diffuse adenocarcinoma cells, indicating a possible tumor suppressor role in the gastric epithelium [129].

As previously mentioned the -160 A/C promoter polymorphism of *CDH1* has also been investigated as increasing GC risk.

Until there is further understanding regarding the genetic variability amongst individuals who develop GC, the clinical interpretation of low-to-moderate penetrance genes associated with GC susceptibility, will remain difficult. Even if validated, the relative risks associated with the -160 A/C *CDH1* polymorphism and other germline polymorphisms such as in *PSCA*, are not high enough to be used to triage screening. Thus at this point, they do not appear clinically relevant [129].

Additionally, the interplay between environmental risk factors and the host's genetic background will also need to be considered. As previously eluded to, the polymorphisms; IL-1B -31 T+, in the gene encoding IL-1β and IL-1RN*2*2, in the gene encoding the receptor for IL-1β are thought to increase levels of IL-1β when the host is infected with *H. pylori*, leading to hypochlorhidria and increased GC risk. The example of the IL-1β response to *H. pylori* infection highlights the importance of understanding gene-

environment interactions to identify potentially modifiable risk factors such as *H. pylori* infection.

1.1.1.3.2   Hereditary Diffuse Gastric Cancer

The report of a Maori family with multiple cases of DGC inherited in a highly penetrant, autosomal dominant manner was first published in 1964 [130]. Three decades later this large family and two other Maori families with similar histories, were analyzed using genetic linkage analysis to define a region on the long arm of chromosome 16 that included the *CDH1* locus [21]. Armed with this information and the knowledge of the role of somatic mutations of *CDH1* in sporadic GCs, Guilford identified *CDH1* germline truncating mutations in all three families [21]. This discovery has led to the subsequent identification of many more HDGC families of different ethnicities caused by novel or recurrent germline *CDH1* mutations or deletions [4, 90, 92, 94, 123, 131-148].

The International GC Linkage Consortium (IGCLC) was launched soon after the discovery of *CDH1* as a susceptibility gene for DGC. This created an international, multidisciplinary collaboration to develop a unified approach to the research and clinical management of the new syndrome HDGC [133]. The collective experience in testing over 160 probands from around the world has been that roughly half of these families can be accounted for by germline mutations or large deletions in *CDH1*. In families with HDGC, the risk for DGC appears to be independent of the common risk factors mentioned earlier.

Individuals harboring germline E-cadherin mutations have a lifetime risk of developing GC of 40-67% for males and a 63-83% for females [94, 131]. In both penetrance studies, females had a higher risk of developing GC [94, 131]. However, as we continue to extend family histories and find new HDGC families, recent unpublished data by the collaborative efforts of the IGCLC suggest that the risk for GC in males and females may be more similar than originally estimated (unpublished data). The average age of developing DGC is 38 years [74]; however, the range extends from fourteen years of age up to 85 years of age [131]. The factors that determine the age of onset in a family remain to be

elucidated. Thus, until it is understood what factors put people at higher risk for early-onset disease, appropriate screening should commence at least 5-10 years prior to the earliest reported diagnoses of cancer.

As of yet, no other genes have been associated with HDGC. Candidate gene studies in Portuguese families without *CDH1* mutations did not find germline mutations in SMAD or caspase-10 [122]. The authors did identify a germline mutation in *TP53* in a family with multiple cases of GC, although the histology of these cancers was not available [122]. Likewise there were no germline mutations in the candidate genes *RUNX3* and *HPP1* in German GC families [123]. Again, these investigators also found a germline *TP53* mutation in a 52-year-old proband with DGC and a family history of GC, leukemia (age 17) and hepatocellular carcinoma (age 34) in three first-degree relatives [123]. Germline *MET* mutations have also been found in two Korean probands with GC. The first had IGC with no age or family history while the second occurred in a proband with DGC from a family selected based on the criteria of two first- or second-degree relatives affected with GC, at least one of whom was diagnosed with cancer before the age of 50 years [149]. Molecular testing for germline mutations in *MET* and other putative candidate genes such as *CTNNB1*, encoding β–catenin, in our *CDH1*-negative HDGC families has been negative (unpublished data). Even though mutations in *CDH1* may not be detected in all HDGC families, it has been shown that the majority of HDGC families display an imbalance of allele-specific *CDH1* expression, thus still implicating the locus in a proportion of *CDH1* mutation-negative HDGC families [150]. It is therefore possible that families with a compelling history of HDGC in whom coding mutations or deletions have not been identified, could have pathogenic mutations in regulatory or other non-coding regions of the *CDH1* gene [150].

1.1.1.3.3   Identification Of At-Risk Individuals

The frequency with which *CDH1* germline mutations are detected in families with HDGC varies regionally, being higher in regions where there are low incidences of GC [4,

[91, 94, 122, 123, 132]. In 1999, the definition of HDGC set forth by the IGCLC was any family meeting either of the following criteria: (1) two or more documented cases of DGC in first- or second-degree relatives, with at least one diagnosed under the age of 50 years, or (2) three or more cases of documented DGC in first- or second-degree relatives, regardless of age of onset [74]. Using the initial selection criteria, the detection rates for germline mutations of *CDH1* have varied from as low as 11% [151] in high-incidence countries like Portugal to 30% in low-incidence areas such as North America [4]. To reflect the growing experience with HDGC, the updated IGCLC guidelines extend *CDH1* genetic testing to families with two cases of GC in which one case is histopathologically confirmed as DGC and diagnosed before the age of 50 [152]. In addition, the guidelines endorse genetic testing of *CDH1* in families with both LBC and DGC, with one diagnosed before the age of 50, and in probands diagnosed with DGC before the age of 40, with no family history of GC [4, 132, 152]. Recently we surveyed the incidence of *CDH1* aberrations in our HDGC families combined with HDGC families from different parts of the world that had either (1) three or more DGC in first degree relatives diagnosed at any age or (2) two or more GC in first-degree relatives with at least one DGC diagnosed before age 50 years [91] and found that aberrations in *CDH1* occur in 46% of families. Keeping in mind that the majority of families came from areas of low GC incidence, this detection frequency likely overestimates the global contribution of *CDH1* mutations to HDGC families meeting these criteria, which likely lies around 25-30%.

1.1.1.4 Genetic Testing For CDH1

1.1.1.4.1   Genetic Counseling

Full screening of the *CDH1* gene is recommended in an individual fulfilling the HDGC criteria. DNA can generally be extracted from blood leukocytes, lymphocytes in saliva, or, with more difficulty and less accuracy, from normal tissue from paraffin blocks. Due to the problems with obtaining good quality DNA from paraffin blocks, an effort is

always made to test DNA from living individuals. The decision to undergo genetic testing should only be made following adequate genetic counseling. There should be pre- and post-genetic testing counseling available that should provide the patient with information regarding HDGC, its mode of inheritance, and penetrance estimates of developing DGC and LBC. A discussion regarding the management options following a positive result (identification of a germline *CDH1* mutation or deletion) should be presented in the pre-genetic counseling appointment. Additionally, the patient should be made aware of the general risks and benefits of genetic testing.

The discussion of genetic testing should include ensuring that they understand the limitations of the analysis. While a negative result could indicate that the cancers in the family are unrelated to *CDH1*, it could also occur if a particular genetic abnormality of *CDH1* was not detected by the assay, resulting in a false-negative outcome. Thus, following a negative diagnostic test, cancer screening in the proband and blood-related family members should continue as before. Due to the uniqueness of each family's mutation, predictive testing can only become available to other members of the family at-risk once a mutation is found in an affected person or obligate carrier. Carrier testing of unaffected individuals allows for risk stratification and focusing of high intensity screening in only those who are at risk.

In those who test negative for the family mutation, the risk of DGC and LBC returns to that of the general population and therefore screening for these individuals can be relaxed to population guidelines.

The psychosocial effects of genetic testing should be recognized, where some individuals may experience anxiety and distress relating to the results of the testing with regard to their personal and/or family risk of inherited cancer. This can potentially cause psychological distress in the individual and can affect family relationships.

As with most adult-onset genetic conditions, predictive testing is not generally offered to minors. However, as there are reports of individuals as young as 14 years of age being affected with DGC [21], with the assent or consent of the parents or guardians and the appropriate consent from the minor, there are exceptions that can be made on a case-by-

case basis. In this scenario, predictive testing would be used in order to determine if high-intensity surveillance would be necessary.

1.1.1.4.2 Methods Of Testing:

1.1.1.4.2.1 Mutation Screening

As germline *CDH1* mutations are heterozygous, various screening techniques designed to detect heterozygosity in the DNA, have allowed targeted sequencing of exons displaying changes. Single strand conformation polymorphism (SSCP) rapidly detects single nucleotide substitutions in PCR amplicons by resolving differences in the electrophoretic mobility of the single-stranded amplicons [153]. The sensitivity of SSCP for mutation detection can be as high as 95% depending on the protocol [153], however SSCP requires highly stringent gel electrophoresis conditions.

Denaturing high-performance liquid chromatography (DHPLC) is an alternative method of mutation screening with improved sensitivity and capacity over SSCP. DHPLC detects heteroduplexes of the mutated and WT sequence upon partial denaturation and reannealing. The heteroduplexes are distinguished from the matched normal homoduplexes by their different melting temperatures on high-performance liquid chromatography. In both methods, exons in which sequence variations are detected are then bidirectionally sequenced to identify the heterozygous change. The popularity of these methods compared to direct sequencing of the gene was their lower cost. However, as sequencing costs are now a fraction of what they were 10 years ago, most laboratories have abandoned such techniques and use direct sequencing.

1.1.1.4.2.2   Sequencing

Currently in our laboratory, we screen for heterozygous mutations of *CDH1* by bidirectionally sequencing the entire coding portion of *CDH1* including intron-exon boundaries [154]. The mutations range from small insertions and deletions to single base substitutions all of which can cause frameshifts or splicing abnormalities and lead to truncation of the protein or instability of the mRNA through nonsense mediated mRNA decay. Truncating mutations are assumed to be pathogenic, whereas missense mutations that result in changes in an amino acid are harder to interpret in terms of their potential effect on E-cadherin's function, as distinguished from harmless variations in the gene. Computer software programs are used to predict the effect of a mutation on splicing and with regard to whether or not the amino acid change might affect the function of the protein. Although in general these predictions need to be validated by functional assays. Another test for pathogenic germline mutations in *CDH1* is that they should segregate with affected family members.

Functional characterization of a potentially pathogenic variant in *CDH1* is usually carried out by expression of a corresponding cDNA in a cell lines that do not usually express E-cadherin [4,207]. The effect of expressing the E-cadherin with the variant amino acid in this cell line can then be compared with the effect of expressing the WT protein. E-cadherin function can then be assessed by assays studying cell-aggregation, and cell-invasiveness [4, 207]. Expression of the WT E-cadherin reverses the abnormalities in the E-cadherin negative cell line whereas expression of the mutated E-cadherin exhibits none or partial restoration of E-cadherin function. Pathogenic mutants of E-cadherin only partially reverse the defects in the cell lines such as decreased cell aggregation and increased invasiveness [4, 249].

Direct assessment of mutations potentially involved in splicing can be by RNA analysis. If normal fresh frozen gastric tissue is not available for RNA extraction, *CDH1* is also expressed in leukocytes and mucosal epithelial cells of the mouth, therefore RNA extraction from blood or saliva samples is also possible. RT-PCR is performed on the

patients RNA to create the coding DNA in order to determine if abnormal transcripts are present[207].

Minigene assays can also be used to determine splicing effects of a mutation, by creating a expression construct that harbors the exon with the mutation of interest surrounded by its neighboring introns and exons, the identification of unexpected transcripts indicates that the mutation alters normal splicing[94].

1.1.1.4.2.3   Large Deletion Analysis

Mutation-negative cases are subjected to multiplex ligation-dependent probe amplification (MLPA), a method that enables detection of copy number variation in genomic sequences[91]. Using this technique, our group has identified large deletions in *CDH1* that segregate with disease in 6.5% of HDGC *CDH1* mutation-negative families [91]. Overall large deletions of *CDH1* account for approximately 4% of HDGC [91].

1.1.1.4.2.4   Testing Stratification

In Newfoundland, an island province located off of the east coast of Canada, we recently identified a founder mutation in several different branches of a large family [21]. In light of the isolated population and our discovery of four other mutations in different families of Newfoundland heritage, we currently test families of Newfoundland heritage using a stepwise approach, consisting of an initial screen for these identified mutations prior to full *CDH1* sequencing

1.1.1.5 Clinical Management

1.1.1.5.1   Management For The Risk Of Gastric Cancer

Due to the highly penetrant nature of HDGC caused by mutations in *CDH1*, at-risk individuals should have annual surveillance endoscopy with greater than 30 random biopsies, beginning in their early twenties [14, 63, 152]. A detailed description of surveillance protocols can be found the latest consensus guidelines from the ILCGC [152]. The necessity for multiple biopsies is supported by the finding that increasing numbers of random biopsies taken on surveillance endoscopy positively correlate with detection of invasive foci of DGC [15]. The decision of when to start surveillance is based on the average age of DGC diagnosis being 40 years, although there are families in which individuals as young as 14 years of age have been diagnosed [133]. Thus screening of at-risk individuals should generally begin 5-10 years prior to the earliest cancer diagnosis in the family. At-risk individuals are those who are known to carry mutations in *CDH1* or those who belong to HDGC families and *CDH1* mutation status is not known.

Several other screening modalities have been tested including chromoendoscopy [10], PET scan [155], endoscopic ultrasound, stool for guaiac, abdominal CT, and multiple random stomach biopsies [14]. Unfortunately these do not reliably detect DGC, as demonstrated by the finding of multiple small cancer foci in six out of six gastrectomy specimens from *CDH1* mutation carriers only a week following an unremarkable panel of these investigations [14]. Despite the inability of endoscopy to reliably detect very small cancer foci, it has a greater likelihood of identifying clinically relevant cancers of more advanced stage that are more likely to metastasize. Therefore regular surveillance by endoscopy with multiple random biopsies still remains an important alternative to gastrectomy [14] and should be strongly recommended in those delaying PTG or electing against it.

1.1.1.5.2   Prophylactic Total Gastrectomy

Prophylactic total gastrectomy is recommended for *CDH1* germline mutation carriers. PTG is achieved by Roux-en-y esophagojejunostomy [14] with extreme caution as to obtaining adequate proximal margins to ensure all of the gastric mucosa has been removed. The chief argument for undertaking such a dramatic risk-reduction strategy is that multiple PTGs carried out in germline *CDH1* mutation carriers, have retrospectively become curative surgeries upon the finding of multiple small foci of invasive DGC within the resected organs [5-17, 63].

PTG is a major operation where, beyond surgical complications such as anastomotic leakage, strictures, or septic complications, there is a virtually 100% morbidity rate for complications such as altered eating habits, loss of weight, and diarrhea [6]. In a young and healthy individual, the risk of mortality with total gastrectomy in an experienced surgeon's hands is estimated to be less than 1% [63]. These estimates are below those quoted in the literature (3.5%), but these are based upon total gastrectomies performed with curative intent for clinical GC in an older patient demographic [156].

Management by a multidisciplinary team approach that includes a dietician, gastroenterologist, geneticist and general surgeon is extremely important in order to counsel the patient adequately regarding the risks, benefits, and clinical sequelae of this major operation [157]. This surgery has a major impact on the patient's nutritional status and ability to maintain adequate caloric intake and normal vitamin and mineral stores without appropriate supplementation. Thus ongoing follow-up with the multidisciplinary team to monitor and correct any abnormal nutritional parameters is essential. Expected deficiencies post-gastrectomy include vitamin B12 deficiency, due to the removal of the production source for intrinsic factor, required to absorb the vitamin. There is also an expectation for the malabsorption of iron, calcium, folate, and the fat-soluble vitamins, underscoring the importance of the involvement of a multidisciplinary team to monitor for this. The morbidity that can be expected post-gastrectomy usually worsens in the first 3-6 months post-gastrectomy but then gradually improves [14]. Due to the weight loss and nutritional implications, prophylactic gastrectomy is not generally recommended until the

growth period is finished. However this decision must also be weighed against the age of the youngest person in the family diagnosed with GC [5]. In families where there are cases of early-onset GC, prophylactic gastrectomy should be considered sooner on a case-by-case basis in combination with earlier commencement of regular endoscopic screening prior to surgery. In the past we have been hesitant to recommend gastrectomy in females prior to completion of childbearing; however, we have recently been acquiring encouraging evidence to suggest that women can successfully carry healthy pregnancies post-gastrectomy [158].

To date there have not been any reports of cancer in a member of an HDGC family post-PTG.

1.1.1.6 Aberrations Of CDH1 And Lobular Breast Cancer

In addition to the high lifetime risk of GC, in females within HDGC families there is an increased lifetime risk of breast cancer (39% - 52%) [94, 131]. In HDGC families there is particular association with the lobular breast cancer (LBC). The average age of onset for breast cancer was found to be 53 years [131].

We have reported two novel germline *CDH1* mutations; one in a family with hereditary LBC and no known history of GC, and the other in a family in which LBC was the predominant cancer diagnosis [159, 160]. No genotype-phenotype relationships have been determined for the mutations seen in hereditary LBC or LBC-associated HDGC families. As breast cancers related to germline *CDH1* mutation carrier status correlate with the lobular subtype, the capacity exists to identify potential *CDH1* mutation carriers based on morphologic grounds. To date, our data show that pathogenic *CDH1* variants are very rare in women with LBC either (1) diagnosed at a young age or (2) with a family history of breast cancer[161]. It is likely that improved detection rates will depend upon more stringent selection criteria such as multiple early-onset cases of LBC in first-or second-degree relatives, or alternatively multiple cases of LBC in addition to a history of GC.

1.1.1.6.1 Epidemiology And Pathology Of LBC

In North America, breast cancer is the most common cancer diagnosis in women where 1:9 women will develop the cancer in their lifetime [452]. The majority of primary breast cancers are adenocarcinomas, where infiltrating ductal carcinoma (IDC) accounts for the majority of breast cancer diagnoses and LBC only comprises about 10% of cases. LBC characteristically has a loose, ill-defined architecture as compared with IDC [162]. Instead of forming discrete glandular structures, the malignant cells in LBC exhibit infiltrative behavior and dissociate from the ductal unit to become isolated and highly dispersive, invading the stroma in single files [101]. Signet ring cells analogous to those seen in DGC, are also seen in LBC and like DGC, LBC characteristically stains negative for E-cadherin [163].

1.1.1.6.2 Sporadic Breast Cancer

In addition to its role in GC, E-cadherin also plays a similar role in LBC. There are striking similarities between the behavioral and morphologic phenotypes of both the DGC and LBC. Both share features such as poor differentiation, and a high mucin content giving rise to a signet ring appearance. Individual cancer cells are also non-cohesive, highly dispersive and invasive. Thus sporadic LBC cells look and behave in a similar fashion to DGC where 86% stain negatively for E-cadherin [101]. Indeed, *CDH1* mutations can also be found in 56% of LBC tumor specimens [164]. In sporadic LBC, the majority of mutations are truncating [164], and the second hit is usually by LOH or promoter methylation [164, 165].

1.1.1.6.3 Hereditary Breast Cancer

Hereditary breast cancer accounts for 5-10% of breast cancer cases where a significant proportion of cases are caused by germline *BRCA1* or *BRCA2* mutations [166]. Other breast cancer susceptibility genes include *TP53* (LFS), and *PTEN* (Cowden syndrome), *ATM*, *BRIP1*, *PALB2* and *CHEK2* [167, 168].

Germline *CDH1* mutations have been shown to have a role in hereditary LBC. The potential association of LBC to HDGC was postulated soon after there appeared to be an increased incidence of breast cancer in the HDGC syndrome. This was on the basis of known *CDH1* aberrations in sporadic LBC [133]. Keller et al. initially described an LBC and a DGC in a *CDH1* mutation carrier [169]. Further supportive evidence came from the identification of HDGC families in which there was an overrepresentation of the LBC subtype [4, 132]. This observation has led to efforts to determine whether or not *CDH1* is a breast cancer susceptibility gene, distinct from its GC risk. The risk seems to be only for female breast cancer as there have not been any reports of male breast cancer associated with HDGC families. By screening for germline mutations of *CDH1* in LBC probands selected based on young age or family history of breast cancer, we confirmed the association of LBC with germline mutations of *CDH1* and reported a novel germline *CDH1* truncating mutation (517insA) in a LBC family with no known history of GC [159].

1.1.1.6.4 Lobular Breast Cancer Risk

1.1.1.6.4.1 Screening

Currently there is not enough data on women with germline *CDH1* mutations and the development of breast cancer to determine the best risk-reduction and breast cancer screening strategies. Thus, recommendations for LBC risk management for women who are known carriers of *CDH1* mutations or those that have an unknown mutation status,

are derived from the experiences with managing other highly penetrant familial breast cancer syndromes. In accordance with recommendations for screening other highly penetrant hereditary breast cancer syndromes, these women should be referred to a high-risk breast clinic, be offered annual screening mammograms and breast MRI; perform monthly breast self-examination and have semi-annual clinical breast examination starting by age 30-35 years, or 5-10 years prior to the earliest breast cancer diagnosis in the family [152, 160, 170]. The American Cancer Society recommends MRI in addition to mammography in women with a lifetime risk of breast cancer greater than 20-25% [171]. Thus, the 39-52% lifetime risk of breast cancer in women conferred by germline *CDH1* mutations [94, 131] well exceeds their minimum range. LBC is difficult to detect by mammography because they do not form masses or develop calcifications, thus the use of MRI in this hereditary cancer syndrome where there is a particular susceptibility to LBC is attractive. Furthermore, there is evidence to suggest some increased detection of LCIS [172].

1.1.1.6.4.2   Chemoprophylaxis

Most LBCs are estrogen-receptor positive [162], and as both tamoxifen and raloxifene have been shown to reduce the risk of estrogen-receptor positive [173, 174] breast cancers in randomized trials, this is a conceivable strategy for chemoprevention [14], although at this time is unproven. Of theoretical benefit to *CDH1* mutation carriers is that the risk reduction with both agents, was greatest in women with lobular carcinoma in situ [175].

1.1.1.6.4.3   Prophylactic Mastectomy

Prophylactic mastectomy has been very effective as a primary risk reduction strategy in women with *BRCA1* or *BRCA2* mutations, reducing their risks up to 90% [176]. Prophylactic mastectomy may also be considered in *CDH1* mutation-positive women,

however at this time not enough data exist to recommend this as a primary risk-reduction strategy in *CDH1* mutation carriers. It would likely be a logical alternative for those women who have previously undergone treatment for breast cancer in one breast, or those who have withstood multiple false-positive biopsies requiring further confirmatory biopsies. Although prophylactic mastectomy can significantly decrease a woman's risk of developing breast cancer, women undergoing the procedure are at risk of a range of physical complications and potential psychological sequelae thus necessitating full counseling prior to the woman making a decision regarding the surgery [177]. The counseling should include the risk of possible altered perception of the body and the sexual relationship and the possibility of a negative physical impact of surgery [178].

1.1.1.7 Screening For Risk Of Other Cancers In CDH1

Although there have been reports of signet ring colon cancer in families with germline *CDH1* mutations [90, 132], currently there is not enough evidence to recommend colon cancer screening in all HDGC families. In HDGC families in which there is an additional family history of colon cancer, in particular of the signet ring cell subtype, it would be prudent to undertake more intense colon cancer screening such as commencing screening by colonoscopy every 3-5 years beginning at age 40 years or 10 years younger than the youngest colon cancer (whichever is younger)[152]. Thus, at this stage these families should be judged on a case-by-case basis.

Whether germline *CDH1* mutation carriers are at higher risk of other cancers still remains to be elucidated. Various other cancers have been reported in isolated families [94, 148]. Prostate cancer has been reported in a germline *CDH1* mutation carrier [134], and the -160 C/A *CDH1* polymorphism has also been implicated in association with the disease in Europeans and Asians [179], however currently, there is no conclusive association with this or other cancers.

1.2 Utilizing Distinct Phenotypes; The Application Of Next-Generation Sequencing To Disease Gene Identification

I began my PhD examining the association between germline mutations in *CDH1* and the predisposition to LBC, in addition to the DGC risk. HDGC is a prototypic example of the genotype-phenotype correlation between germline and somatic mutations in *CDH1* as they relate to two seemingly unrelated cancers, DGC and LBC. Based on the known association of LBC with germline mutations in *CDH1* as seen in HDGC families and in families with multiple LBCs, with one such family described in this thesis, I began by testing the hypothesis that germline mutations in *CDH1* would be found in a significant proportion of women with hereditary LBC, unselected for GC. I found that this was not the case and that the frequency of germline mutations in *CDH1* in women with early onset or hereditary LBC, unselected for GC, was very low (less than 2%). This finding did not discount the association of LBC and germline mutations in *CDH1*, but rather it informed us that families with multiple LBCs were likely rare. This study was performed using DHPLC to screen exons cheaply for heterozygous variants that could then be Sanger sequenced. I performed the DHPLC screen in half of the cases and sequenced the positive cases accordingly. Over the course of the study the cost of Sanger sequencing had become more economical, and therefore the other half of the samples were sent to the Genome Sciences Centre for sequencing.

Seeing as the finding of germline *CDH1* mutations in the selected cohort with LBC, were infrequent, there were few mutations to functionally validate. Therefore, I decided to move the focus of my thesis towards investigating other examples of cancer-related genotype-phenotype correlations. Intriguingly, there is overlap between the genes that cause Mendelian developmental disorders and those that predispose to cancer, with the most extreme examples being those that have both developmental defects or congenital abnormalities and a predisposition to cancer. Furthermore, there are a growing number of genes known to cause severe developmental disorders when disrupted in the germline that also become drivers of sporadic cancer following acquisition of somatic mutations. An example of a gene known to cause an MCA syndrome and also known to be recurrently somatically mutated in various hematologic malignancies is *PTPN11*. I had

the opportunity to investigate the multiple granular cell tumors (MGCT) that developed in a patient LEOPARD syndrome caused by a recurrent germline mutation of *PTPN11*. To test whether, *PTPN11* may function as a tumor suppressor in these multiple tumors, I looked for LOH of the WT allele. To ask whether the recurrent mutation occurred as a high frequency somatic mutation in association with sporadic granular cell tumors (GCT), I genotyped paraffin DNA from 10 sporadic granular cell tumor cases. I was only able to amplify a PCR product in six of 10 cases, due to poor quality of the extracted paraffin DNA. I did not detect the recurrent variant in any of the six cases. Another example of a gene that causes disease when altered in the germline or somatic scenario, is *FOXL2*. A recurrent mutation, c.402C>G, in *FOXL2* is associated with almost all adult-ovarian granulosa cell tumors. The specific somatic mutation is associated with a clear tissue-specific predilection for disease, although when *FOXL2* is non-specifically mutated in the germline, it causes a congenital anomaly syndrome. The final example of a genotype-phenotype correlation I studied was with regard to germline mutations in *BRCA1* and *BRCA2* and the association with high-grade serous ovarian carcinoma. Each of these examples, support the notion that there is utility in defining the genotype-phenotype correlations in the sporadic and hereditary cancer setting, as each paradigm of tumor development informs the other and establishing these correlations helps guide targeted genetic testing.

These initial studies were all performed with candidate gene sequencing used alone or in combination with DNA screening technologies, such as DHPLC or HRM, devised to lower the costs of mutation screening while maintaining high sensitivity for variant detection. Although these technologies are cheap and high throughput, as evidenced by the HRM analysis in this thesis, there is still the need for repeat assays and orthogonal validation. Newer massively parallel sequencing technologies have made sequencing orders of magnitude cheaper. Furthermore, the ability to barcode and multiplex samples allows for the sequencing of many samples at once, thus it is conceivable that newer sequencing technologies that can genotype in one step, will replace these multi-step mutation-screening methods. Following my work defining genotype-phenotype correlations with candidate genes, I had the opportunity to use next-generation

sequencing, to try and determine the underlying genotype that correlated with a distinct phenotype in a Mendelian disease gene discovery experiment.

Mendelian disorders are defined by single gene mutations that cause recognizable patterns of phenotypes or abnormalities, otherwise known as syndromes. Previously, identification of disease genes has been through a lengthy process of linkage analysis of multiple large pedigrees followed by candidate gene sequencing in regions of linkage. Recently, the advancements in sequencing technology have revolutionized the way in which Mendelian genes are being discovered [182]. This new technology allows for rapid identification of candidate variants though an unbiased approach. By comparing the entire genome or just the protein coding regions (exome) of individuals who share a phenotype, one can look for genomic elements that are consistently altered. In the case of exome sequencing, if one assumes the disease will likely be due to rare variants that alter protein coding, one can use bioinformatic tools to select genes that harbor rare variants in multiple individuals. These techniques were initially applied to discover the genetic susceptibilities in families with congenital disorders due to single gene mutations. The power of this technology was first demonstrated by the ability to rediscover the causative gene, *MYH3*, in Freeman Sheldon Syndrome by performing whole exome capture (WEC) and sequencing of four unrelated individuals with the syndrome and eight Hapmap individuals [183]. Subsequently, Miller Syndrome became the first Mendelian disorder to be explained by identification of the causative gene, *DHODH*, using similar methods [184, 185]. Since then, there have been an increasing number of reports of this technology being used for genetic diagnosis [186] and identification of disease genes in previously unresolved genetic syndromes [187]. These technologies have also been utilized in determining the underlying causation of hereditary cancer syndromes [188-201]. Furthermore, some Mendelian disorders have been found to be caused by germline mutations in genes more well-known to be somatically mutated in cancer [202].

We used exome sequencing of multiple related individuals to determine the genetic basis of the family's Mendelian disease. It is conceivable that we could have arrived at the underlying diagnosis in a cheaper fashion using only high-density SNP arrays to define

the region of linkage, in conjunction with exome sequencing of a single individual. However, as this was our first attempt at exome sequencing to determine an underlying Mendelian disease gene, we were not sure of how many false positive variant calls we would detect. To eliminate the potential effect of sample-specific false positives, we sequenced multiple affected individuals and required that candidate variants would need to comply with the apparent autosomal recessive Mendelian mode of inheritance in the family. This experiment served as a proof of principle to show that we were able to use next-generation sequencing in Mendelian disease gene discovery and provided the impetus to apply these technologies to cancer susceptibility syndromes.

Hereditary cancer syndromes behave like any other Mendelian disorder and show clear tissue-specific phenotypes that allow for the provision of clinical genetic testing based on the recognition of associated patterns of sites and types of cancer. Furthermore, the same genotype-phenotype correlations seen in the hereditary cancer setting are also echoed in the somatic mutation profiles of the corresponding sporadic cancers. Thus, the observation and identification of recurrent aberrations of particular genes and pathways in either the hereditary or sporadic cancer setting, shed light on the key biologic processes as they relate to particular tumor types. Cancer is a common disease where highly penetrant syndromes account for less than 5-10% of familial cancer. Recently, genome wide association studies of thousands of patients have also identified multiple loci associated with familial cancer, albeit with much lower risk. These studies do not immediately identify the causative genes [180, 181]. Identification of the genetic causes of highly penetrant autosomal dominant syndromes has enabled germline mutation testing of asymptomatic family members that has transformed the management of these syndromes. The impacts have been far reaching with regard to risk stratification, prophylactic surgery and tailored therapeutic strategies. It is evident that the identification of further disease genes that predispose to familial cancer would greatly benefit unaffected family members, however until now, this has not been technically feasible.

In my thesis I present the analytic strategy and early results of analysis of next-generation sequence data from a family with a newly defined autosomal dominant syndrome predisposing individuals to gastric adenocarcinoma and proximal polyposis of the

stomach (GAPPS). To conclude, I present the current and future applications of these technologies to the management of hereditary cancer and general oncology more broadly. I also present some of the caveats to genome-wide sequencing with regard to the potential for discovery of incidental findings and some of the related ethical issues with regard to the return of results.

My thesis began in a time when sequencing single genes was difficult and improving patient selection for specific gene tests by pathology and family data was of great importance. Typically, clinical Sanger sequencing of genes costs two to three thousand dollars per gene. Considering the extensive differential diagnosis for some disease presentations, as in the case of the family with spondyloepiphyseal dysplasia (SED) and retinitis pigmentosa (RP) described in my thesis, the costs of clinical genetic testing can balloon. Thus relative to the costs of clinical sequencing, exome sequencing is cheap and essentially functions as a multiplex assay. Having arrived at the next-generation sequencing era when sequencing is comparatively very cheap, the much greater issue now becomes how to deal with the data, the variant annotation and designation of disease-causing mutations.

CHAPTER 2: The Association Of Lobular Breast Cancer And Germline Mutations In *CDH1* In Early-Onset Or Familial Lobular Breast Cancer

In this chapter I have shown that lobular breast cancer is associated with hereditary diffuse gastric cancer by the demonstration of a *CDH1* germline mutation in a predominantly lobular breast cancer family. Therefore I sought to test the hypothesis that women selected for lobular breast cancer diagnosed under the age of 45 or with a family history of breast cancer harboured pathogenic germline mutations in *CDH1*.

2.1 The Association Of Lobular Breast Cancer And Germline Mutations In *CDH1*

2.1.1 Lobular Breast Cancer And Diffuse Gastric Cancer: Loss Of E-Cadherin

Currently mutations in single genes account for approximately 5-10% of breast cancer [166], where the high penetrance genes such as BRCA 1 and 2 account for 3-8%, and *TP53* and *PTEN* as seen in Li Fraumeni and Cowden syndrome together only account for <0.1% of breast cancer diagnoses [167]. Other low penetrance genes such as CHEK2 have been identified, however, there remains a proportion of hereditary breast cancer not yet explained. LBC accounts for 10% of all breast cancers compared to the other major histologic subtype, invasive ductal carcinoma (IDC). Several factors suggest that LBC has a stronger hereditary basis relative to IDC, such as the higher frequency of bilateral disease [162], and also excess familiality of LBC observed in population studies [203]. LBCs compose only 3% and 9% of the breast cancer tumor types seen in germline *BRCA1* and *BRCA2* mutation carriers, respectively [204], illustrating that the genetic risk factors for the majority of cases are unaccounted for by these genes.

The histology of LBC is characterized by infiltrative cancer cells that are isolated and highly dispersive in stromal tissue [101]. The pathologic appearance is remarkably similar to DGCs and both LBC and DCG demonstrate characteristic mucinous, signet ring cells. This is not unexpected as E-cadherin staining is absent in 85% of sporadic invasive LBC

[163] and somatic *CDH1* mutations have been identified in 56% of sporadic LBCs [164]. Furthermore, in IDC, somatic *CDH1* mutations are not found [164] and loss of E-cadherin expression is an uncommon feature. As loss of E-cadherin expression is a distinctive trait of both LBCs and DGCs, it likely contributes to the unique histopathologic features shared by the two cancers.

There are some differences with regard to the nature of the mutations seen in LBC and DGC. Mutations associated with sporadic LBC have generally been found to be nonsense or frameshift mutations [105] which encode truncated, non-functional proteins, whereas in sporadic DGC, mutations have generally been found to be splice site and in-frame mutations [68]. In sporadic LBC, mutations in *CDH1* are spread throughout the gene [68] compared with the mutations seen in sporadic DGC that tend to cluster in exons 7-9. Compared with germline mutations found in HDGC families without a history of LBC, germline mutations associated with LBC also appear to occur throughout the gene (Figure 4). Another difference between the molecular genetics of the two types of cancers is that in sporadic LBC, silencing of E-cadherin expression is generally accomplished by a mutation in one allele in combination with LOH or promoter hypermethylation in the remaining allele [165]. This is in contrast to sporadic DGC, where biallellic inactivation is achieved by mutations in one allele in concert with promoter hypermethylation in the other [205]. In keeping with the genotype-phenotype correlations seen in sporadic LBC, we recently identified a truncating germline *CDH1* mutation in an LBC family where analysis of the tumor was suggestive of partial LOH in the WT allele [159]. Herein we describe a predominantly breast cancer family segregating a germline mutation in *CDH1*, predicted to disrupt splicing and similar to previously reported HDGC-associated mutation in a family uniformly affected with GC. Moreover, a previous study reported a germline missense mutation in a proband with LBC although did not detail family history, or functionally characterize the missense mutation [206]. These differing examples demonstrate the need for further studies of germline mutations in LBC families in order to determine the genotype-phenotype correlations in this subset.

Figure 4 HDGC-associated CDH1 germline mutations.

Mutations shown above *CDH1* gene schematic occur in families with DGC history and those below *CDH1* occur in families with additional LBC history. In addition to the known *CDH1* germline mutations compiled by Kaurah et al., the recent mutation in an LBC family and novel mutation from this thesis are shown and identified below the symbol denoting mutation type. Reproduced from [Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer., Schrader KA, Masciari S, Boyd N et al, 7, 73-82, 2008] with permission from Springer.
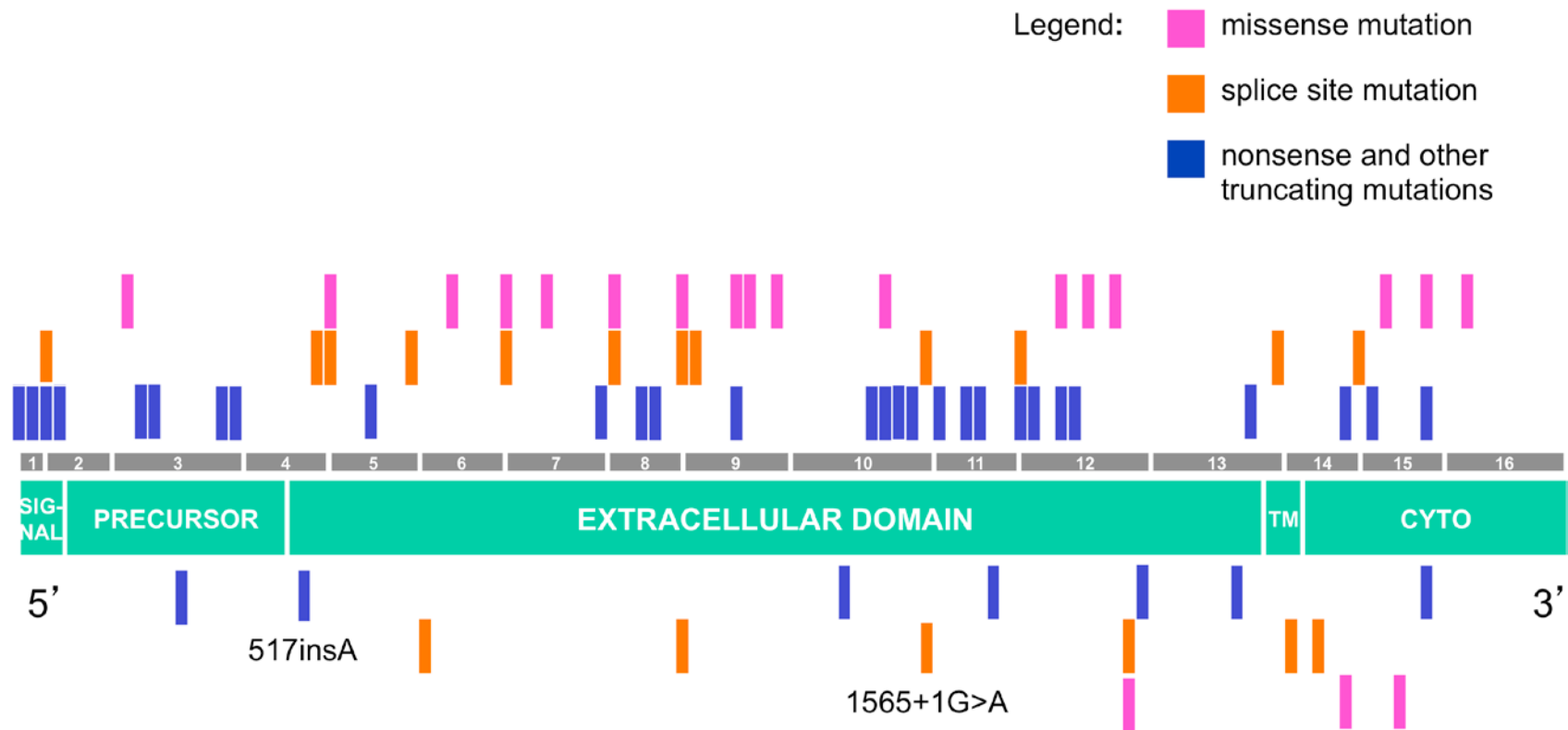
2.1.2 Lobular Breast Cancer And HDGC

Breast cancer has been observed in HDGC kindreds. Occasionally clustering of LBC cases within HDGC families has led to the misclassification of the families as *BRCA1*- and *BRCA2*- mutation negative breast cancer kindreds [94]. In 1998, Keller described the first case of histologically defined LBC in association with HDGC [169]. Since then, several more HDGC families with associated breast cancer were reported where it was observed, that these cases were LBCs when pathology was available [4, 90, 169, 207].

Prior to establishment of the association between HDGC and LBC, several efforts to determine whether *CDH1* was a breast cancer susceptibility gene were attempted in view of the well-recognized phenotype of loss of E-cadherin expression displayed by the breast cancer subtype. For various reasons these studies failed to demonstrate the link. Rahman et al. examined 65 cases of lobular carcinoma *in situ*, however did not pre-screen the cases based on family history and included a wide age range, from 26-71 years, not necessarily in keeping with the usual age of onset seen in hereditary cancer syndromes[208]. Salashor examined 19 breast cancer cases showing LOH at the *CDH1* locus, however of those, only 3 were confirmed to be only LBC or mixed pathology [209]. Lei examined 13 familial LBC cases, however did not define the extent of the family history [210].

Penetrance data based on 11 families, estimated the cumulative risk for LBC for females in HDGC families to be 39% (95% CI, 12%-84%) by 80 years of age [131]. More recently we have published an estimated cumulative risk for breast cancer for females by the age of 75 years as being 52% (95% CI, 29%-94%) from analysis of 4 predominantly GC pedigrees from Newfoundland with the 2398delC *CDH1* founder mutation [94]. This is with the caveat that LBC risk for *CDH1* mutation carriers has been assessed within high-risk HDGC families, leading to a potential ascertainment bias and underestimation of the role of *CDH1* mutations in LBC development. To accommodate for this we have begun analysis of *CDH1* mutations within familial LBC families or those families ascertained

through a relatively young index case with confirmed LBC and have found germline *CDH1* mutations in these kindreds [159].

Our group has reported a novel germline *CDH1* truncating mutation (517insA) in an LBC family with no known history of GC [159]. Within this thesis we report a germline *CDH1* mutation in a second family in which breast cancer is the predominant cancer diagnosis. The management of HDGC in all patients with a particular focus on the management of the breast cancer risk associated with germline *CDH1* mutations will be discussed.

2.1.3 Case Report

Presented herein was the second *CDH1* mutation-positive family to be reported in which the predominant cancer was lobular breast cancer (Figure 5).

Reproduced from [Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer., Schrader KA, Masciari S, Boyd N et al, 7, 73-82, 2008] with permission from Springer.
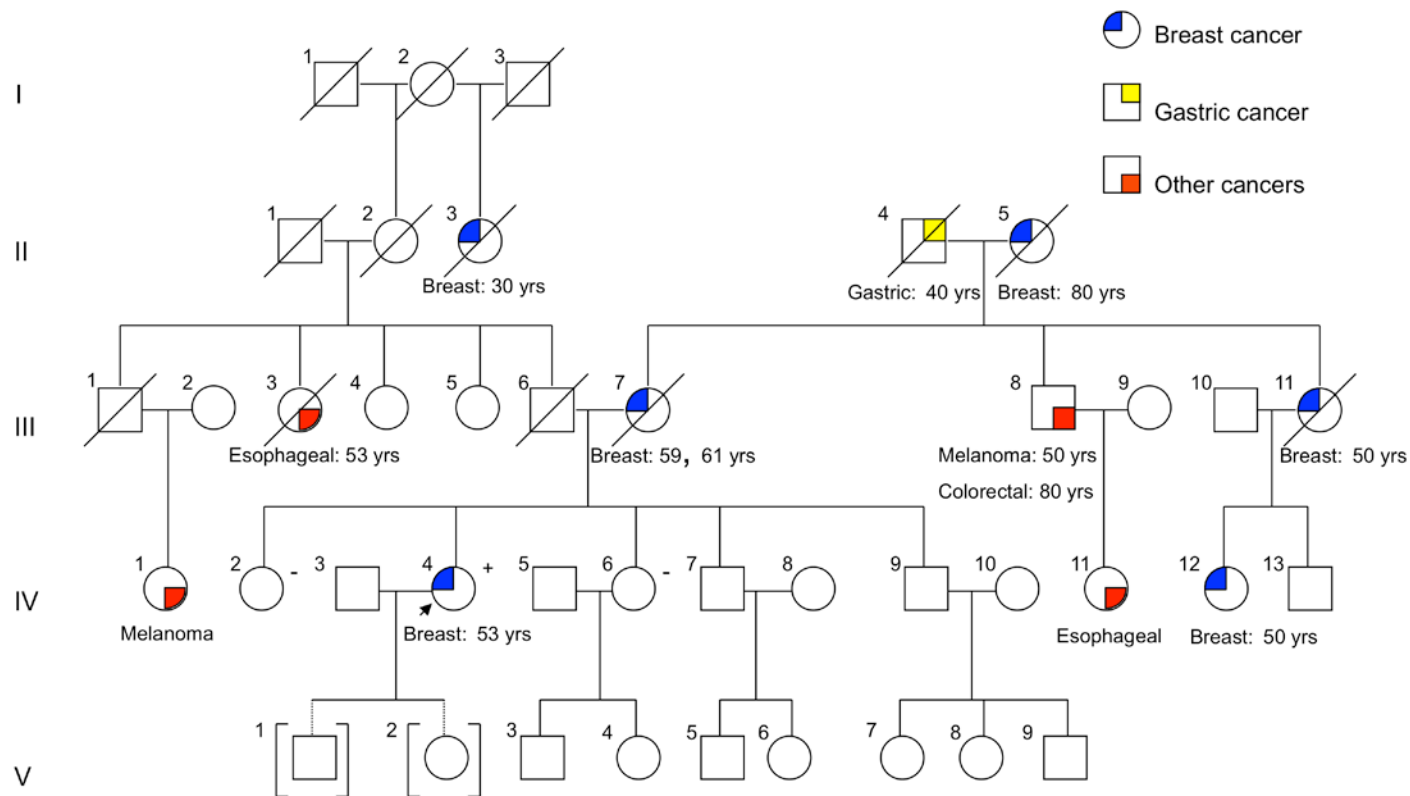
2.1.3.1 Methods

The described family was referred to the ongoing HDGC study at the British Columbia Cancer Agency from a cancer genetics clinic in Seattle, WA, USA. Informed consent was obtained from the proband by the referring genetic counsellor following ascertainment of a detailed cancer family history and appropriate genetic counselling prior to germline mutation testing. Approval for the HDGC study is by the clinical research ethics board of the University of British Columbia.

The proband (IV-4) was diagnosed with widely metastatic LBC at age 53 years (Figure 5). Her family, of European ancestry, had a history of breast cancer diagnoses occurring in an autosomal dominant fashion on the maternal side of the family where her mother, aunt, and first cousin developed breast cancer in their 50's. Due to her high-risk pedigree *BRCA1*, *BRCA2*, and *PTEN* genetic testing was undertaken in clinical testing labaoratories and all were negative. *CDH1* testing was also pursued. Our laboratory carried out molecular genetic testing for *CDH1* on a research basis. This consisted of amplification of all 16 exons of *CDH1* for DHPLC analysis and bidirectional Sanger sequencing of the amplicons identified with traces that differed from the homozygous WT *CDH1* amplicon [4].

2.1.3.2 Results

For exon 10 of *CDH1*, the initial amplicon failed and was therefore analyzed by direct sequencing and thus revealed a heterozygous donor splice site mutation, IVS 10+1 G>A (Figure 6). The mutation (1565+1G>A) is in the same conserved position as a previously reported heterozygous mutation (1565+1G>T) which was found in an Arabian HDGC family with no recorded history of breast cancer [141]. Due to its position at a donor splice site, this mutation is regarded as pathogenic [211].

Figure 6 Sequence from family carrying c.1565+1G>A CDH1 germline mutation.

Reproduced from [Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer., Schrader KA, Masciari S, Boyd N et al, 7, 73-82, 2008] with permission from Springer.
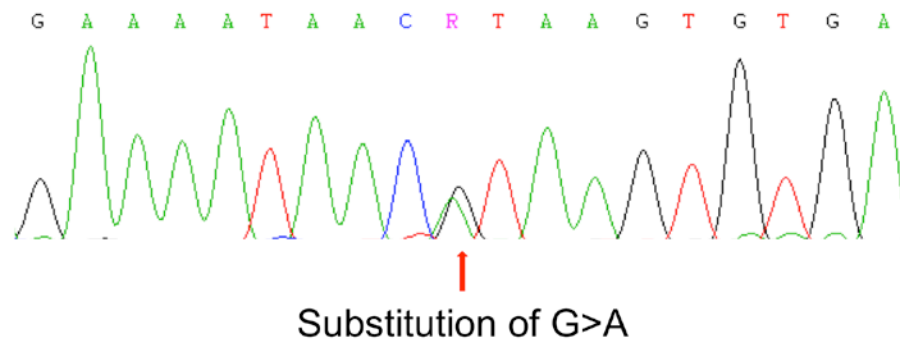
The proband's medical oncologist appropriately advised the patient not to worry about her GC risk for the time being but to focus on her treatment for breast cancer. The proband's sisters (IV-2 and IV-6) participated in all aspects of the proband's genetic consultation. They were appropriately concerned about their risk of breast cancer, but had not thought much about the possibility of getting GC until the *CDH1* mutation was found. IV-2 and IV-6 had predictive genetic testing for the *CDH1* mutation and both were found to be negative. Other family members are being informed about the availability of predictive genetic testing.

2.1.4 Clinical Implications Of *CDH1* Associated LBC Risk

At this time, it seems reasonable to conclude that at least four groups of women can potentially be at risk for future LBC in particular: women with LBC and a family history of breast cancer, women with a known *CDH1* mutation, women from families with diffuse GC in whom no *CDH1* mutation has yet been identified; and women with a germline *BRCA2* mutation. Since there has not yet been a large population based study of the prevalence of *CDH1* mutations among women with LBC, it is premature to recommend genetic evaluation to women with a family history of breast cancer unless, at the very least, one of the breast cancers can be shown to have been lobular. Additional research can be expected to provide better guidance for these families. Guidelines for the management of the LBC risk are presented earlier in this thesis (section 1.1.1.6.4).

2.1.5 Conclusion

We conclude that *CDH1*-associated LBC should be a new entry into the growing plethora of breast cancer prone syndromes, which clearly reflect the profound genotypic and phenotypic heterogeneity of breast cancer (Table 1). The lack of shared genetic risks for most breast and GI cancers was demonstrated through a recent study of 13,023 genes in 11 breast and 11 colon cancer cell lines in which the only commonly mutated gene

between these two cancer types is *TP53* [213]. This likely reflects underlying differences in the biology of these diseases, however also highlights the unique nature of germline mutations in the *CDH1* gene which are strongly associated with specific histologically defined subtypes of breast and GI cancer, namely LBC and DGC which are both part of the HDGC syndrome.

Table 1 Other syndromes with familial susceptibility to breast and gastric cancers.

AD = autosomal dominant, AR=autosomal recessive. Reproduced from [Hereditary Diffuse Gastric Cancer: Association with Lobular Breast Cancer., Schrader KA, Masciari S, Boyd N et al, 7, 73-82, 2008] with permission from Springer.

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
BRCA2 Hereditary Breast/Ovarian Cancer	AD	*BRCA2*	breast	Metanalysis of 22 studies	BC is considered an integral tumor of the syndrome with an average cumulative risk in carriers by age 70 years of 45% (95% confidence interval (CI) 33%-54%) [214].
			ovary		
			larynx	Case series of 29 *BRCA2*-families	Adjusted standard incidence rate of 576.8 (95% CI 473.7-702.2) for breast cancer (BC) and 597.7 (95% CI 449.7-794.4) for GC [215].
			prostate		

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
			pancreas	Cohort study of 3728 individuals from 173 breast-ovarian cancer families with *BRCA2* mutations	Relative risk (RR) for GC was 2.59; 95% CI = 1.46-4.61)[127].
				Case series of 35 Ashkenazi Jewish GC patients	6174delT mutation was found in 2/35 cases (5.7%, p=0.06; OR 5.2 95%CI 1.2-22)[216].
				Population-based series of 649 unselected incident cases of ovarian cancer screened for germline mutations in *BRCA1* and *BRCA2*.	GC incidence in relatves of cases with *BRCA2* mutations was 4.9%, relative risk 6.2, 95% CI 2.0-19 [217].

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
BRCA1 Hereditary Breast/Ovarian Cancer	AD	*BRCA1*	breast ovary prostate	Clinic-ascertained case series of 483 *BRCA1* mutation carriers in 147 families	BC is considered an integral tumor of the syndrome with an average cumulative risk in carriers by age 70 years of 72.8% (95% confidence interval [CI] = 67.9% to 77.7%) [126]. Cumulative age-adjusted lifetime risk of gastric cancer was 5.5% (95% CI = 3.4% to 7.5%) [126].
Peutz-Jeghers Syndrome	AD	*STK11*	gastrointestinal (GI) tract	Metanalysis of 6 studies	GC is considered an integral tumor of the syndrome with a RR of 213 (95% confidence interval 96-368) [121]. BC is considered an integral tumor of the syndrome with a relative risk of 15.2 (95% CI 7.6-27) [121].
Cowden Syndrome	AD	*PTEN*	breast thyroid endometrium	Case series / Review Case report	BC is considered an integral tumor of the syndrome with an incidence of 22-50% [218, 219]. GC in situ has been reported in a patient with Cowden

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
					Syndrome [220].
Li-Fraumeni Syndrome	AD	*p53* *CHK2*	breast adrenal cortex connective tissue kidney nervous system pancreas white blood cells	Case series Review Case reports/ Case series Meta-analysis of 10 studies	BC is frequently found in families with this cancer susceptibility syndrome [221]. Chompret expanded the spectrum of cancers to include GC [222]. Germline *TP53* mutations have been found in GC families without *CDH1* mutations [122, 123, 223]. The 1100delC *CHEK2* allele estimated odds ratio 2.34; 95% CI 1.72–3.20; P=.0000001 [224], however germline mutations in GC kindreds have not been identified [225].
Familial Adenomatous Polyposis	AD	*APC*	colon and rectum duodenum	Case report/ review	Literature review by Shimoyama et al. totalled 30 reported cases of GC and FAP in the literature [125].

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
			thyroid	Case series	47% to 49% of primary BCs had promoter hypermethylation at the *APC* locus [226, 227].
			pancreas	Case series	23% of LBCs have been shown to have LOH of *APC* [143].
Lynch Syndrome (Hereditary Nonpolyposis Colon Cancer (HNPCC))	AD	*MSH2* *MLH1* *MSH6* *PMS1* *PMS2*	colon and rectum	Case series	GC accounted for 5% of cancers in families harboring *MLH1* or *MSH2* mutations. Familial RR with *MSH2* mutations, 2.7, P = 0.050 [117].
			endometri um		
			stomach	Case report	*MLH1* mutations in large kindred segregated with BCs exhibiting MSI [228].
			small intestine		
			urotheliu m	Case series	A slight increased incidence of BC was seen in *MLH1* mutation carriers [229].
			kidney		
			ovary	Case report	Germline *MSH2* mutation carrier with BC exhibited LOH for *MSH2* in tumors analyzed [230].
			skin		

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
			pancreas	Case series	Analysis of primary invasive BCs demonstrated that 25% of tumors were immunonegative for MSH2 staining [231].
			brain white blood cells biliary tract		
Ataxia-telangiectasia (AT)	AR	*ATM*	white blood cells	Review	Mutations causing AT in homozygotes, confer susceptibility to BC in heterozygotes, where women with *ATM* mutations have a ~2-fold risk of BC and ~15% of these women will develop the disease [232].
				Case report/ Case series	GC has been reported in association with the syndrome [233-235].

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
				Case series	There is evidence of excess risks of GC in heterozygotes (RR = 3.39, 95% CI = 0.86 to 13.4) [236].
Xeroderma pigmentosum	AR	*XPA, ERCC3(XPB) XPC ERCC2(XPD) DDB2(XPE) ERCC4(XPF) ERCC5(XPG) POLH (XP-V)*	skin eyes	Case reports	BC and GC have both independently been reported with the syndrome [237, 238].
Werner Syndrome	AR	*WRN*	connective tissue skin thyroid	Case report/ review Case contol	GC has been reported in association with the syndrome [239]. There are no reports of BC in association with Werner syndrome. Although, there is evidence supporting *WRN* as a low-penetrance familial BC susceptibility gene, *WRN* Cys1367Arg polymorphism with familial

Syndrome	Mode of inheritance	Associated gene(s)	Sites of primary cancer(s)	Type of study	Evidence for association with the syndrome
					breast cancer (OR = 1.28, 95% CI 1.06-1.54) and high-risk familial breast cancer (OR = 1.32, 95% CI 1.06-1.65) [240].

With the recent demonstration of a *CDH1* mutations in a family ascertained through an index case of LBC and in view of the additional new mutation in a predominantly breast cancer family that we have described here, the evidence for establishing LBC as part of the HDGC syndrome is strong. There now is a need for establishing the prevalence of *CDH1* mutations in LBC families to avoid the ascertainment bias generated from only looking at cases from families identified because of their family history of GC. It is not currently known what the risk of GC is in these families which present predominantly as having a susceptibility to breast cancer and therefore identification of *CDH1* as a true susceptibility gene for LBC could result in *CDH1* screening and effective risk reduction strategies for selected breast cancer families and further studies examining their risk for gastric and other cancers.

Most hereditary cancer syndromes are associated with cancer risk involving multiple organs. Here we have discussed germline *CDH1* mutations and the risks with regard to DGC and LBC, however as the recognized spectrum of related cancers broadens, more affected families will be identified and successfully managed with regard to avoidance of

specific cancer risks. Longer life expectancy in individuals with penetrant mutations could potentially lead to the development of different, later onset disease as yet to be identified in these kindreds. This represents a particular challenge in hereditary cancer practice as the clinical community is segregated into organ specific specialties and will require that a multi-disciplinary approach be employed for the management of the high-risk patients in these families where the *CDH1* mutation has been identified.

2.2 The Frequency Of Germline Variants In *CDH1* In Early-Onset Or Familial Lobular Breast Cancer

The following experiment was undertaken to determine the frequency of germline mutations in *CDH1* in individuals with LBC. This experiment has since been replicated to some extent, whereby the authors came to similar conclusions [161]. Of note an individual with a germline deletion was identified as outlined in the letter by Newman et al. responding to the initial article [161]. Furthermore, a germline 16q22.1 large deletion of *CDH1* was identified in a proband with psychomotor delay and dysmorphic features, inherited from his mother diagnosed with invasive LBC at age 35 years [241].

2.2.1 Introduction

Previously, we identified one carrier of a germline truncating *CDH1* mutation among 23 women with LBC known not to carry germline *BRCA1* and *BRCA2* mutations [159]. This case series included women diagnosed at a young age (<45 years) and women diagnosed at any age with a family history of breast cancer but without a family history of GC (1/23 or 4.3%)[159]. The same mutation was subsequently confirmed in a relative of the mutation carrier who also had LBC. This coincidence of *CDH1* mutations and hereditary LBC led us to assess the prevalence of *CDH1* mutations in a larger series of women with early-onset LBC or a family history of breast cancer, consistent with hereditary LBC.

2.2.2 Materials And Methods

2.2.2.1 *Patient Accrual*

Three-hundred and twenty-seven LBC cases were identified through three different sources. The Breast Cancer Family Registry (Breast CFR) is an National Cancer Institute (NCI)-sponsored resource which includes six population- based and clinic-based family registries and a collection of samples and data from more than 12,500 families with and without breast cancer [242]. Samples (n=168) were obtained from the Northern California, New York, Australia, Philadelphia, and Ontario sites of the Breast CFR. The Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer (kConFab) provided specimens (n=33) from families with a strong history of breast cancer, recruited from family cancer clinics in Australia and New Zealand [243]. The 126 remaining samples were collected through a Breast Cancer Research Foundation (BCRF)-funded Breast Cancer Genetics Consortium, a group of high-risk cancer clinics which included Dana-Farber Cancer Institute, Baylor College of Medicine, Beth Israel Deaconess Medical Center, Georgetown University, Massachusetts General Hospital, Memorial Sloan Kettering Cancer Center, Stanford University, the University of Chicago, the University of Pennsylvania, and the Hospital Vall d'Hebron, Barcelona, Spain. The inclusion criteria for the identification of the eligible LBC cases differed slightly between these groups. For cases from the Breast CFR and kConFab, eligibility for this study required a female case with documented invasive lobular or mixed (lobular and ductal) breast cancer, not known to carry germline *BRCA1* and *BRCA2* mutations, and either: (1) diagnosed before age 45 years, or (2) at any age but with two or more cases of breast cancer in first- or second-degree relatives. For cases from the BCRF-funded Breast Cancer Genetics Consortium, eligible women had a diagnosis of invasive lobular or mixed (lobular and ductal) breast cancer, were not known to carry germline *BRCA1* and *BRCA2* mutations, and were either: (1) diagnosed before age 45 years, or (2) diagnosed at any age but with at least two or more cases of breast cancer in first- or second-degree relatives, and third-degree relatives in the paternal lineage, and with no reported family history of GC. The characteristics of the LBC cases screened for mutations in *CDH1* are summarized in Table 2. All cases had provided written informed consent and the study protocols were approved by the institutional review board at each participating center. DNA was

extracted at the molecular laboratories for some of the collaborating centers using standard procedures (Qiamp DNA Blood Midi kit; Qiagen, Valencia, CA, USA) and anonymized genomic DNA samples were sent to the Centre for Translational and Applied Genomics (CTAG) at the British Columbia Cancer Agency (BCCA) where the analysis of *CDH1* was performed.

Table 2 Criteria for ascertainment for CDH1 mutation analysis.

The criteria for ascertainment were (1) a patient with a history of lobular or mixed ductal and lobular pathology whose *BRCA1* and *BRCA2* mutation status was negative or unknown and either diagnosed before age 45 years or (2) diagnosed at any age but with two or more cases of breast cancer in first- or second-degree relatives. Reproduced from [Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers., Schrader KA, Masciari S, Boyd N et al, 48, 64-68, 2011] with permission from BMJ Publishing Group Ltd.

Breast Cancer Genetics Consortium							
Criteria for ascertainment	Patients (n=120)	Sex	Median age (years)	Age range (years)	Novel non-synonymous variants in criteria	BRCA1/2 negative (n=66)	BRCA1/2 unknown (n=54)
Criteria 1	66	F	40	(28–44) (n=66)	2	36	34
Criteria 2	54	F	54	(45–72) (n=54)	1	30	20
Breast CFR							
Criteria for ascertainment	Patients (n=165)	Sex	Median age (years)	Age range (years)	Novel non-synonymous variants in criteria	BRCA1/2 negative (n=147)	BRCA1/2 unknown (n=18)
Criteria 1	142	F	40	(31–44) (n=51)	3	130	12
Criteria 2	23	F	57	(45–79) (n=20)	0	17	6
kConFab							
Criteria for ascertainment	Patients (n=33)	Sex	Median age (years)	Age range (years)	Novel non-synonymous variants in criteria	BRCA1/2 negative (n=33)	BRCA1/2 unknown
Criteria 1	6	F	40	(37–43) (n=6)	0	6	0
Criteria 2	27	F	57	(45–77) (n=28)	0	27	0
All samples							
Criteria for ascertainment	Patients (n=318)	Sex	Median age (years)	Age range (years)	Novel non-synonymous variants in criteria	BRCA1/2 negative (n=246)	BRCA1/2 unknown (n=72)
Criteria 1	214	F	40	(28–44) (n=123)	5	172	46
Criteria 2	104	F	56	(45–79) (n=101)	1	74	26

2.2.2.2 *Preparation Of DNA And CDH1 Sequencing*

To accommodate the limited amount of DNA available, genomic DNA samples were subjected to whole-genome amplification using the GenomiPhi DNA amplification kit (GE Healthcare Bio-Sciences Inc., Quebec, Canada) as performed in our earlier study [159]. Exons and intron-exon boundary splice junctions of half the study cohort were amplified and screened for heterozygous base changes by the denaturing high pressure liquid chromatography (DHPLC) [159]. The primer sequences and conditions used have previously been described [4]. Exons displaying DHPLC changes consistent with a heterozygous variation were reamplified and PCR products were purified (Qiagen MinElute; Qiagen, Mississauga, ON). Bidirectional sequencing was then performed (Big Dye Terminator V.3.1 Cycle Sequencing Kit; Applied Biosystems, Foster City, California, USA) and analyzed (ABI Prism 310 Genetic Analyzer). Germline mutations in *CDH1* appear as heterozygous sequence changes, with the exception of large deletions, which are not detected by sequencing. Sequencing of all exons and intron-exon boundaries were carried out on the remaining samples using validated primer sets [244], at the Genome Science Centre on a service basis or in our laboratory. DHPLC is highly

2.2.2.3 *Deletion Analysis*

Multiplex ligation-dependent probe amplification, previously used to identify large-scale deletions in *CDH1*, has been described [91]. MLPA could only be performed on 134 samples for which sufficient germline DNA was available.

2.2.2.4 *Mutation Validation*

Basic Local Alignment Search Tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to compare sample sequences to the NCBI cDNA NM_004360.3 and the genomic DNA NG_008021.1, GI:190341080. Web-based software programs were used to look for predicted effects on splicing (Splice Site Prediction by Neural Network hosted by the Berkeley Drosophila Genome Project) and the predicted effects of amino acid changes on

protein structure (Sorting Intolerant From Tolerant, SIFT, software version 2, Fred Hutchinson Cancer Research Center, Seattle, Wash).

2.2.2.5 *Protein Structure Analysis*

Protein structure analysis and preparation of structural models was performed using the PyMol software package (DeLano Scientific). Structural co-ordinates were obtained from X-ray crystal structures of the ectodomain of C-cadherin (PDB code 1L3W) [245].

2.2.2.6 *Functional Characterization*

2.2.2.6.1 *Site Directed Mutagenesis*

To examine the effects of the missense mutations on protein localization, WT E-cadherin-WTpcDNA3.1 plasmids [246] were mutated using QuickChange site-directed mutagenesis kit (Stratagene, Cedar Creek, Texas) as per the manufacturer's instructions to create each of the novel non-synonymous variants we identified. The corresponding forward primer sequences are included following each variant: c.1223C>T, A408V, 5'-CCCCAATACCCCAGTGTGGGAGGCTGTAT-3'; c.8C>G, P3R, 5'-CTTACCATGGGCCGTTGGAGCCGCAGC-3'; c.88 C>A, P30T, 5'-GGAGCCCTGCCACACTGGCTTTGACGC-3'; c.1813A>G, 5'-ACTATATTCTTCTGTGAGGGGAATCCAAAGCCTCAGG -3'; c.1297G>A, D433N, 5'-CCACAAATCCAGTGAACAACAATGGCATTTTGAAAACAGCA-3'). The mutated plasmids were validated by DNA-sequencing.

2.2.2.6.2 *Cell Culture*

The immortalized ovarian surface epithelial cell line, IOSE-80pc, that does not express E-cadherin [247] and the ovarian carcinoma cell line, OVCAR-3, that highly expresses E-cadherin [247], were maintained in a 1:1 mixture of MCDB 105 medium and Medium 199

(Sigma, St. Louis, MO, USA), supplemented with 5% fetal bovine serum (FBS) (Gibco BRL).

2.2.2.6.3 *Transient Transfections*

Aliquots of IOSE-80pc cells were transfected in parallel with the E-cadherin variants under investigation; c.1223C>T, A408V; c.8C>G, P3R; c.88 C>A, P30T; c.1813A>G, R695G; c.1297G>A, D433N and control plasmids. These controls included the empty vector LacZ as a control for the transfection procedure, wild-type (WT) E-cadherin and known loss of function E-cadherin mutants (c.1018A>G and c.2494G>A) that cause HDGC. Transient transfections were performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's guidelines.

2.2.2.6.4 *Fluorescence Microscopy*

E-cadherin staining was performed on subconfluent cell monolayers cultured on glass coverslips. Cells were washed once in media alone and then fixed in 4% paraformaldehyde (PFA) for 20 minutes at room temperature. Cells were washed with phosphate-buffered saline (PBS) before and after permeabilization with 0.2% Triton X-100 for 10min at room temperature. To decrease non-specific background, cells were incubated with serum-free protein block (Dako, Carpinteria, CA) for 30 minutes prior to incubation with a mouse monoclonal antibody [1:500] to the extracellular domain (HECD-1, 205601 Calbiochem) in PBS at room temperature. Primary antibody was detected using Alexa Fluor 488 goat anti-mouse IgG (Invitrogen, Carlsbad, CA) [1:500] for 60 minutes at room temperature. Cells were counterstained with DAPI, 4',6-diamidino-2-phenylindole dihydrochloride (DAPI, Sigma-Aldrich, St. Louis, MO) [1:1000] prior to mounting on microscope slides with glycerol. Slides were analyzed using the Axioplan 2, Zeiss (MetaSystems, Isis) camera VAC-30054 and pictures were obtained at 40x magnification.

2.2.3 Results

2.2.3.1 *Sequencing*

Germline DNA from 327 eligible LBC cases was analyzed for variants in *CDH1*, but for 9 samples several exons failed to amplify, yielding incomplete results. Sequence analysis for heterozygous variants in the 318 cases with complete results did not detect any protein-truncating mutations. MLPA analyses in 134 cases did not reveal any large deletions in *CDH1*.

We did find 10 cases with non-synonymous variants. One non-synonymous change, c.1774G>A p.A592T, was found in two cases and is a known germline variant that is not associated with risk of familial breast cancer or HDGC [123, 248]. The variant, c.2494G>A, p.V832M, which had previously been identified in a case with HDGC and was functionally characterized as a pathogenic mutation [140, 249], was found in a female who was diagnosed with LBC at age 43 years and had a family history of ductal breast cancer in a sister and unspecified breast cancer in a maternal aunt. Segregation analysis has not yet been performed. The remaining non-synonymous variants were novel and did not appear in any public databases. These variants were: c.8C>G, p.P3R; c.1223C>T, p.A408V; c.1297G>A, p.D433N; c.1813A>G, p.R605G and c.88 C>A, p.P30T, which was found in two cases not known to be related. There was no family history of GC in any of the cases who carried novel non-synonymous variants (Table 3).

There was no known family history of GC in these patients. Criteria for ascertainment for *CDH1* mutation analysis. The criteria for ascertainment were (1) a patient with a history of lobular or mixed ductal and lobular pathology whose *BRCA1* and *BRCA2* mutation status was negative or unknown and either diagnosed before age 45 years or (2) diagnosed at any age but with two or more cases of breast cancer in first- or second-degree relatives. Reproduced from [Germline mutations in *CDH1* are infrequent in women with early-onset or familial lobular breast cancers., Schrader KA, Masciari S, Boyd N et al, 48, 64-68, 2011] with permission from BMJ Publishing Group Ltd.

Non-synonymous variant	Criteria 1 or 2	BRCA1/2 mutation status	Age at diagnosis	Family history (age at diagnosis)
c.8C→G, p.P3R	1	Negative	38 years	Maternal aunt=breast cancer (46 years) Maternal aunt=breast cancer (67 years) Maternal cousin=breast cancer (42 years) Mother=retroperitoneal tumour Paternal grandmother= breast cancer
c.88 C→A, p.P30T (two patients)	1	Unknown	40 years	
	2	Negative	47 years	Paternal aunt=breast cancer (40 years) Female paternal cousin=breast cancer (40 years) Male paternal cousin=breast cancer (50 years) Female paternal cousin=breast cancer (47 years)
c.1223C→T, p.A408V	1	Negative	44 years	No cancers
c.1297G→A, p.D433N	1	Negative	41 years	Paternal grandmother=intestinal cancer Maternal grandmother= lung cancer Maternal grandfather= mouth cancer
c.1813A→G, p.R605G	1	Unknown	42 years	Mother=breast cancer (60 years) Maternal uncle=pancreatic cancer (64 years)
c. 2494G→A, p.V832M (known missense mutation in HDGC)	1	Negative	43 years	Sister=ductal breast cancer Maternal aunt=breast cancer Paternal uncle=leukaemia Paternal grandmother=colon cancer

Nine unreported novel silent changes were identified: five synonymous variants in exons and four variants in introns. Two of these novel changes were found in more than one case (data not shown).

We performed several tests to assess the likelihood that any of the non-synonymous variants resulted in a loss of normal function. Web-based software (SIFT) that predicts whether the amino acid change conferred by non-synonymous variants might alter protein structure, and thus function, indicated that all but one variant c.8C>G, p.P3R, which occurred in the signal peptide of the pre-protein, should be tolerated and therefore is unlikely to be pathogenic. Moreover, web-based software (BDGP) did not predict alteration of splicing by any of the novel synonymous or non-synonymous variants or intronic variants identified.

The likely pathogenicities of the novel non-synonymous variants were further assessed by analyzing the predicted effects of amino acid changes on the three dimensional structure of E-cadherin. As the coordinates of the three dimensional structure of the ectodomain of E-cadherin were not available, we used the model of the closely related paralog, C-Cadherin, to predict likely changes in the structure. One of the mutations, c.1223C>T, p.A408V, changes the alanine residue, which is well-conserved in this family of proteins to bulkier valine, and is located in the calcium ion binding extracellular domain 3. Surface modeling of the mutated protein indicated that this bulky valine could conceivably alter the binding pocket of one of three calcium ions that mediate homotypic cadherin domain interactions (Figure 7). Another mutation, c.1297G>A, p.D433N, was also found to be located in close proximity to this calcium-binding site (Figure 7). As the c.8C>G, p.P3R variant occurs in the signal peptide of the precursor protein and had been predicted to be pathogenic, we hypothesized that this variant could result in mislocalization or lack of expression of E-cadherin on the cell-surface. To test this hypothesis, we expressed normal E-cadherin or each of the mutated versions of the

protein, in cells lacking endogenous E-cadherin. As seen in Figure 8, E-cadherin mutated with the c.8C>G, p.P3R variant did exhibit membrane localization, indicating that protein localization was not grossly affected by this variant. Additionally, the other novel non-synonymous variants also demonstrated normal membrane localization (data not shown). However, as the levels at which we expressed E-cadherin were not physiological, it is possible that subtle effects of the mutations could have been missed.

Figure 7 Proximity of two variants to a key calcium-binding site in the extracellular domain of E-cadherin.

Protein modeling of extracellular domain 2 (EC2) and extracellular domain 3 (EC3) based on crystal structure of C-cadherin. Positions of c.1223C>T, A408V (A254) and c.1297G>A, D433N (D279). Note the proximity of the amino acid changes to a calcium-binding site. Numbering refers to the amino acid position in mature protein (i.e. following cleavage of pro-peptide 1-154). (PDB ID 1L3W). Reproduced from [Germline mutations in *CDH1* are infrequent in women with early-onset or familial lobular breast cancers., Schrader KA, Masciari S, Boyd N et al, 48, 64-68, 2011] with permission from BMJ Publishing Group Ltd.
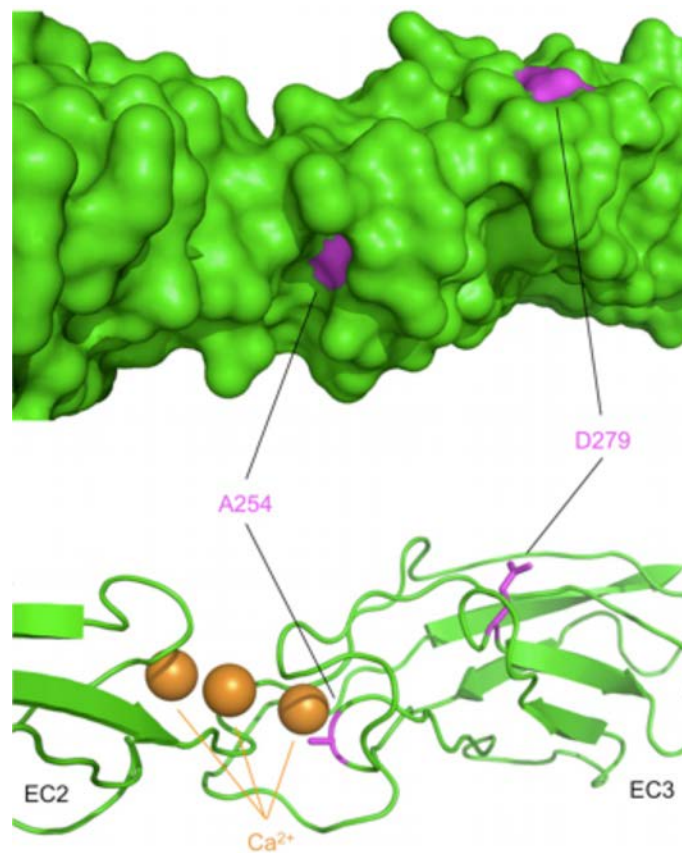
Figure 8 Immunofluorescent staining of E-cadherin mutated with the c.8C>G, P3R variant shows localization to the cell-membrane.

Immunofluorescent staining of E-cadherin shows cell-membrane localization in cells expressing endogenous E-cadherin or following transient transfection with the WT and mutant, c.8C>G, P3R, pcDNA3 *CDH1* constructs. (A) Untransfected parental IOSE-80PC cells, negative for endogenous E-cadherin. (B) OVCAR3 is an ovarian carcinoma cell line that over-expresses endogenous E-cadherin. Panels C and D show E-cadherin expression at day 3 post-transient transfection of IOSE-80PC cells with pcDNA3 *CDH1*: WT (C); c.8C>G, P3R (D). Reproduced from [Germline mutations in *CDH1* are infrequent in women with early-onset or familial lobular breast cancers, Schrader KA, Masciari S, Boyd N et al, 48, 64-68, 2011] with permission from BMJ Publishing Group Ltd.
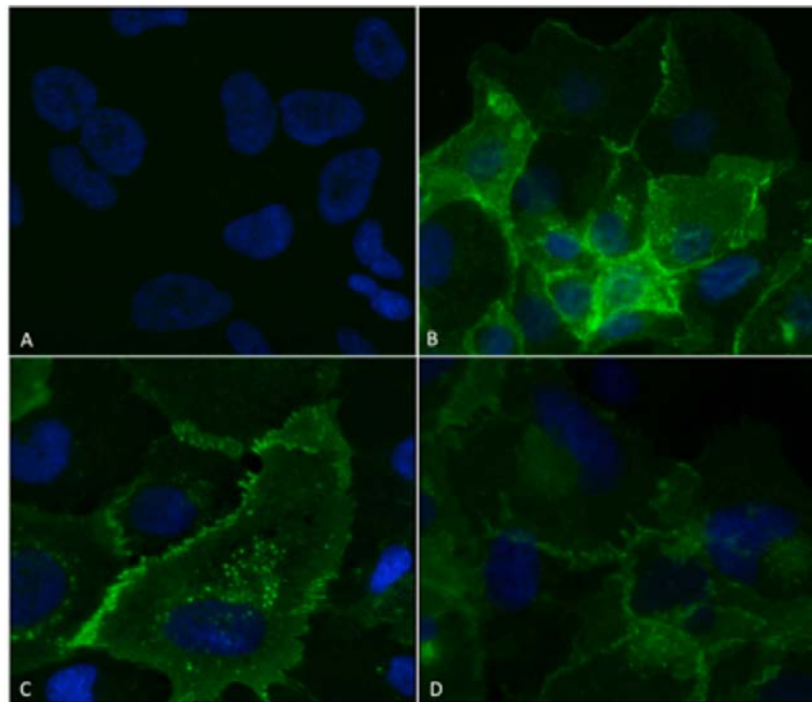
Taking into account the *in vitro* and *in silico* analysis, four non-synonymous variants (c.8C>G, p.P3R; c.1223C>T, p.A408V; c.1297G>A, p.D433N, and c.2494G>A, p.V832M) are considered potentially pathogenic (4/318 or 1.3%). If we only consider the subset of cases who have been tested and found not to carry *BRCA1* or *BRCA2* mutations, the prevalence of potentially pathogenic variants is 1.6% (4/246).

2.2.4 Discussion

Germline mutations in *CDH1* are associated with a substantively increased risk of LBC [159]. This study found that the prevalence of potentially pathogenic *CDH1* variants is low in early-onset and familial LBC cases who do not report a clear family history of DGC. The large sample size increases the likelihood that the results in this setting are precise. This study highlights the utility of publicly available registries as valuable resources of clinically- and epidemiologically-annotated families with accompanying germline DNA for future research in this field.

It remains possible that *CDH1* mutations are present in rare families with multiple LBCs even without GC. Although the cases in the present study had confirmed LBC, we were unable to confirm the pathology of the breast cancers in the relatives, which remained unspecified in the majority of the cases. Additionally, as 72 cases (23%) were not tested for mutations in *BRCA1* and *BRCA2* (Table 2), it is possible that some *BRCA1* and *BRCA2* mutation carriers were included in this study. The likelihood however is low, as the majority of early-onset and familial breast cancers are not accounted for by germline mutations in *BRCA1* and *BRCA2* [243, 250]. We had previously reported a pathogenic truncating *CDH1* mutation in an LBC case and her mother, who had both developed LBC before age 45 years [159]. However, our data suggest that *CDH1* associated LBC without GC must be very rare, as so few were identified in the present study among women highly selected for early-onset LBC or LBC with additional breast cancer in the family. It might still be prudent to consider germline *CDH1* testing in families with confirmed multiple cases of early-onset LBC, even in the absence of a family history of GC. In such families, and in those with a reported but unspecified history of a cancer occurring in the abdomen, the possibility of ovarian cancer would lead to *BRCA1* and then *BRCA2*

testing, and the possibility of DGC should lead to consideration of *CDH1* testing. For women with LBC, it is important to look for a family history of GC so that HDGC families will be recognized and offered appropriate management for their risk of DGC.

In our study, the pathogenic germline variant p.V832M was identified in an LBC case without a family history of GC. This variant was initially found to segregate with disease in a Japanese family where the proband had DGC at age 61 years and four of seven siblings, the mother, and a niece, all had unspecified GC. Functional characterization in Chinese hamster ovary cells demonstrated reduced cell aggregation and increased invasive properties of the mutant compared to WT E-cadherin [249]. Though this effect was not reproduced in functional characterization undertaken in human squamous epithelial cells[251], further work has demonstrated a mechanism by which this mutation might confer a pathogenic effect; through loss of type Ig phosphatidylinositol phosphate kinase (PIPK1γ) -binding, causing abnormal E-cadherin trafficking and adherens junction formation [252].

The novel non-synonymous variants in this study were not confirmed by our *in vitro* and *in silico* studies to be pathogenic, although further investigation needs to be done on the suggestive evidence that the variants c.1223C>T, p.A408V and c.1297G>A, p.D433N might interfere with calcium-dependent homophilic binding. Also, a novel, presumably rare variant (c.88 C>A, p.P30T) was shared by two LBC cases from one of the high-risk breast cancer clinics: this could imply that this variant is linked to the disease and that these two women are distantly related. Alternatively, this may represent a rare variant not associated with LBC, whose distribution in the normal population frequency will become known as the genomes of more people are sequenced. Data from the 1000 Genomes Project may also be helpful in the interpretation of the significance of these variants, through demonstration of the full profile of normal variation within *CDH1* and their distribution in and across populations.

The Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [16 (04, 2013) accessed], contains exome

sequence data from a total of 13,006 chromosomes from 2203 African-American and 4300 European-American unrelated individuals, totaling 6503 samples from individuals with heart, lung and blood disorders, not selected for cancer history. Based on exome sequence data generated from this cohort, variants in *CDH1* are distributed across the gene. Total variation of the sequence data surrounding the coding portions of *CDH1* can approximately be divided in thirds comprising intronic, silent, and missense variants. No splice or nonsense variants are reported and there is only a single frameshift variant recorded in the entire dataset. Missense variants have minor allele frequencies below 2% and a few silent and intronic variants have minor allele frequnecies as high as 40%.

Two of the variants that were identified in this study are now present in the ESP; c.88 C>A, p.P30T (rs139866691) and c.1223C>T, p.A408V (rs13813586) with a minor allele frequencies of 0.16%/0.00%/0.10% and 0.00%/0.02%/0.01%, respectively and listed in the order of European American/African American/all populations. The following updated *in silico* analysis of the variants uses Polyphen 2 that predicts the possible impact of an amino acid substitution on the structure and function of a human protein[427]. The c.88 C>A, p.P30T variant was predicted to be "probably damaging" with a score of 0.986 (sensitivity: 0.74; specificity: 0.96) and the c.1223C>T, p.A408V variant was predicted to be "probably damaging" with a score of 0.995 (sensitivity: 0.68; specificity: 0.97). Polyphen 2 predictions for the other variants are listed; c.1297G>A, p.D433N was predicted to be "benign" with a score of 0.023 (sensitivity: 0.95; specificity: 0.81), c.8C>G, p.P3R was predicted to be "benign" with a score of 0.001 (sensitivity: 0.99; specificity: 0.15), c.1813A>G, p.R695G was predicted to be "benign" with a score of 0.180 (sensitivity: 0.92; specificity: 0.87). The previously reported variant, c.2494G>A, p.V832M (rs35572355), was present in EVS with a minor allele frequency of 0.012%/0.02%/0.015% (European American/African American/all) populations. Polyphen 2 predicted the impact of the p.V832M to be "probably damaging" with a score of 1.000 (sensitivity: 0.00; specificity: 1.00).

Indeed, two of the non-synonymous variants identified in this study have since been reported in a large dataset of individuals not selected for cancer history and the frequency

of these variants is in keeping with other rare *CDH1* missense variants seen in the normal population, thus it is still not clear if these variants contribute to susceptibility to LBC.

Thus, although a combination of LBC and DGC is strongly indicative of germline mutations in *CDH1*, in the absence of a history of DGC, *CDH1* mutations appear to be extremely rare. It is possible that *CDH1* mutations would be more often identified in families with multiple documented invasive lobular or mixed ductal/LBCs in the absence of DGC, but such families are uncommon. Therefore, a history of early-onset or familial LBC should trigger specific questions around a history of cancer occuring in the abdomen that could represent gastric or ovarian cancer.

CHAPTER 3: Distinct Tumor Phenotypes Relate To Aberration Of Specific Pathways Due To Germline Or Somatic Variants

Following determination of the low frequency of germline mutations in *CDH1* associated with lobular breast cancer in women selected based on young age or family history of breast cancer, there were few mutations to followup functionally. Therefore, I sought to study other examples of strong genotype-phenotype correlations. In this chapter I first investigate the apparent association between multiple granular cell tumors and germline mutations in genes associated with the RAS-MAP kinase family. Support for this association, comes from the occurrence of multiple tumors in a patient with an underlying germline mutation in *PTPN11* as the basis of her congenital syndrome and subsequent literature review and investigation of her tumors.

An example of the overlap that can exist between MCA syndromes and tumor predisposition is within the RAS-MAP kinase family [253]. Based on the known association of tumor predisposition within this pathway, it was conceivable that MGCT occurred in association with LEOPARD syndrome [OMIM#151100], a MCA syndrome caused by perturbation of *PTPN11*. Furthermore, the dermatology literature had also reported GCT in solitary or multiple forms in other cases of RAS-pathway associated syndromes, such as Noonan Syndrome. *In vitro* work suggested that the recurrent T468M LEOPARD syndrome–associated mutation, also found in the patient who is the subject of this chapter, acts in a dominant-negative manner; compared to the activating mutations associated with Noonan Syndrome and sporadic leukemia [254]. With that said, the finding of allelic heterogeneity between Noonan syndrome and LEOPARD Syndrome in other genes argues against vast differences in the functional consequences of the causative variants [255]. Therefore, it is currently accepted that *in vivo*, the main mechanism of aberration of the RAS-pathway in Noonan and LEOPARD Syndrome is through activation of the pathway. More recently germline deletions of *PTPN11* have been linked to non-syndromic multiple enchondromatosis[256]. Altogether, these various genotype-phenotype correlations are examples of the many ways in which genes can demonstrate tissue-specific functions and phenotypes. In the case of *PTPN11*, germline loss-of-

function mutations in association with inactivation of the WT allele can lead to aberrant chondrocyte growth, whereas activating mutations in *PTPN11* can cause juvenile myelomonocytic leukemia (JMML) and acute myeloid leukemia [257].

3.1 Multiple Granular Cell Tumors Are Associated With LEOPARD Syndrome

Granular cell tumors are uncommon entities thought to be of Schwann cell origin [258]. They usually occur as benign solitary lesions in the skin or tongue, although MGCT can arise in up to 30% of cases [259]. Eight percent of GCT are familial and appear to be associated with MGCT [259]. In this setting malignant GCT have been reported in a father and son [259].

LEOPARD syndrome (LS) [OMIM #151100] is a rare autosomal dominant syndrome characterized by multiple lentigines, electrocardiogram (EKG) abnormalities, ocular hypertelorism, obstructive cardiomyopathy, pulmonary stenosis, genital abnormalities in males, retardation of growth and sensorineural deafness [260, 261]. LS is commonly caused by mutations in *PTPN11*, which encodes the protein-tyrosine phosphatase non-receptor type 11, or SHP-2 [262, 263]. SHP-2 acts upstream of the Ras/*RAF1*/Erk MAP kinase signaling pathway which promotes cellular differentiation and proliferation and oncogenesis [253]. Noonan syndrome (NS), Neurofibromatosis type 1 (NF1), Costello syndrome (CS), and cardio-facio-cutaneous syndrome (CFC) are other syndromes caused by aberrant signaling in this pathway. The first three of these are associated with varying risks of malignancy [253]. Isolated malignancies have been reported in association with five cases of LS [263-267]. However, due to the rarity of LS, it is difficult to determine whether there is an increased risk of cancer.

The *PTPN11* mutations that cause LS result in loss-of-function [268, 269] and a dominant-negative effect on the activity of Erk MAP kinase [268]. These contrast with the activating mutations of *PTPN11* which cause 50% of NS [270] and predispose these individuals to an increased risk of JMML [257]. Activating mutations in *RAF1* also cause NS [255, 271], although, paradoxically, a gain-of-function mutation in *RAF1* was also found in a patient with LS [255]. It might be predicted that an increased risk of malignancy in LS might be

confined to these rare cases of LS caused by gain-of-function mutations in *RAF1*. However, here we report the first molecularly defined case of LS associated with MGCT, and show that it is caused by a previously characterized loss-of-function mutation in *PTPN11*. This suggests that the role of SHP-2 in regulation of the Ras/MAP kinase pathway and tumorigenesis may be complex and supports evidence that GCT may be caused by dysregulation of the Ras/MAP kinase pathway.

3.1.1 Case Report

The female patient was born at term following an uncomplicated pregnancy. Birth weight was 2.9 kg. There were no neonatal concerns. Poor weight gain and frequent episodes of otitis media occurred during infancy. At four months of age, pulmonary stenosis and a ventricular septal defect were suspected by cardiac auscultation during an admission for pneumonia. Echocardiography in her late teen years showed mild pulmonary stenosis without evidence of hypertrophic cardiomyopathy. Dark freckling was noted from an early age, with some lesions being present from birth. At 8 years of age, two raised non-malignant moles were removed from her scalp and two similar lesions were removed from her back a few years later. During her adolescence, she developed MGCT of the skin and subcutaneous tissues. To date, fifteen tumors have been removed, ten from her upper body and five from her lower body. At 13 years of age, she was found to have a thoracic scoliosis requiring corrective surgery at age 16 years following progression of the curve from 35 to 55 degrees. At 16 years of age, she required cautery for recurrent epistaxis and also described Raynaud's phenomenon. Audiometry at 23 years of age revealed mid-range frequency hearing loss bilaterally, with the left side more severely affected than the right. Family history was unremarkable.

When examined at 23 years of age, she had tightly curled auburn hair. Height was below the 3rd centile, and weight was at the 15th centile. Her head circumference was at the 98th centile. Her eyes were normally set with minimal ptosis on the left. There was bilateral overfolding of the helices. On examination of the chest, no pectus abnormality was noted. There was a prominent second heart sound and a soft systolic murmur. She had extreme flexibility in the small joints of her hands and in her hips and knees. There were profuse lentigines and numerous keloid scars where the various skin lesions had

been removed. There were three GCT on her upper back (Figure 9). Based on history and examination, LS was suspected and she met the minimal criteria for diagnosis suggested by Voron et al. [272]. As MGCT have previously been reported in a patient with *PTEN* hamartoma tumor syndrome (PHTS) [273], germline analysis of *PTEN* was done and was negative for mutations. Histology and immunohistochemistry of the excised nodular lesions, which expressed the protein S100, were consistent with GCT (Figure 10 and 11). The final diagnosis was of MGCT of skin and subcutaneous tissues.

Figure 9 Three salmon colored nodules are demonstrated with two overlying the left scapula and the third crusted nodule located medially.

Note the profuse lentiginosis and midline scoliosis repair scar. Reproduced from [Multiple granular cell tumors are an associated feature of LEOPARD syndrome caused by mutation in *PTPN11*., Schrader KA, Nelson TN, De Luca A et al, 75, 185-189, 2009] with permission from John Wiley and Sons.

The cells have abundant amphophilic granular cytoplasm, cytologically bland nuclei that vary somewhat in size, and some nuclei have small nucleoli. (H&E stain; original magnification: × 100). Reproduced from [Multiple granular cell tumors are an associated feature of LEOPARD syndrome caused by mutation in *PTPN11*., Schrader KA, Nelson TN, De Luca A et al, 75, 185-189, 2009] with permission from John Wiley and Sons.
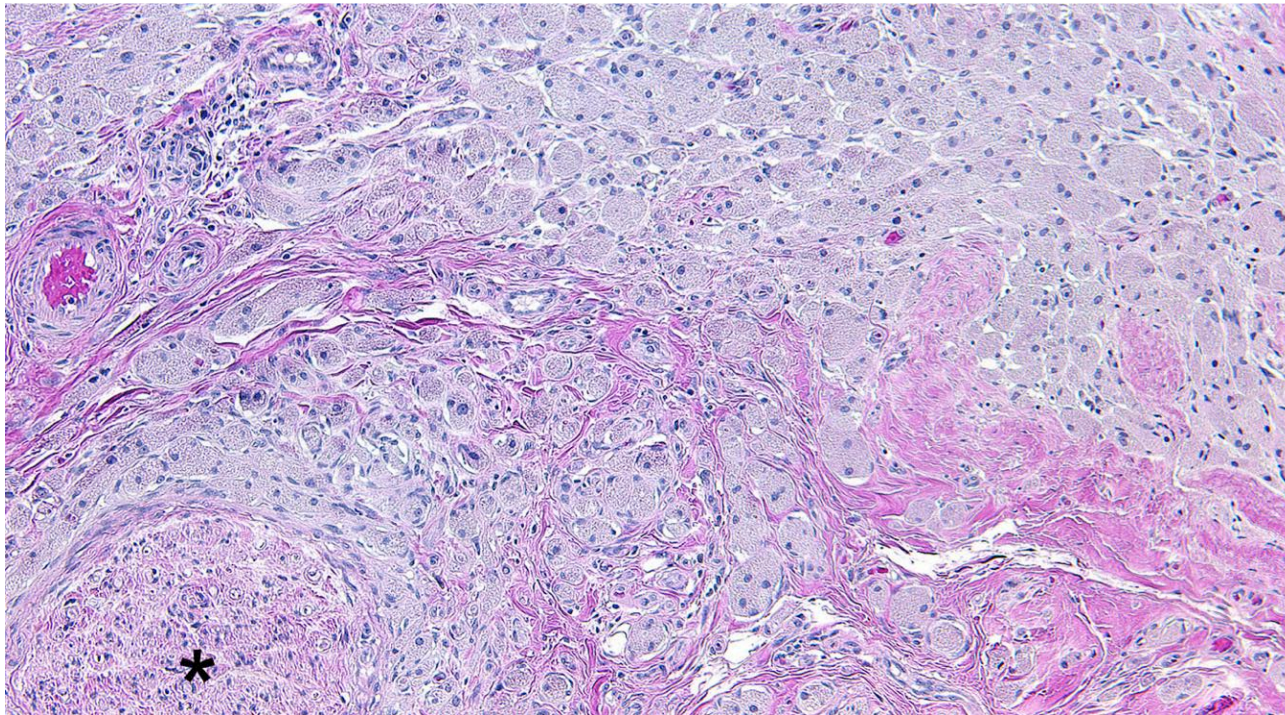
Stained with S100. (Original magnification: x 100). Reproduced from [Multiple granular cell tumors are an associated feature of LEOPARD syndrome caused by mutation in *PTPN11*., Schrader KA, Nelson TN, De Luca A et al, 75, 185-189, 2009] with permission from John Wiley and Sons.
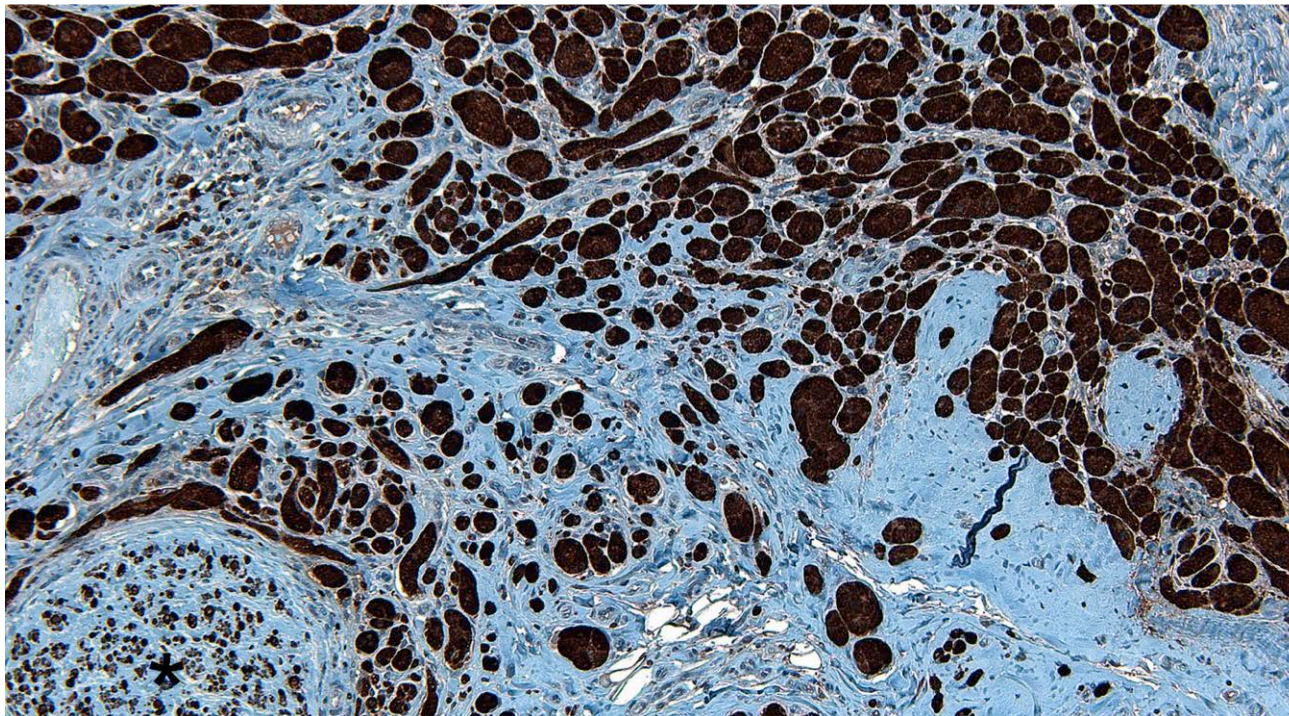
3.1.1.2 Methods And Results

To confirm the clinical diagnosis of LS, exons 7, 12, and 13 of the *PTPN11* gene were selectively amplified from leukocyte genomic DNA using primers flanking the intron-exon boundaries corresponding to the VariantSEQr™ primer sequences available online at the NCBI probe database. (http://www.ncbi.nlm.nih.gov/sites/entrez?db=probe&cmd=search&term=VariantSEQr). Bidirectional sequencing analyses of exons 7, 12, and 13 were carried out according to manufacturer's recommended VariantSEQr™ protocol, using an AB3130 and using SeqScape software for analysis (Applied Biosystems). These analyses revealed a previously reported missense mutation, c.1403C>T in exon 12 of *PTPN11*, that substituted methionine for the threonine at position 468 [262]. Analysis of six of the patient's tumors showed no LOH for the missense mutation as compared with surrounding normal epithelium and blood. Likewise, microsatellite analysis of six of the patient's tumors for LOH, using 6 markers tightly linked to the *NFI* locus (D17S841, D17S1863, D17S635, D17S1166, 3' *NF1*-1, 3' *NF1*-2) did not reveal LOH. Additionally, FISH analysis[298] in two GCT samples was negative for hemizygous deletion at the *PTEN* locus.

3.1.1.3 Discussion

In 1971, Selmanowitz reported a mother and daughter affected with LS. Both had strabismus and systolic cardiac murmurs, while the mother had bilateral hearing loss and hypertelorism. Of note, the daughter also had MGCT. It was postulated that the rare syndrome and the rare multiple tumors had a common origin [274]. However, the constitutional molecular lesion in these patients was not defined. Since Selmanowitz's report, two cases of MGCT have been clinically diagnosed with NF1 [275, 276] and another two cases have been clinically diagnosed as NS [277, 278]. The association of MGCT with LS, and the previous reports linking MGCT with other syndromes related to the RAS/MAP kinase pathway, led us to observe that many of the collated cases of MGCT

associated with syndromic features [279, 280] also exhibited other phenotypic features shared by the neuro-cardio-facial-cutaneous syndromes [281]. These included lentiginosis [282-284], cryptorchidism and ptosis [285], pulmonary stenosis and small joint hyperextensibility [286], hypertelorism and small joint hyperextensibility [279], and short stature [280, 287, 288]. Moreover GCT originate from Schwann cells, which seem to be particularly vulnerable to aberrations in the Ras/MAP kinase pathway [289]. Schwann cells are key components of neurofibromas [289], complex tumors of mixed cell origin. Neurofibromas commonly occur in association with loss-of-function of *NF1*, a negative regulator of the Ras/MAP kinase pathway [289].

As MGCT have been reported in association with NF1, we investigated the possibility that the T468M *PTPN11* mutation, in contrast to its reported dominant-negative effect, had, in the precursors of MGCT, an activating effect on the Ras/MAP kinase pathway that could have led to MGCT in conjunction with LOH of *NF1*. However, we did not detect LOH at the *NF1* locus in analysis of six of the patient's tumors. Additionally, there was no evidence of LOH for *PTPN11* in six GCT, making a classic tumor suppressor role for *PTPN11* in the cells from which GCT arise unlikely, although other perturbations of *PTPN11* cannot be excluded. If this T468M mutation of *PTPN11* acts as reported, as a dominant-negative and suppresses the Ras/Erk MAP kinase pathway, GCT might be caused by decreased Erk MAP kinase activity, leading to decreased differentiation and thus abnormal prolongation of proliferation. This mechanism would also be in keeping with the previous reports of malignancy in LS patients with constitutional *PTPN11* mutations that have since been shown to behave as loss-of-function mutations [263-265, 268, 269]. Thus, we propose that MGCT be recognized as an associated feature of LS, and as an example of yet another tumor related to aberrant signaling within the Ras/MAP kinase pathway. It is likely a somatic hit in another gene(s) is needed for the development of a GCT. However, the fact that only occasional patients with LS, NS or NF1 due to mutations in the Ras/MAP kinase pathway develop multiple GCT, suggests either the additional influence of modifier genes or environmental exposures. As the reports of malignant GCT in families with MGCT [259] are suggestive of a causal relationship between the benign and malignant forms of this disorder, we recommend that the lesions

be closely monitored for change or excised due to the low, but potential risk of malignancy.

We also explored other oncogenic mechanisms. Loss-of-function mutations of *PTPN11*, albeit not this mutation, have been shown to increase tyrosine phosphorylation of Gab1 resulting in activation of the PI3K pathway [269]. This could have been significant as MGCT occurred in a case of PHTS caused by a loss-of-function mutation in *PTEN*, which inhibits the PI3K pathway, although LOH for *PTEN* was not demonstrated in the tumors [272]. We ruled out a germline *PTEN* mutation in our patient, although it is conceivable that MGCT could evolve from the combination of germline mutations in the Ras/MAP kinase pathway and somatic alterations of the PI3K pathway. This, however, was not supported by assessment of two of the patient's tumors, which were negative for somatic deletion of *PTEN*, although more subtle perturbation of *PTEN* such as a point mutation or promoter methylation cannot be excluded. Further insights into the altered genomic regions involved in GCT formation might come from more global analyses such as array comparative genomic hybridization of the tumor and germline DNA.

In summary, MGCT should be considered an associated feature of LS and this finding implicates the Ras/MAP kinase pathway in the development of GCT.

3.1.2 Screening Of Sporadic Granular Cell Tumors To Determine If The Same Recurrent Germline Variant Plays A Role In The Sporadic Counterpart

To determine whether the recurrent p.T468M variant may have an important role in somatic tumor genetics of sporadic GCT, we identified 10 cases of sporadic GCT in unrelated patients that had been resected at Vancouver General Hospital. The tumors were microdissected and DNA was extracted from FFPE tissue. Primers flanking exon 12

of *PTPN11*, corresponding to the VariantSEQr™ primer sequences available online at the NCBI probe database.

(http://www.ncbi.nlm.nih.gov/sites/entrez?db=probe&cmd=search&term=VariantSEQr) were used to amplify the tumor DNA. DNA of sufficient quality to sequence was obtained from six of 10 of the samples. DNA extracted from paraffin can vary in its quality. Sanger sequencing for the recurrent p.T468M mutation was undertaken and sequencing in six of the six samples with good quality DNA, showed absence of the variant.

3.1.3 Future Directions In Establishing The Role Of *PTPN11* In Granular Cell Tumor Development

While the *in vitro* studies showed that this variant, *PTPN11* p.T468M inhibited EGF-evoked Erk activation[254], the *in vivo* state of Erk activation remains to be investigated in this patient's GCT. Demonstration of activated Erk would be in keeping with the general consensus that, despite the inability to activate EGF-evoked Erk, the overall effect of the variant is likely to activate the RAS-pathway. To answer this question, we have access to a frozen tissue specimen from the patient for parallel studies with the paraffin tissue blocks to determine whether there is increased phospho-ERK levels as compared with normal tissues and control specimens, since *PTPN11* is upstream of ERK and upregulation of the pathway results in ERK activation, measured by levels of phosphoERk [268]. If p.T468M functions as a dominant negative mutation, one would predict phospho-ERK to be downregulated in all of the patients tissues, conversely if it is a constitutively activating mutation we should see an overall increase in the phospho-ERK expression in the tumor and normal patient tissues. It is possible that if indeed there is a differential expression of phospho-ERK in the tumor as compared with he normal tissue this may be due to other molecular mechanism underlying the tumor pathology.

3.2 The Specificity Of The *FOXL2* C.402C>G Somatic Mutation: A Survey Of Solid Tumors.

The second example of the specificity of genotype-phenotype relationships is demonstrated by the recurrent somatic mutation in *FOXL2* associated with sporadic granulosa cell tumors [291]. A somatic mutation in the *FOXL2* gene is present in almost all (97%; 86/89) morphologically defined adult-type granulosa-cell tumors (A-GCTs) [291]. This *FOXL2* c.402C>G mutation changes a highly conserved cysteine residue to a tryptophan (p.C134W). It has been found in a minority of other ovarian malignant stromal tumors but not in benign ovarian stromal tumors or unrelated ovarian tumors or breast cancers [291]. Of note, germline loss-of-function mutations in *FOXL2* cause blepharophimosis epicanthus inversus, an MCA syndrome characterized by shortened palpebral fissures, ptosis and epicanthus inversus [OMIM 605597].

To determine the specificity of this mutation for the disease, we studied the DNA of other cancers and cell lines for the presence of c.402C>G *FOXL2*. We screened DNA from 752 tumors of epithelial and mesenchymal origin and 28 ovarian cancer cell lines and 52 other cancer cell lines of varied origin. We found the *FOXL2* c.402C>G mutation in an unreported A-GCT case and the A-GCT-derived cell line KGN[294]. All other tumors and cell lines analyzed were mutation negative. In addition to proving that the KGN cell line is a useful model to study A-GCTs, these data show that the c.402C>G mutation in *FOXL2* is not commonly found in a wide variety of other cancers and therefore it is likely pathognomonic for A-GCTs and closely related tumors.

3.2.1 Introduction

Malignant A-GCTs are malignant sex cord-stromal tumors known for their genomic stability and varied prognosis [290]. Until recently, there has been little insight into the molecular characteristics of A-GCTs. Using whole-transcriptome paired-end RNA sequencing; we identified a somatic missense mutation (c.402C>G, p. Cys134Trp) in the

Forkhead transcription factor gene, *FOXL2* [291]. This mutation was present in 97% of 89 morphologically identified A-GCTs [291]. *FOXL2* has been shown to be crucial for granulosa-cell differentiation [292]. This was the first association of a somatic mutation in *FOXL2* associated with cancer, however aberrant expression of *FOXL2* has been reported in juvenile granulosa-cell tumor of the testis [293]. The mutation was also found at a lower frequency in two other related ovarian stromal tumors; 21% (3/14) thecomas and 10% (1/10) juvenile-type GCTs were mutation positive [291]. This single, recurrent mutation suggests that it is characteristic of granulosa-cell tumors, and its high frequency implies that it is potentially a driver in disease initiation.

To determine the specificity of this somatic mutation, high resolution melting or polymerase chain reaction (PCR) -based allelic discrimination was used to screen a diverse collection of tumors and ovarian tumor cell lines. Additional cytogenetic analysis was performed to demonstrate the stable karyotype of the A-GCT cell line, KGN [294].

3.2.2 Materials And Methods

The Curve Analysis (HRM) was conducted at Novartis. Samples for HRM were purchased as either DNA or tissue blocks from vendors who provided unlinked anonymized specimens collected in accordance with applicable review boards approval, regulations and laws. Novartis does not require an ethical review committee for samples collected in this manner. Control DNA, used to validate the HRM assay, was extracted from anonymized tumor specimens compiled by the frozen tumor bank, OvCaRe (Ovarian Cancer Research), under written informed consent. Approval for analysis of these samples for the *FOXL2* mutation was obtained through the joint Clinical Research Ethics Board of the BC Cancer Agency and the University of British Columbia.

Seven hundred and fifty-two tumor DNA samples, of epithelial and mesenchymal origin (Table 4) were screened with HRM on the LightScanner™ instrument (Idaho Technology Inc., Salt Lake City, Utah) [295, 296]. For each tumor block, malignant cells composed >50% of the cellularity. Matched normal adjacent tissue was available for all cases. The assay was designed to detect sequence variants in the region from Ile102 to Phe138 in *FOXL2* (NP_075555.1). Since *FOXL2* is a single exon gene, PCR primers were placed

in the coding region (forward primer 5' AGAAGGGCTGGCAAAATAGC, reverse primer 5' GCCGGTAGTTGCCCTTCT) resulting in a 150 base pair (bp) amplicon.

Table 4 Summary of tumor types screened by High Resolution Melt Curve Analysis (HRM).

Sequence data is available for all screen positive samples. *Variants seen on HRM screen but not confirmed by sequencing (HRM false positive results).Reproduced from [The Specificity of the *FOXL2* C.402c>G Somatic Mutation: A Survey of Solid Tumors., Schrader KA, Gorbatcheva B, Senz J et al, 48, e7988, 2009] in accordance with open access and creative commons licensing.

	Total cases n=752 (excluding controls)	Normal by HRM	Confirmed positive for *FOXL2* c.402C>G mutation out of HRM positive cases
Ovarian cancer negative controls	14	11	0/3*
Ovarian A-GCT positive controls (including an unreported A-GCT case and the A-GCT cell line, KGN)	13	0	13/13
Bladder Cancer	40	40	
Breast Cancer	74	71	0/3*
Carcinoid Cancer	8	8	
Cervical Cancer	16	16	
Colorectal Cancer	77	75	0/2*
Endometrial Cancer	12	12	
Esophageal Cancer	21	21	
Gastric Cancer	90	89	0/1*

	Total cases n=752 (excluding controls)	Normal by HRM	Confirmed positive for *FOXL2* c.402C>G mutation out of HRM positive cases
Head & Neck Cancer	28	26	0/2*
Hepatic (HCC & Cholangiocarcinoma)	14	14	
Lung Cancer (All types)	125	123	0/2*
Melanoma	31	31	
Ovarian Cancer	32	32	
Pancreatic Cancer	4	4	
Prostate Cancer	37	37	
Renal Cancer	52	51	0/1*
Leiomyosarcoma	15	15	
Malignant fibrous histiocytoma-pleomorphic sarcoma	8	8	
Rhabdomyosarcoma	2	2	
Liposarcoma	4	4	
Fibrosarcoma	1	1	
Testicular Cancer	19	18	0/1*
Thyroid Cancer	42	42	

The primary screen used whole genome amplified (Qiagen Repli-G kit) DNA derived from frozen tissue blocks of untreated primary tumors. All samples which had an aberrant melting curve or which failed to amplify in the initial screen were followed up with a repeat HRM assay using unamplified DNA prepared from tumor and adjacent normal tissue. Tumor samples which were repeat positive for an aberrant melting curve were sequenced in duplicate, and the resulting sequence trace files were analyzed for mutations using the phrap/phred/consed software package (www.phrap.org). DNA from 27 ovarian tumor samples previously genotyped for the mutation using a previously validated TaqMan real-time PCR-based allelic discrimination assay (Applied Biosystems, Foster City, CA) specific for the *FOXL2* c.402C>G mutation [291] were used to validate the performance of the HRM assay. This included an unreported A-GCT case and the cell line KGN.

To establish the specificity of the *FOXL2* c.402C>G mutation in ovarian cancer cell lines, we used the same TaqMan real-time PCR-based allelic discrimination assay to genotype 28 ovarian cancer cell lines and 52 cancer cell lines of different tissue origin for the *FOXL2* c.402C>G mutation (Table 5).

To assess the cytogenetic profile of KGN, we utilized 24-color FISH (24XCyte, MetaSystems, Cat. D-0125-120-MC) and analyzed the results using the Axioplan 2, Zeiss,(MetaSystems, Isis), camera VAC-30054.

Table 5 Cell lines screened by TaqMan real-time PCR-based allelic discrimination assay for the FOXL2 c. 402 C>G mutation.

Ovarian cancer cell lines are italicized and ovarian granulosa-cell- derived lines are underlined. Reproduced from [The Specificity of the *FOXL2* C.402c>G Somatic Mutation: A Survey of Solid Tumors., Schrader KA, Gorbatcheva B, Senz J et al, 48, e7988, 2009] in accordance with open access and creative commons licensing.

	TaqMan for *FOXL2* c.402C>G (n=80)	Sequenced for *FOXL2* c.402C>G	Tumor Categories
A2780	-		Ovarian - High grade serous
CAOV3	-		Ovarian - Carcinoma not otherwise specified (NOS)
OV90	-		Ovarian - High grade serous
OVCAR3	-		Ovarian - High grade serous
OVCAR4	-		Ovarian - Carcinoma NOS
OVCAR5	-		Ovarian - Carcinoma NOS
OVCAR8	-		Ovarian - Carcinoma NOS
PE01	-		Ovarian - High grade serous
PEA1	-		Ovarian - High grade serous
PE014	-		Ovarian - Moderately-differentiated serous adenocarcinoma
<u>*SVOG*</u>	-		Ovarian - SV40 transformed normal human granulosa-cells
TOV21G	-		Ovarian - Clear cell
<u>*KGN*</u>	+	*FOXL2* c.402C>G	Ovarian - Granulosa-cell
CAOV4	-		Ovarian - Carcinoma NOS
DF-15	-		Ovarian - Carcinoma NOS
DF-22	-		Ovarian - Carcinoma NOS
DF-23	-		Ovarian - Carcinoma NOS
EFO-21	-		Ovarian - Carcinoma NOS
EFO-27	-		Ovarian - Mucinous
ES-2	-		Ovarian - Clear cell

	TaqMan for *FOXL2* c.402C>G (n=80)	Sequenced for *FOXL2* c.402C>G	Tumor Categories
FU-OV-1	-		Ovarian - Carcinoma NOS
Hey-A8	-		Ovarian - Carcinoma NOS
OVMANA	-		Ovarian - Clear cell
PA-1	-		Ovarian - Teratocarcinoma
RMUG-S	-		Ovarian - Mucinous
SK-OV-3	-		Ovarian - Carcinoma NOS
IGR-OV1	-		Ovarian - Carcinoma NOS
TOV-112D	-		Ovarian - Endometrioid
CCRF-CEM	-		Leukemia
HL-60(TB)	-		Leukemia
K-562	-		Leukemia
MOLT-4	-		Leukemia
RPMI-8226	-		Leukemia
SR	-		Leukemia
A549/ATCC	-		Non-Small Cell Lung
EKVX	-		Non-Small Cell Lung
HOP-62	-		Non-Small Cell Lung
HOP-92	-		Non-Small Cell Lung
NCI-H226	-		Non-Small Cell Lung
NCI-H23	-		Non-Small Cell Lung
NCI-H322M	-		Non-Small Cell Lung
NCI-H460	-		Non-Small Cell Lung
NCI-H522	-		Non-Small Cell Lung
COLO 205	-		Colon
HCC-2998	-		Colon
HCT-116	-		Colon

	TaqMan for *FOXL2* c.402C>G (n=80)	Sequenced for *FOXL2* c.402C>G	Tumor Categories
HCT-15	-		Colon
HT29	-		Colon
KM12	-		Colon
SW-620	-		Colon
SF-268	-		CNS
SF-295	-		CNS
SF-539	-		CNS
SNB-19	-		CNS
SNB-75	-		CNS
LOX IMVI	-		Melanoma
MALME-3M	-		Melanoma
M14	-		Melanoma
MDA-MB-435	-		Melanoma
SK-MEL-2	-		Melanoma
SK-MEL-28	-		Melanoma
SK-MEL-5	-		Melanoma
UACC-257	-		Melanoma
UACC-62	-		Melanoma
786-0	-		Renal
A498	-		Renal
ACHN	-		Renal
CAKI-1	-		Renal
RXF 393	-		Renal
SN12C	-		Renal
TK-10	-		Renal
UO-31	-		Renal

	TaqMan for *FOXL2* c.402C>G (n=80)	Sequenced for *FOXL2* c.402C>G	Tumor Categories
PC-3	-		Prostate
DU-145	-		Prostate
MCF7	-		Breast
MDA-MB-231/ATCC	-		Breast
MDA-MB-468	-		Breast
HS 578T	-		Breast
BT-549	-		Breast
T-47D	-		Breast

3.2.3 Results

All 11 previously reported *FOXL2* c.402C>G mutation-positive A-GCT specimens as well as an unreported A-GCT case and the A-GCT cell line, KGN, validated the HRM assay by demonstrating a variant melt curve distinct from the common (WT) pattern. None of the 14 *FOXL2* c.402C>G mutation negative samples exhibited this variant melt profile. However, three of the 10 high grade serous ovarian cancers showed an alternative variant profile; sequencing confirmed them to be false positives.

The primary screen of 752 whole genome amplified tumor DNA samples yielded 24 samples (~4%) with a variant profile distinct from that seen in association with the *FOXL2* c.402C>G mutation-positive A-GCT specimens, as well as 41 with an indeterminate profile and 29 samples that failed. The secondary screen was performed on this set of 94 samples using unamplified genomic DNA derived from tumors and matching normal specimens. Eighty-two of the samples were found to be false positives where there was no variant profile seen between the tumor and normal DNA. Twelve of the samples remained indeterminate and were subsequently sequenced and confirmed to be false positives.

The granulosa-cell line KGN that was derived through long-term passage of a recurrent A-GCT [294],was the only cell line found to harbor the mutation. The mutation was not present in an SVOG granulosa-cell line, immortalized by SV40 [297] or 26 other ovarian cancer derived cell lines. Unlike most ovarian cancer derived cell lines, KGN shows relative genomic stability (Figure 12). In addition to deletion of 7q, it is monosomic for chromosome 22 that is the most frequent cytogenetic abnormality seen in A-GCTs [298]; another feature demonstrating its similarities to A-GCTs.

24-color FISH demonstrates the tumor cell line's karyotype 45, XX, 7q-, -22 consistent with the original publication[294]. *FOXL2* is located at 3q23. Images were obtained using the Axioplan 2, Zeiss, (MetaSystems, Isis), camera VAC-30054. Reproduced from [The Specificity of the *FOXL2* C.402c>G Somatic Mutation: A Survey of Solid Tumors., Schrader KA, Gorbatcheva B, Senz J et al, 48, e7988, 2009] in accordance with open access and creative commons licensing.
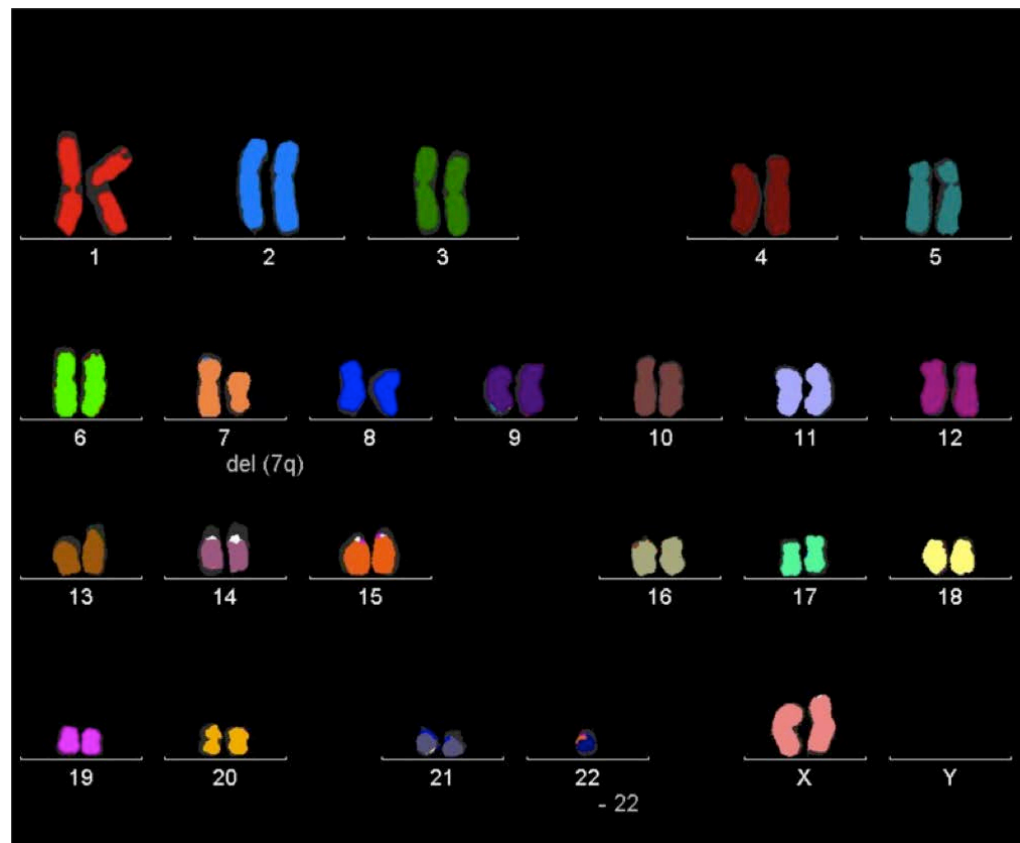
3.2.4 Discussion

Loss-of-function germline mutations in *FOXL2* are associated with blepharophimosis–ptosis–epicanthus–inversus syndrome [BPES;OMIM#110100]; an autosomal dominant developmental disorder characterized by eyelid malformations and premature ovarian failure due to a dysfunction of granulosa-cells [299]. The *FOXL2* c.402C>G mutation is seen in the heterozygous state in most A-GCTs [291]. Unlike in BPES, where germline *FOXL2* mutations are spread across the gene [300], the somatic *FOXL2* mutation in A-GCTs involves the same bp in all cases. This favors a specific functional consequence such as a dominant negative effect or a change or gain of function as opposed to a generic loss of function and the ultimate impact of this mutation is oncogenic. Additionally, immunohistochemical data indicating that *FOXL2* expression is maintained in the nuclei in A-GCTs, that were heterozygous or appeared to be hemizygous or homozygous for the mutation, implies that this mutation does not affect protein localization [291].

Analysis of 28 various ovarian cancer-derived cell lines demonstrated that the mutation was only present in the granulosa-cell tumor cell line, KGN, suggesting that it is molecularly akin to A-GCTs. The presence of the missense mutation in the well-characterized A-GCT cell line, KGN, is in keeping with the high frequency of the somatic mutation in A-GCTs and supports the use of this cell model to study the properties of this ovarian sex cord stromal tumor. This cell line has been used in a number of elegant studies that have addressed the question of the function of *FOXL2* [301] and the effects of *FOXL2* missense, haploinsufficient or hypomorphic mutations associated with BPES [302-304]. Further dissection of these phenomena with attention to the possible confounding effects of this mutation in one copy of the endogenous gene may elucidate the function of this missense mutation in the granulosa-cell tumor.

The absence of the *FOXL2* c.402C>G mutation in this large series of common epithelial malignancies such as lung, colorectal, breast, gastric, bladder, thyroid, prostate, melanoma and ovarian carcinoma, in addition to a range of less frequent tumors, implies a high specificity of this recurrent mutation for ovarian sex cord stromal tumors. This study does not exclude the possibility that the mutation could be found in other rare or related neoplasms such as testicular stromal tumors. As the mutation was not found in

non-GCT ovarian tumor cell lines and the SV40 transformed granulosa-cell line, SVOG, provides further support of its likely role in A-GCT disease initiation. Considering the extremely high frequency of this mutation in morphologically selected A-GCTs (97%)[291], these data provide further evidence suggesting that the mutation is also specific for this tumor type and could be useful as a diagnostic test. Further studies will be required to determine the relevance of the mutation in other sex cord stromal tumors of the ovary, however, it is possible that all mutation positive tumors could ultimately be considered to be a single entity of which the major component would be A-GCTs.

3.3 Serous Ovarian Cancer And *BRCA1* And *BRCA2* Germline Mutation Status

The third and final example of germline or somatic aberrations in particular genes, giving rise to specific tumor pathology, is that of *BRCA1* and *BRCA*2 aberrations in high-grade serous epithelial ovarian cancer. It has been observed that ovarian cancer associated with germline *BRCA1* and *BRCA2* germline mutations is usually high-grade serous ovarian, although, population studies examining this have rarely incorporated rigorous pathology review or molecular and immunohistochemical studies to ensure correct classification of the epithelial ovarian carcinomas. Therefore, I examined the rate of germline mutations in an unselected cohort of women with non-mucinous serous ovarian carcinoma and demonstrate that with adequate pathology review, and in some cases with reclassification of ovarian pathological classification, germline mutations in *BRCA1* and *BRCA*2 are restricted to the high-grade serous epithelial ovarian subtype.

3.3.1 Germline Mutations In *BRCA1* And *BRCA2*

Germline mutations in *BRCA1* and *BRCA2* predispose to autosomal dominant hereditary breast and ovarian cancer, where determination of carrier status in women presenting with ovarian cancer may impact prognosis and influence treatment [317]. Female carriers can be offered risk-reducing surgeries or can elect for heightened surveillance for breast and ovarian cancer. These implications extend to their families. Morphologic classification of epithelial ovarian carcinoma delineates five major histological subtypes: high-grade serous, clear cell, endometrioid, mucinous, and low-grade serous [306]. These constitute distinct clinical entities with differing presentation, prognoses, molecular and immunohistochemistry profiles [305-307], reflecting aberration of specific molecular pathways such as DNA damage repair, cell cycle control or promotion or control of growth and proliferation [306, 308]. Seventy percent of epithelial ovarian carcinoma is of the high-grade serous subtype [306]. Germline *BRCA1* and *BRCA2* mutations account for 12-15% of invasive epithelial ovarian carcinoma [217, 309, 310]. Ovarian cancers in *BRCA1* and *BRCA2* mutation carriers show almost exclusive association with epithelial ovarian

carcinoma, in contrast to the low likelihood of germ cell or stromal tumors seen in this population [311-313]. High-grade serous histology is reported in 77-93% of *BRCA1* and *BRCA2* mutation carriers in population-based series of invasive ovarian carcinoma [217, 309, 314, 315]. Our objective was to estimate the incidence of germline *BRCA1* and *BRCA2* mutation carriers in a consecutive series of women with nonmucinous ovarian cancer unselected for a personal or family history of breast cancer or ovarian cancer in British Columbia.

3.3.2 Materials And Methods

From 2004-2009, patients were recruited from the Vancouver General Hospital and British Columbia Cancer Agency in Vancouver, British Columbia, Canada. Ethical approval was obtained from the joint Clinical Research Ethics Board of the BC Cancer Agency and the University of British Columbia. All women who had been fully staged and were undergoing debulking surgery (primary or delayed) for cancers of ovarian/peritoneal/fallopian tube origin were approached for informed consent for the banking of tumor tissue. Pathology review was performed in all cases. Women with borderline ovarian tumors or mucinous histology were excluded and those with endometrioid, clear cell, or serous carcinomas were referred to the BC Cancer Agency Hereditary Cancer Program genetic counselors to discuss and perform germline *BRCA1* and *BRCA2* testing. Cascade genetic counseling of family members of *BRCA1* and *BRCA2* positive patients revealed from this study was offered through the same program. Carcinomas of mixed epithelial sub-types were included in the study, and those with a serous component were classified as high-grade serous. Press et al. previously characterized the *BRCA1* and *BRCA2* abnormalities in the tumors of 49 of these cases [316] and more recently the detailed immunohistochemical and molecular characterization of the primary tumors and clinical and outcome data related to the full cohort (n=131) has been published [317].

The sequence of *BRCA1* and *BRCA2* was determined from peripheral blood derived genomic DNA via standard bi-directional dideoxy sequencing of the entire coding and proximal intronic regions. The presence of large genomic rearrangements, conferred by

deletions or duplications, was determined via MLPA according to the manufacturer's protocol (MRC Holland).

Family history acquired at the time of initial consultation by the gynecologic oncology surgical team (non-genetic counselor) and by the genetic counselor at the time of counseling for *BRCA1* and *BRCA2* mutation testing was reviewed compared with BC Cancer Agency's referral guidelines for hereditary breast and/or ovarian cancer (Table 6) (http://www.bccancer.bc.ca/HPI/CancerManagementGuidelines/HereditaryCancerProgram/referralinformation/hboccriteria.htm). Family histories were considered in first and second-degree relatives. Whether they would have been referred based on the finding of invasive, non-mucinous epithelial ovarian cancer alone in the proband, versus whether they would have met other criteria for testing, was reviewed and compared to germline mutation status. This enabled us to determine the number of mutation carriers that were identified based on their history of ovarian cancer alone.

*Criteria relating to family history.

** History of cancer in cousins and more distant relatives from the same side of the family may also be relevant, although for the purpose of this studywas not included. Modified from the BC Cancer Agency. HBOC criteria. Available at: www.bccancer.bc.ca/HPI/CancerManagementGuidelines/HereditaryCancerProgram/referralinformation/hboccriteria.htm. Retreived April 24, 2012. Reproduced from [Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy., Schrader KA, Kalloger S et al, 120, 235-40, 2012] with permission from Wolters Kluwer Health.

Previously known *BRCA1* or *BRCA2* mutation in a close family member	* Confirmed *BRCA1* or *BRCA2* gene mutation
Proband	* Breast cancer at age 35 years or yournger, or
	* Two breast cancers with at least one occurring at age 50 years or younger, or
	Invasive, non-mucinous epithelial ovarian cancer, including cancer of the fallopian tubes or primary peritoneal cancer at any age
**History in first- or second- degree relatives on same side of family	* One breast cancer + one ovarian cancer, or
	* One male breast cancer + breast or ovarian cancer, or
	* Two breast cancers, both at age 50 years or younger, or
	* Two ovarian cancers, or
	* Three breast cancers, one diagnosed at age 50 years or younger
Ashkenazi Jewish heritage	* Personal or family history of breast or ovarian cancer

Statistical analysis was performed using SPSS 20.0.0 and OpenEpi 2.3.1 (OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version 2.3.1. www.OpenEpi.com, updated 2011/23/06, accessed 2012/04/10). Ninety-five percent confidence intervals (95% CI) for mutation frequencies were calculated under the assumption of binomial distributions of the observed numbers of cases using Fisher's Exact test. Age comparisons were performed using a Mann-Whitney U test.

3.3.3 Results

Of 131 women who participated in genetic counseling, 26 women (20%) were found to harbor germline mutations in either *BRCA1* (n=19) or *BRCA2* (n=7) (Figure 13). Pathology review revealed that all 26 of these women had high-grade serous ovarian cancer (Table 7). Therefore, if we only consider the 103 cases of high-grade serous, referral to the Hereditary Cancer Program for genetic counseling and the option of *BRCA1* and *BRCA2* testing based on the criteria of high-grade serous histology alone, would have revealed a *BRCA1* and *BRCA2*-germline mutation rate of 25% in this cohort.

Figure 13 Germline BRCA1 and BRCA2 mutations found within the study cohort of 131 women with nonmucinous epithelial ovarian cancer are spread across each gene.

Schematic diagrams of A. *BRCA1* and B. *BRCA2* genes. Exons are demarcated by vertical lines, and locations of mutations are approximated by circles above the gene and are labeled by arrows. Closed circles indicate a personal or family history consistent with hereditary breast and ovarian cancer, and open circles indicate sporadic ovarian cancer. All *BRCA1* and *BRCA2* germline mutations were found in patients with high-grade serous histology. Reproduced from [Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy., Schrader KA, Kalloger S et al, 120, 235-40, 2012] with permission from Wolters Kluwer Health.
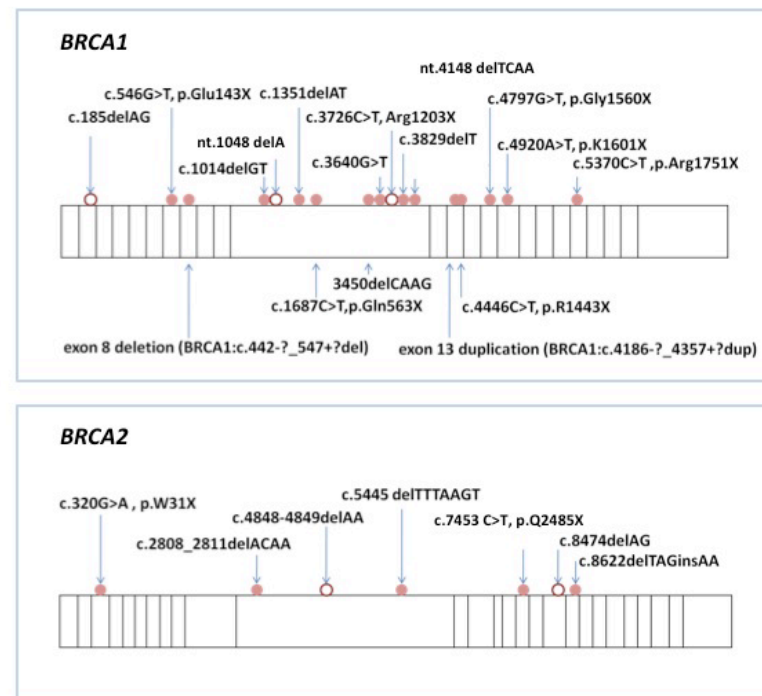
Total number (n) of cases for each histology and mean age $\pm$ standard deveiation are shown. In columns 3, 4 and 5 the number of cases, mutation frequency (percentage of total no. of cases), and 95% confidence intervals are given. Reproduced from [Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy., Schrader KA, Kalloger S et al, 120, 235-40, 2012] with permission from Wolters Kluwer Health.

Histologic Subtype	No. of Cases	Age (y)	BRCA1 Mutation-positive cases	BRCA2 Mutation-positive cases	BRCA1 and BRCA2 Mutation-positive cases
High-grade serous	103	58 $\pm$ 11	19, 18% (12-27)	7, 7% (3-14)	26, 25% (17-34)
Combined non-high grade serous	28	52 $\pm$ 11	0	0	0, 0 (0-12)
Low-grade serous	5	58 + 8	0	0	0, 0 (0-52)
Clear cell	12	49 + 12	0	0	0, 0 (0-26)
Endometrioid	11	52 + 10	0	0	0, 0 (0-28)
TOTAL	131	57 + 11	19	7	26, 28% (13-28)

Current guidelines for referral to the BC Cancer Agency Hereditary Cancer Program include the criterion of offering genetic counseling and testing to individuals with non-mucinous ovarian cancer diagnosed at any age. The importance of this criterion is evident, where even with a comprehensive family history obtained by a genetic counselor, germline *BRCA1* and *BRCA2* mutation carriers with high-grade serous histology and without a personal or family history suggestive of hereditary breast and ovarian cancer would not be identified. By showing the number of germline *BRCA1* and *BRCA2* mutation carriers with high-grade serous epithelial ovarian carcinoma meeting only the criterion of ovarian cancer in the proband, Figures 14 and 15 outline cases that would be captured by family histories (by genetic counselor or non-genetic counselor) as compared to histology-driven referral. Figures 14 and 15 show the difference between family history ascertained by the gynecologic oncology surgeon (denoted as non-genetic counselor) and the genetic counselor.

Further delineation of the *BRCA1* and *BRCA2* mutation carriers according to whether they met criteria for testing based on their history of nonmucinous ovarian cancer alone as opposed to whether they met other British Columbia Cancer Agency Hereditary Breast or Ovarian Cancer (BCCA HBOC) criteria or both. These results are based on history ascertained by the genetic counselor. Reproduced from [Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy., Schrader KA, Kalloger S et al, 120, 235-40, 2012] with permission from Wolters Kluwer Health.

Further delineation of the *BRCA1* and *BRCA2* mutation carriers according to whether they met criteria for testing based on their history of nonmucinous ovarian cancer alone as opposed to whether they met other British Columbia Cancer Agency Hereditary Breast or Ovarian Cancer (BCCA HBOC) criteria or both. These results are based on history ascertained by the gynecologic oncology surgeon or other health professional (nongenetic counselor). Reproduced from [Germline *BRCA1* and *BRCA2* Mutations in Ovarian Cancer: Utility of a Histology-Based Referral Strategy., Schrader KA, Kalloger S et al, 120, 235-40, 2012] with permission from Wolters Kluwer Health.
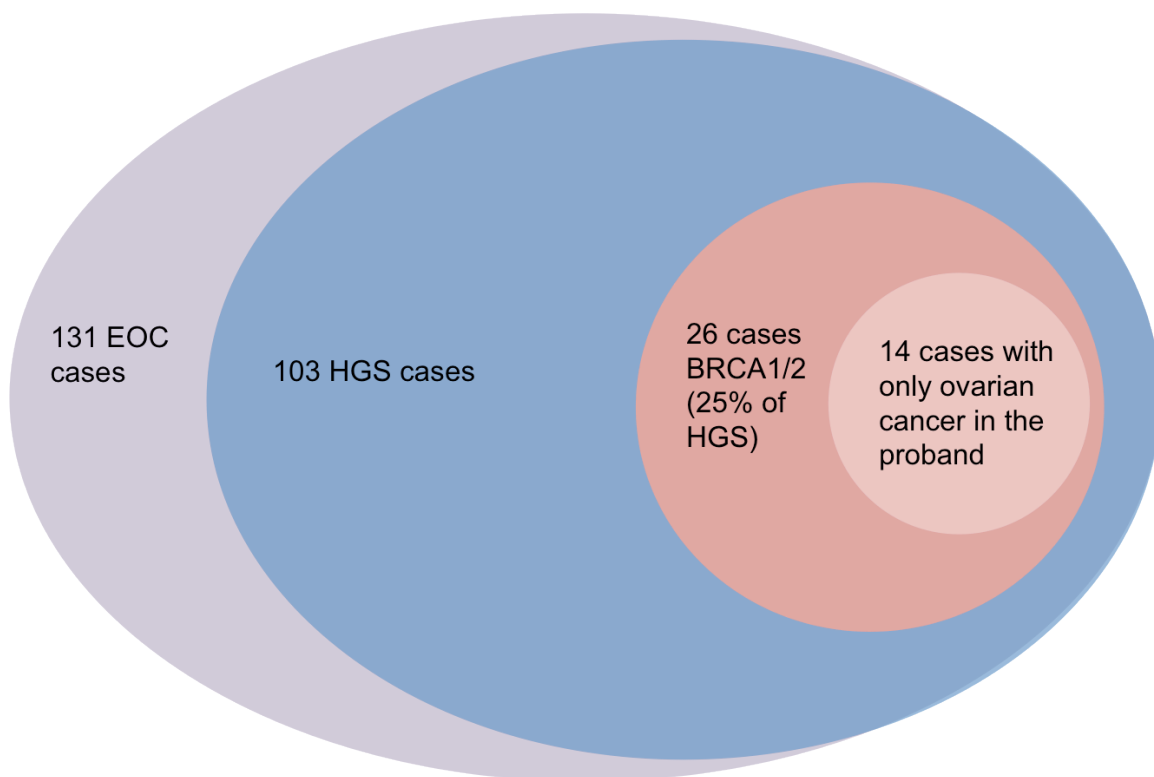
In both scenarios (Figure 14 and 15), the light pink circle demonstrates the number of women potentially missed if just relying on the family history criteria for hereditary breast, ovarian, or breast and ovarian cancer (denoted by an asterisk in Table 6) for referral for genetic counseling and testing of *BRCA1* and *BRCA2* mutation carrier status.

If the first family history was taken by another health care provider other than the gynecologic surgeon (e.g. general practitioner), this was also recorded as non-genetic counselor. Restricting *BRCA1* and *BRCA2* testing to women with family histories of hereditary breast and ovarian cancer, as ascertained by the surgeon, missed 14 mutation carriers, lowering detection rates to 9% (12/131), or 11.6% (12/103) if only considering the patients with high-grade serous histology. This improved to 16% (21/131) or, 20.4% (21/103) when ascertained by the genetic counselor. Therefore, upon reviewing the histories as obtained by the non-genetic counselor compared with those ascertained by the genetic counselor, it was apparent that 35% (9 of 26; 95%CI,17-56%) of *BRCA1* and *BRCA2* mutation carriers with high-grade serous epithelial ovarian carcinoma, would have been missed if referrals for testing had only been based on having a family history suggestive of hereditary breast and ovarian cancer (Figure 15). Even with the more extensive history obtained by a genetic counselor, 19% (5 of 26; 95%CI, 6-39%) of mutation carriers did not meet the familial component of hereditary breast and ovarian cancer criteria and prior to the inclusion of the criterion addressing singleton cases of non-mucinous epithelial ovarian carcinoma in the proband, would not have been offered testing. Capture of these probands was secondary to our research study protocol where counseling and testing was offered to all patients, and what was an emerging practice-change to consider referral for all non-mucinous ovarian carcinomas. If considering only cases with high-grade serous epithelial ovarian carcinoma and without a family history of hereditary breast and ovarian cancer, 11% of women (five of 46; 95%CI,4-24%), were germline *BRCA1* and *BRCA2* mutation carriers. Furthermore, the number of mutation-positive women without an apparent family history increased to 14 individuals when family history was taken by a non-genetic counselor. The average age of ovarian cancer diagnoses in women with high-grade serous epithelial ovarian carcinoma who did not

have a family history of hereditary breast and ovarian cancer was 51 years in mutation-positive cases (ranging from 44 to 61 years) and 61 years in mutation-negative cases (ranging from 37 to 84 years). This difference was only marginally significant when compared using a Mann-Whitney U test (p= 0.042). Of the two mutation carriers that harbored the recurrent c.185 delAG mutation, only one was reported to be of Ashkenazi Jewish descent. Two unrelated individuals carried exon 13 duplications.

3.3.4 Discussion

These results suggest that germline *BRCA1* and *BRCA2* mutations are confined to the high-grade serous histologic subtype, reflecting our own and other's experiences of frequent aberration of the *BRCA1* and *BRCA2* pathways within this subtype, but not in non-high-grade serous subtypes[317]. The same series reported herein was also analyzed for somatic mutations in *BRCA1* and *BRCA2* with five additional women discovered to have mutations of clinical significance in their tumors (germline DNA normal). *BRCA1* and *BRCA2* germline or somatic mutation frequency of high-grade serous cancers was therefore 30% (31/103) and similar to the reported findings by Hennessy et al (somatic and germline mutations in *BRCA1* and *BRCA2* in 23% of high-grade serous epithelial ovarian carcinoma) [318]. Methylation of *BRCA1* was found in an additional 20% of high-grade serous tumors bringing the total percentage of high-grade serous tumors showing *BRCA1* and *BRCA2* abnormalities to 50% [317].

Population-based studies have shown the incidence of germline *BRCA1* and *BRCA2* mutations associated with high-grade serous histology to be 17-18% [310, 314]. Although studies have identified *BRCA1* and *BRCA2* mutations in subtypes other than high-grade serous, the classification of tumors was based on pathology reports without review and use of current diagnostic criteria [217, 309, 314]. In the Cancer Genome Atlas, germline *BRCA1* and *BRCA2* mutations were found in only 15% (47 of 316 cases with high-grade serous ovarian carcinomas), where the lower mutation rate may have related to case selection and the technical aspects of the sequencing and mutation analysis [319]. Recently, Arnold et al, reported a 22.2% frequency of germline *BRCA1* and *BRCA2* mutations in patients with high-grade epithelial ovarian carcinoma with a serous component and no

Ashkenazi Jewish heritage [320]. Walsh et al reported a 25% *BRCA1-* and *BRCA2-* mutation frequency in patients with high-grade serous epithelial ovarian carcinoma, although also identified mutations in undifferentiated carcinomas, endometrioid and clear cell subtypes, possibly relating to the method and extent of pathology review that was not specified [321]. With the evolution of histopathological subtype diagnosis for ovarian carcinoma, and the appreciation of defining molecular abnormalities, specific subtypes can now be very reproducibly diagnosed with high interobserver agreement, using a combination of morphologic appearance and immunophenotyping [305, 322, 323]. In our cohort, we found germline *BRCA1* and *BRCA2* mutations to be exclusive to the high-grade serous subtype, suggesting that non-high-grade serous epithelial ovarian carcinoma are at very low risk of hereditary breast and ovarian cancer, and thus may not require referral to the Hereditary Cancer Program for genetic counseling. These results will need to be validated in a larger population-based series that also includes careful pathological review of the primary tumors, however, if the association holds true, a histology-based referral to a Hereditary Counseling Program should be incorporated in to the current referral schema. This would result in cost-savings and improved utilization of resources, preventing unnecessary referrals in non-high-grade serous cases where the incidence of *BRCA1* and *BRCA2* mutations of clinical significance is low or nil. It may also, as outlined, reduce the likelihood of missed probands when referral is based on family history alone. Thorough pathology assessment, using morphologic, molecular and immunohistochemical methodology [305, 307], is essential to the success of this strategy. Although clinical criteria for referral for hereditary breast and ovarian cancer syndrome exist and include referral of women with epithelial ovarian, fallopian tube or primary peritoneal cancer [National Comprehensive Cancer Network, (NCCN) guidelines Version 1.2011] also with the serous subtype being specified in the 2009 American College of Obstetricians and Gynecologists guidelines [324], currently there is no mechanism to ensure that they are being referred for genetic counseling regarding their risk for germline *BRCA1* and *BRCA2* mutations. One way to address this would be the inclusion of reflex-recommendations on the pathology report that recommend that genetic assessment be offered to all women with high-grade serous ovarian/fallopian tube/peritoneal carcinoma. This uniform referral strategy may lead to an increased opportunity for determination of

BRCA1 and *BRCA2* carrier status in women belonging to patient-populations that may be under-referred for genetic counseling. Further education in the community for referring family practitioners and general gynecologists regarding the importance of referral for genetic counseling for patients with high-grade serous ovarian carcinoma will also help ensure patients and families potentially at high-risk for breast, ovarian and other related cancers are not missed.

In view of the strong association and high incidence (25%) of underlying *BRCA1* and *BRCA2* mutations in women with high-grade serous ovarian (pelvic) carcinoma, genetic assessment for consideration of *BRCA1* and *BRCA2* germline testing should be offered to all women diagnosed with this histologic subtype of ovarian cancer, regardless of age or family history.

CHAPTER 4: The Application Of Next-Generation Sequencing Technology In Diagnosis Of Well-Defined Phenotypes

Targeted and whole genome sequencing has revolutionized the discovery of novel disease genes [338]. Eventually the low cost of sequencing and growing bioinformatic capacity will permit the routine clinical use of these tools in the rapid diagnosis of Mendelian disorders with heterogeneous phenotypes. Diagnostic algorithms used in clinical practice will evolve to reflect this, resulting in faster diagnosis, rapid expansion of disease phenotypes and better clinical management.

By identifying the shared phenotypes within affected individuals in a family, one can determine the genetic basis for Mendelian disorders by comparing the protein coding regions (the exome) of their genomes to look for shared novel variants. Depending on the pattern of inheritance and known ancestry of the family, the experimental design can be tailored to reflect the likely mode of inheritance of the disorder. To test whether this approach could be successful, exome sequencing of four family members affected with a rare and distinct phenotype of SED and RP was undertaken.

4.1 The Successful Use Of A Next-Generation Sequencing Approach To Diagnose The Genetic Basis Of The Occurrence In The Extended Family With Autosomal Recessive SED And RP.

The family originates from an isolated fishing community in Newfoundland, Canada (Figure 16). Individual family members initially presented with SED leading to multiple surgeries in the 3rd-6th decade. It was later determined that a RP phenotype, presenting in the 3rd and 4th decades with decreased night vision and leading to significant peripheral and central vision loss or blindness by the 5th-7th decade, cosegregated with the SED phenotype. Additionally, affected individuals had a high incidence of corneal abnormalities.

4.1.1 Phenotype Of Spondyloepiphyseal Dysplasia And Retinitis Pigmentosa

Consideration of the potential candidate genes was made prior to analysis, based on the presumed mode of autosomal recessive inheritance and the phenotype of SED and RP in the family from Newfoundland.

Disorders affecting bone composition, chondrocytes and bone signaling can lead to dysplasia of the spine and epiphyses termed SED. Various clinical subtypes of the disorder exist from SED congenita, typified by short-trunk dwarfism evident at birth, barrel chest, kyphosis, pectus carinatum, platyspondyly of the vertebrae, dislocation of the hip, flattened epiphyses and diminished joint mobility at elbows, knees and hips leading to a waddling gait, myopia and retinal detachment [OMIM#183900], is due to autosomal dominant mutations in *COL2A1*, while X-linked SED tarda is due to haploinsufficiency of *TRAPPC2* and is characterized by disproportionate short stature with an increased arm span to height ratio, a short trunk, broad chest, osteoarthritis and limited joint motion in back, hips and knees with typical sparing of the interphalangeal joints [OMIM 313400]. The severity and inheritance pattern of these two examples was not a good match for the family as the bony phenotype was more likely due to a recessive cause and if X-linked disease is considered, then the requirement for joint replacements in females in their 3 and 4th decades was more severe than the mild arthritis that has been reported in carrier females of *TRAPPC2* mutations [325]. However, it remained possible that a more complex situation involving skewed X-inactivation could explain a more severe female phenotype. More in keeping with the family's moderate short stature and varied spectrum of disease ranging from SED tarda to early-onset osteoarthritis beginning in childhood and affecting hips, shoulders, and interphalangeal joints, is SED-tarda with progressive arthropathy caused by autosomal recessive mutations in *WISP3* [OMIM 208230]. The disorder is characterized by short stature, kyphoscoliosis, generalized osteopenia and secondary osteoarthritis with the prominent phenotype being progressive joint problems beginning in early childhood that mimic juvenile inflammatory arthritis, however do not display biochemical or radiologic evidence of rheumatoid arthritis [326].

The severity of joint disease can exceed that of multiple epiphyseal dysplasia (MED), a bony dysplasia affecting the long bones that can cause short stature, joint pain in late childhood, deformity of the hands, feet and knees and scoliosis [326]. In addition to mutations in *COMP, MATN3, COL9A1, COL9A2, COL9A3* causing dominant MED, recessive MED can be caused mutations in *SLC26A2* [OMIM 226900]. *COL9A1* have also been found to cause a recessive form of Stickler syndrome [OMIM 120210], characterized by moderate to severe sensory neural hearing loss, moderate to high myopia with vitreoretinopathy and epiphyseal dysplasia [327]. Of note there have been no documented instances of retinal detachment in the family.

The retina is a complex sensory structure designed to capture light energy and convert it to electrical potentials to be interpreted by the brain. Rod and cone photoreceptors comprise the outer layer and contain photoreactive pigments that absorb light and use signal transduction cascades to transform the photons into electrical potentials. Photoreceptors synapse with second order retinal neurons that further synapse with the ganglion cells [328]. Abnormal structure and function of multiple components within this complex signaling network can lead to degeneration of the rod or cones, eventual visual loss and the appearance called RP, thus RP is genetically heterogeneous [329]. The majority of RP occurs in isolation however, can occur in association with a syndrome [329]. The frequency of RP in isolation is 1/5000 and the most common form of inheritance is autosomal recessive [329]. There are over 20 genes associated with autosomal recessive RP. All forms of inheritance have been described on the island of Newfoundland and within the Burin Peninsula itself (personal communication, J Green). The causative genes have been identified in several of the families [330]. There is no apparent link between families with isolated RP and the family under study in this thesis (personal communication, J Green). Newfoundland families with syndromic forms of RP have also been described, many have been those associated with Bardet Beidl, a genetically heterogeneous multisystem ciliopathy disorder characterized by RP, central obesity, mental retardation, polydactyly and renal dysplasia [OMIM#209900]. The incidence of BBS in Newfoundland is nine times that of Caucasians of northern European ancestry [331]. The inheritance of BBS is autosomal recessive and linkage and molecular studies of these families have helped identify and narrow candidate regions and genes for several BBS

disease loci [332-337] illustrating the genetic complexity that exists within the relatively isolated province.

The symptoms of rod dystrophy are night blindness, loss of peripheral vision and an eventual decrease in visual acuity. Cone degeneration typically produces, central vision loss, photophobia, and decreased color vision. Rod and cone functions can be discretely tested by electroretinogram (ERG) studies and genotype-phenotype correlations are seen associated with several genes. The spectrum of proteins associated with RP range from those with roles in visual transduction, photoreceptor structure, extracellular matrix, retinal development, protein folding, retinoid cycle, disc shedding, transcription factors, ore-mRNA splicing, protein folding/trafficking, cGMP channels and photoreceptor cilia [328].

4.1.2 Expansion Of The Phenotype Of Mucolipidosis Type III Phenotype

4.1.3 Materials And Methods

4.1.3.1 Study Design

Due to the geographic isolation, known consanguinity and autosomal recessive inheritance pattern in the family (Figure 16), it was predicted that the causative mutation would most likely be novel, and would lie within an extended block of linkage that was homozygous in the affected individuals and heterozygous in the unaffected obligate carrier. We performed exome sequencing of three affected individuals (63, 83, 104) and one obligate carrier (94) to look for novel variants that were consistent with this pattern of inheritance in all individuals (Figure 16). The research protocol was approved by the BC Cancer Agency research ethics board and the Memorial University Human Investigation Committee. All participants gave written informed consent.

Figure 16 Identification of the family's candidate mutation by exome sequencing.

Pedigree of the family, indicating disease status. Individuals 63, 83, 104 and 94, shaded in green, red, blue and brown, respectively, were exome sequenced. Due to the geographic isolation, known consanguinity and autosomal recessive inheritance pattern, it was predicted that the causative mutation would most likely be novel, and would lie within an extended block of linkage that was homozygous in the affected individuals. Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.
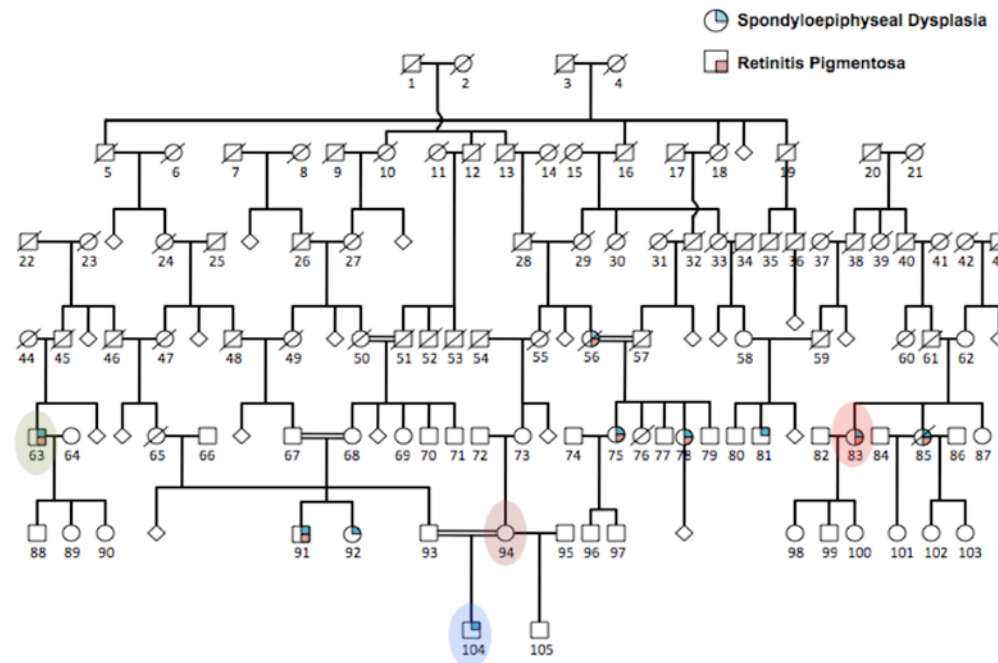
4.1.3.2 Ruling Out Homozygous Microdeletions

Affymetrix Genome-Wide Human SNP Array 6.0 assays were also performed on DNA from the same individuals, to rule out homozygous microdeletions and to confirm blocks of linkage surrounding candidate novel variants. Samples were prepared in a standard manner[448] and analysis was performed using the Affymetrix Genotyping Console™ software.

4.1.3.3 Exome Capture, Sequencing And Bioinformatic Analysis

Exome capture was achieved through solution hybrid selection with the Human All Exon kit SureSelect Target Enrichment System (Agilent) for Illumina Genome Analyzer paired-end sequencing [339]. Two lanes of Illumina paired-end sequencing were prepared for each exome capture library. Short paired-read (75bp) sequences obtained from the Illumina Genome Analyzer were mapped to the reference human genome (NBCI build 36.1, hg18) using MAQ (version 0.7.1) in paired-end mode [340]. The Sequence Alignment/Map (SAM 0.1.7) format was used for downstream processing. Insertion and deletion (indel) information was extracted from the alignment data using the Samtools package [341]. SNVMix, capable of distinguishing homozygous variants, was also used for the inference of SNVs [342], and was modeled on the expectation of normal Hardy-Weinberg frequencies consistent with a diploid genome. Variants were enriched for novel non-synonymous, indel and splice-site variants, by filtering with those already present in dbSNP130, The 1000 Genomes Project, and in in-house exomes.

4.1.3.4 Biochemical Confirmation Of The Pathogenic Variant

Pathogenicity of the candidate variant in *GNPTG*, was confirmed by measuring the serum enzyme activities of hexosaminidase, α-N-acetylglucosaminidase, α-mannosidase, and β-

glucuronidase which are lysosomal enzymes known to be dependent on mannose-6-phosphate targeting. Each of the four lysosomal enzymes was assayed fluorometrically, by incubating serum together with the appropriate 4-methylumbelliferyl (4MU)-glycoside conjugate, at 37°C and acid (4.0-5.0) pH. Reactions were stopped after a 30-120 minutes, by addition of alkaline buffer (1M glycine-NaOH buffer pH 10.0), depending on the enzyme assayed. Free 4MU, released from the substrate by enzymatic hydrolysis, was quantitated by measuring its fluorescence (excitation wavelength 350 nm, emission 440 nm). Biochemical analysis was performed on samples from four affected and three unaffected-carrier family members in parallel with seven unrelated healthy controls. For each enzyme, all serum samples were assayed simultaneously within a single batch.

4.1.3.5 Results

On average, coverage of targeted exons for greater than 10 reads was 87%. Filtering revealed an average of 412 novel non-synonymous variants, 242 novel indels and 60 splice site variants per case.

Based on the assumption of a recessive mode of inheritance, we required candidate variants to be homozygous in all three affected individuals and heterozygous in the obligate carrier. Only two novel variants fulfilled this requirement, a non-synonymous variant in *RPL3L* c.1013G>C, p.338A>G and a 6bp deletion in *GNPTG* c.238-243del, p.80K_81Ydel. These variants were both homozygous in individuals 63, 83, and 104 and heterozygous in the obligate carrier individual 94. Both of these variants were validated and found to segregate with disease in 14 further family members using Sanger sequencing. Five of five affected family members were homozygous for the variants, eight of eight unaffected obligate carriers were heterozygous for each variant and only one of two individuals at 50% risk of being carriers, was found to be heterozygous for both variants and the other individual was not found to carry either of the variants.

Both variants were located within a 3.5Mb region of linkage containing 202 University of California, Santa Cruz (UCSC) genes on chr16:482,030-3,938,263 (hg18). Within the 3.5

MB interval, 134 genes, out of the 202 UCSC genes, had been targeted by exome capture. Coverage of targeted exons in the 3.5 MB region of linkage ranged from 46-59% and there were on average, 5 novel non-synonymous variants, 1 indel, and no splice-site variants per affected individual. Manual review of novel variants within the region revealed two further non-synonymous variants in *MSLNL* that followed the expected pattern of inheritance. These were called in more than one individual, although had not met our full criteria by filtering strategies. The remaining novel variants in the region were not shared. This was confirmed by manual review.

4.1.3.6 Confirmation Of Pathogenicity Of The 6 Base Pair Deletion In GNPTG

The most promising candidate variant was the 6bp deletion, c.238-243del, p.80K_81Ydel in exon 5 of *GNPTG* (Figure 17), an OMIM disease gene associated with mucolipidosis type III gamma (MLIIIγ) [OMIM#252605]. The 6bp deletion removes highly conserved lysine and tyrosine residues. The deletion of each of these amino acids is predicted to be "probably damaging" by Polyphen 1. Further support for pathogenicity was based on the absence of the deletion in 368 chromosomes from an ethnically matched control population previously ascertained through the Newfoundland Familial Colorectal Cancer Registry[449].

Figure 17 Identification of the family's candidate mutation by exome sequencing.

Entropy figures, showing the allele distribution of the position of the mutation and surrounding sequence post-filtering, using a base-calling threshold of 10 and allowing for up to three mismatches per paired-end read. In these figures, the single letter corresponds to the homozygous state, the lack of a letter corresponds to homozygous null at that particular bp position, and two letters corresponds to the heterozygous state. The overall height of the columns indicates the information content and the height of the symbols indicate the relative frequency of each nucleic acid at that position[450]. Below the entropy figures are corresponding screen shots of the BAM files viewed with IGV 1.4.2[451], showing the reference DNA and amino acid sequence above the raw paired-end reads overlying the 6bp deletion in *GNPTG*. Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.
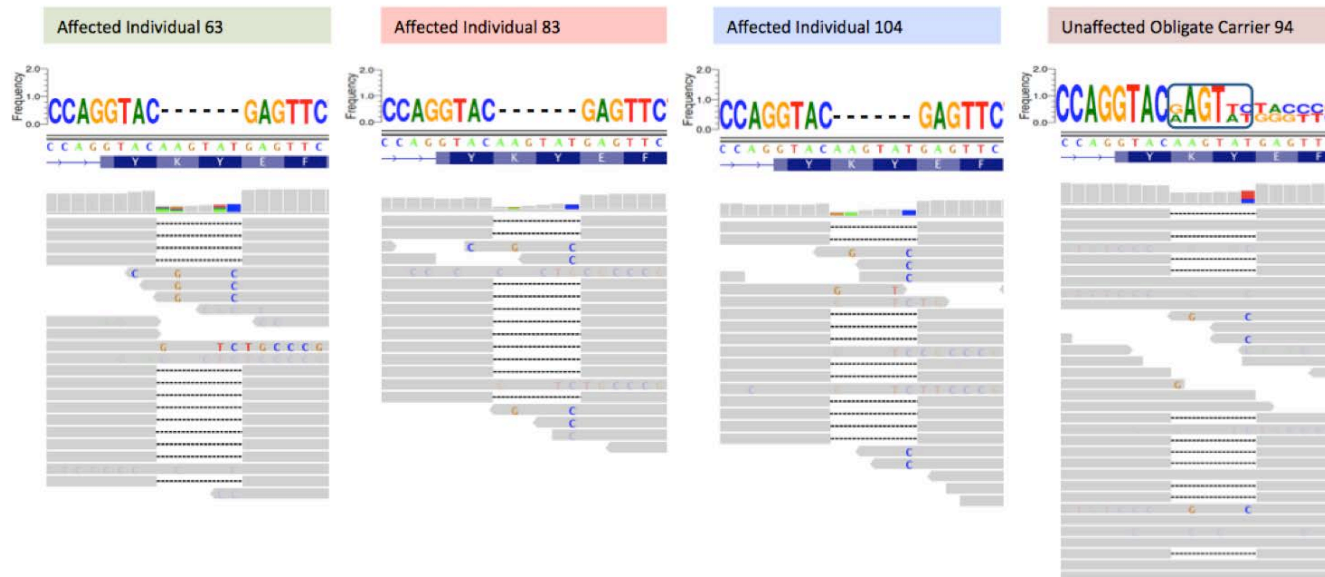
MLIIIγ is caused by biallelic germline loss-of-function mutations in *GNPTG* (Figure 18), which encodes the gamma subunit of GlcNAc-1–phosphotransferase [343-350], while the related disorder MLIII alpha/beta (MLIIIα/β) is caused by mutations in *GNPTAB*, which encodes the alpha and beta subunits of the same phosphotransferase. The phosphotransferase complex plays a crucial role in targeting of many enzymes to the lysosome. In both subtypes of MLIII, loss of phosphotransferase function causes mistargeting of lysosomal enzymes, resulting in abnormal accumulation of various lysosomal substrates in tissues [344, 351].

Figure 18 Identification of the family's candidate mutation by exome sequencing.

(C) Mutations found in *GNPTG* and its encoded protein N-acetylglucosamine-1-phosphotransferase subunit-γ. The 11 exons of *GNPTG* are represented (as numbered blue boxes) above a schematic of the N-acetylglucosamine-1-phosphotransferase subunit γ-protein (the teal segment, with the conserved PRKCSH domain in green). Below the schematic are all of the reported nucleotide mutations, including the novel mutation reported in this thesis (boxed and denoted by an asterisk). Nucleotide position numbering is based on starting with the A in the ATG start codon for *GNPTG* in position 1 (based on the sequence given in record number NM_032520.4 in Entrez Gene; UTR denotes untranslated region). Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.
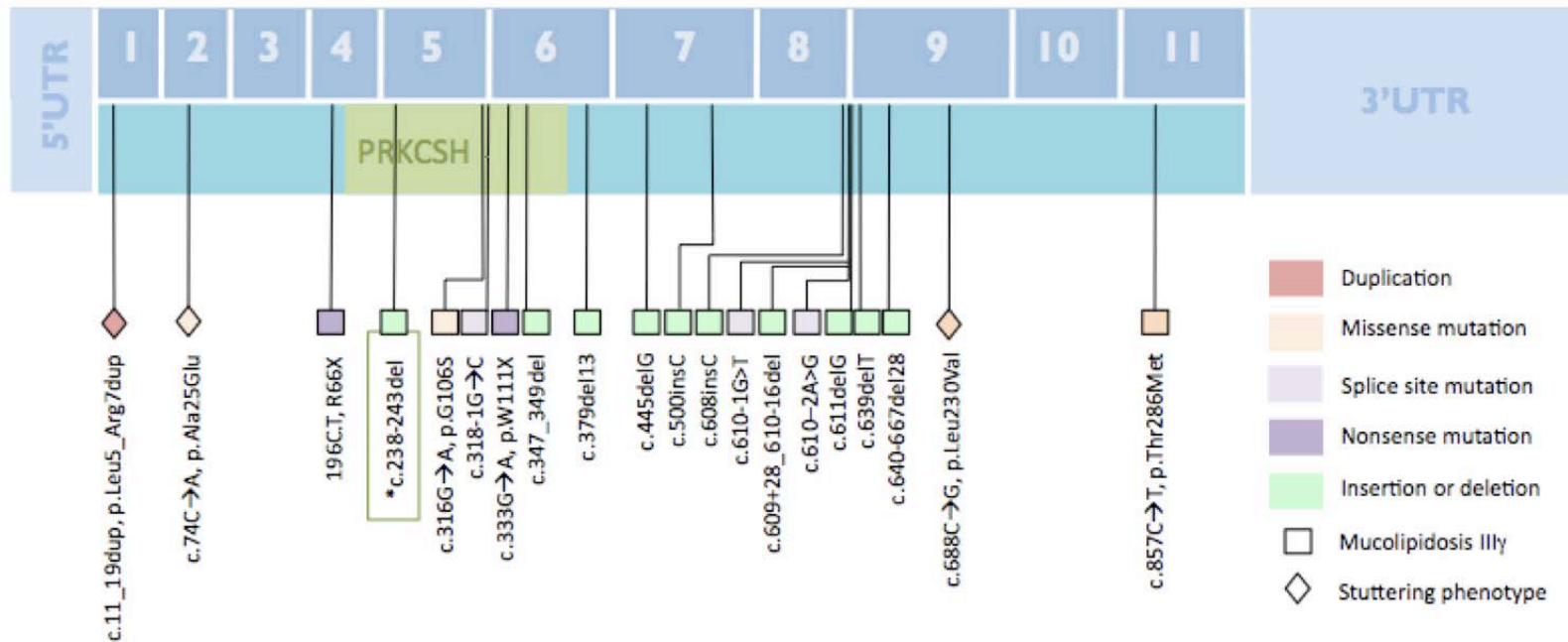
Lysosomal enzyme activities were found to be markedly elevated in serum of affected individuals compared to levels in carriers and healthy controls, indicating enzyme mistargeting and thus confirming the diagnosis of MLIIIγ (Table 8). Carriers did not demonstrate a phenotype.

Table 8 Elevations of serum lysosomal enzyme activities by at least 10-fold have been regarded as biochemical diagnostic criteria for MLIII.

Comparison of serum enzyme activities by least square means (using Tukey HSD) showed statistically significant (p < 0.0001) elevations for all four enzymes in affected subjects versus normal control subjects and versus carriers. Serum enzyme activities in carriers did not differ significantly from normal controls. *All enzyme activities are expressed in nmol/h/ml serum. Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.

	Hexosaminidase	α–N–Acetylglucosaminidase	α–Mannosidase	β–Glucuronidase
Activities* in normal controls (n = 7)	Mean 1267 Range 743–1718	Mean 33 Range 27–40	Mean 51 Range 36–73	Mean 243 Range 129–380
Activities* in carrier individuals (n = 3)	Mean 1166 Range 1036–1374	Mean 36 Range 32–44	Mean 55 Range 52–63	Mean 209 Range 165–286
Activities* in affected patients (n = 4)	Mean 19 139 Range 16 469–21 156	Mean 415 Range 318–470	Mean 7888 Range 5850–10 016	Mean 16 454 Range 13 055–19 017
Range of fold-elevations in affected patients versus mean of normal controls	13–17-fold	10–14-fold	114–196-fold	54–78-fold

Histochemical analysis of a section of individual 104's femoral head, who had undergone joint replacement at age 23 years, revealed microvesicular change in the chondrocytes consistent with the abnormal chondrocyte morphology observed in the mucolipidosis mouse model (Figure 19a).

Figure 19 Demonstration of the associated bone and retinal abnormalities.

Panel 1, section of normal cartilage; panel 2, section of cartilage from the femoral head of individual 104 (A). Note the pronounced microvesicular change in the chondrocytes. Of note, the appearance of the chondrocytes are in keeping with those seen in *GNPTAB*-null mice, compared with *GNPTG*-null mice, which do not display an abnormal chondrocyte morphology [352]. Panel 3, section of the femoral head of individual 104, showing macrophages in the subchondral marrow space with accumulated glycosolated intracellular material and inclusions. All sections were stained with periodic acid–Schiff with diastase digestion (PAS-D) and photographs were taken at ×400 original magnification. (B) Retinal photographs of bone spicules in individual 91 at age 28 years (left) and individual 85 at age 35 years (right). Bone specimen processing is courtesy of Dr Barry Gallagher and Dr Torsten Nielsen who in addition stained and photographed the samples. The retinal photographs are courtesy of Dr Jane Green. Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.
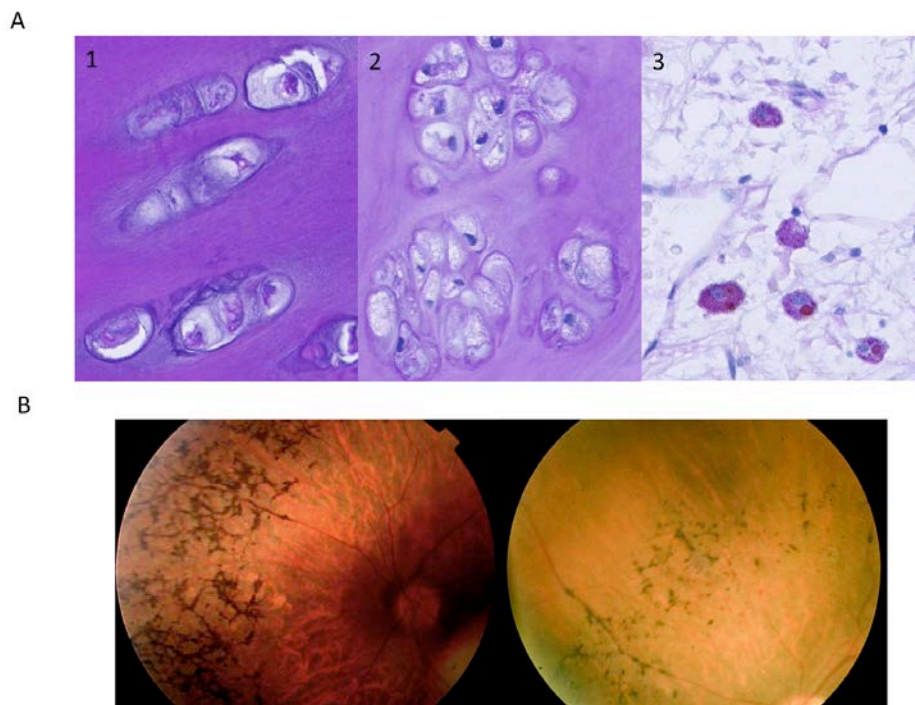
4.1.3.7 Discussion

Mucolipidosis type III is an autosomal recessive, lysosomal storage disorder typically defined by progressive joint stiffness, scoliosis, coarse facies, mild intellectual disability, dysostosis multiplex with progressive destruction of the hip joint, increased lysosomal enzyme levels in serum, reduced lysosomal enzyme levels in cultured fibroblasts, corneal clouding and opacities [343, 353]. Most published clinical descriptions of MLIII predate the genetic distinction between MLIIIγ and MLIIIα/β. As the latter is apparently more common, the specific manifestations and natural history of MLIIIγ have yet to be fully defined. However, the spectrum of disease in MLIII overall is notoriously broad. Furthermore, rare variants in *GNPTAB* and *GNPTG* have recently been implicated in nonsyndromic familial stuttering [354].

The degree of retinal degeneration seen in this family has not yet been reported in association with MLIII (Figure 19b); nor even in MLII, a severe infantile-onset disorder allelic to MLIIIα/β that involves complete ablation of the phosphotransferase activity [355]. Severe retinal degeneration is seen in *GNPTAB* null mice, implying a causal link between loss of phosphotransferase function and retinal pathology [356]. However, *GNPTG* null mice have not been shown to develop retinal disease [352]. This family's mutation in *GNPTG* causes the deletion of a lysine and tyrosine residue in the phosphotransferase g-subunit. This subunit is thought to have a role in regulation of phosphotransferase complex structure and function, rather than a direct role in catalysis [347]. It is possible that the presentation with severe retinal degeneration reflects specific effects of the particular mutation on phosphotransferase function.

Variable presentations of MLIII now need to be considered as a possible etiology for other reported syndromes of SED or spondylometaphyseal dysplasia associated with retinal manifestations [357-359]. In addition, reported findings of low neuraminidase activity in fibroblasts, but normal activity in leucocytes, in patients with spondylometaphyseal

dysplasia with cone-rod dystrophy [360], suggest MLIII as the cause [361]. This broadens the phenotypic spectrum of MLIII even further.

This next-generation sequencing approach allowed a rapid molecular diagnosis of this family's complex phenotype of MLIIIγ, with a newly reported severity of rod-cone dystrophy. This enabled a molecular diagnosis for the family and furthermore has significant clinical implications for the affected family member's medical management with regard to cardiac surveillance, anesthetic considerations and management of potential atlantoaxial instability

The use of exome sequencing to study four individuals from a single family allowed us to rapidly narrow the list of candidate mutations to two, and to correctly diagnose the disease gene using a standard clinical assay. This mutation could have been found by a combined homozygosity mapping and exome sequencing of one affected individual approach, and then performing segregation analysis on each novel variant in the region. However as we expected more false positives, our strategy was to collect multiple exome sequences and filter for novel non-synonymous variants based on their expected homozygous or heterozygous state to rapidly, and without bias, select the true candidate variants.

Traditional linkage and homozygosity mapping approaches identify regions of linkage or identity-by-descent (IBD) that contain multiple candidate genes. In all approaches ranking of candidate genes or variants usually requires some knowledge of the gene functions and possible disease mechanisms. This process can be time-consuming, however is expedited by next-generation sequencing approaches such as the one (Figure 20). Furthermore, as whole genome sequencing costs come down to that of SNP arrays, in the clinical setting of heterogeneous Mendelian disease diagnosis, sequencing of multiple individuals will provide the most direct approach.

Figure 20 Expedited diagnosis of an autosomal recessive condition in a consanguinous family by next-generation sequencing.

Linkage analysis, homozygosity mapping and exome sequencing are three strategies that can be used to search for disease genes. Potential end points are indicated in the red hatched boxes. Homozygosity mapping in consanguineous families can direct the investigator to candidate genes within regions of IBD. However, depending on the degree of consanguinity, there can be multiple regions of IBD and further sequencing to determine the pathogenic variant is required. The exome- sequencing approach can, in one step, provide the investigator with a list of candidate variants. Although separate genome-wide analysis of polymorphisms is still necessary to rule out deletions or duplications missed by exome sequencing, these would be detected by whole-genome sequencing. A combined homozygosity-mapping and exome-sequencing approach is also discussed in the text. Reproduced from [Using next-generation sequencing for the diagnosis of rare disorders: a family with retinitis pigmentosa and skeletal abnormalities., Schrader KA, Heravi-Moussavi A, Waters P et al, 225, 12-8, 2011] with permission from John Wiley and Sons.
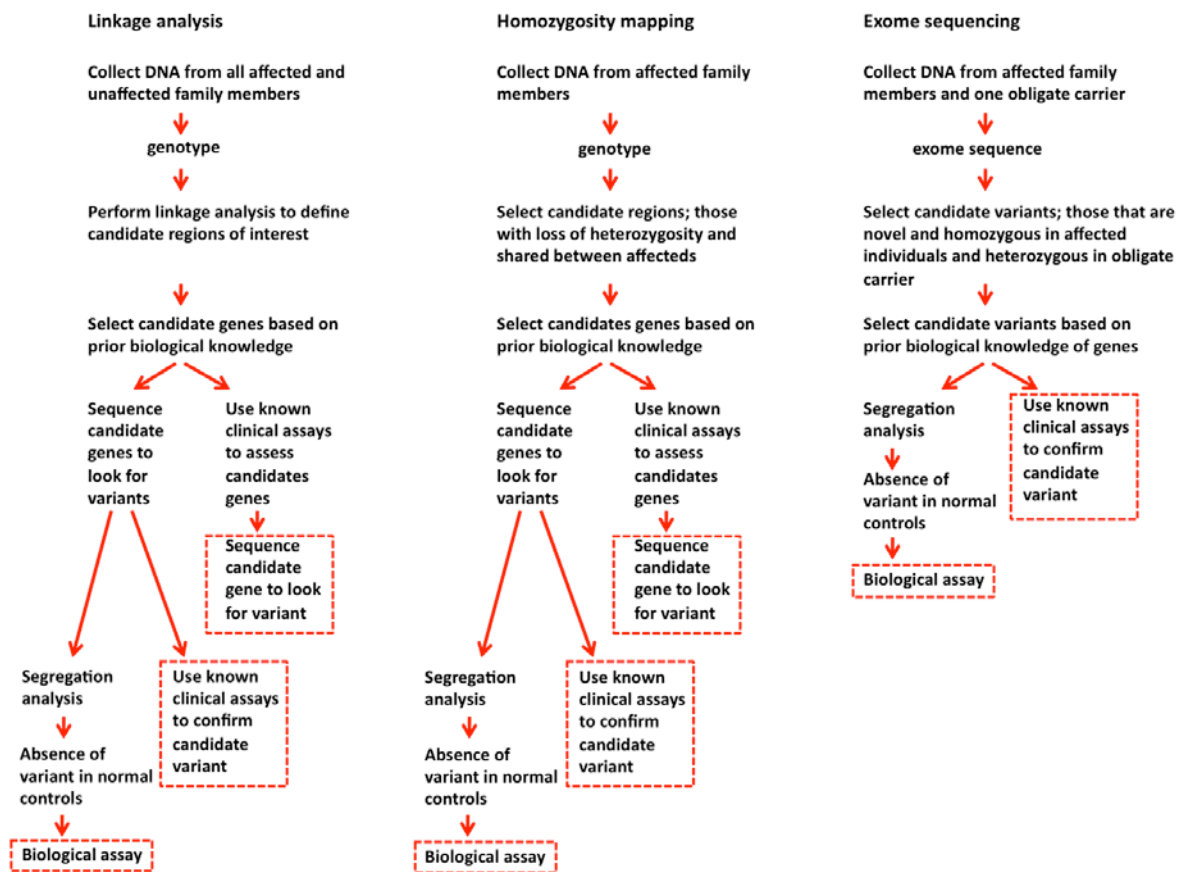
4.2 Strategy For Novel Gene Identification In A Proband With Gastric Adenocarcinoma And Proximal Polyposis Of The Stomach

To illustrate principles of using massively parallel sequencing in the setting of novel gene discovery in hereditary cancer syndromes, the preliminary work outlining the analysis of a family with an autosomal dominant susceptibility to GAPPS is given to demonstrate some key principles.

Fundic gland polyps occur in the fundus and body of the stomach and account for the majority of gastric polyp diagnoses. At the microscopic level, they are characterized by cystic dilated fundic glands, lined with flattened parietal, chief and mucous neck cells [362]. They are identified in 1-2% of upper endoscopies the general population, are 1-5mm in size and can be single or multiple but are usually less than 10 in number [363]. Gastric fundic gland polyposis (>10 lesions) can occur sporadically in association with exogenous factors such as chronic proton pump inhibitors use (PPI), occurring in 7.3% of those on PPI therapy [363, 364]. Alternatively they can occur due to a genetic predisposition. Gastric polyposis occurs in FAP, where fundic gland polyposis is common [365]. Malignant progression of sporadic fundic gland polyps are uncommon, whereas malignancy can arise in the setting of fundic gland polyps caused by germline mutations in *APC* [366-369]. Differences between syndromic and non-syndromic fundic gland polyps can be characterized at the molecular level, where somatic alterations of *APC* occur more frequently in FAP related fundic gland polyps [370], and conversely, somatic mutations in beta-catenin occur more frequently in the sporadic setting [362]. Gastrointestinal poyposis can also occur in the setting of germline mutations in *MUTYH* associated polyposis [371], Peutz Jeghers[372], Generalised Juvenile polyposis[373], germline mutations in *BUB1B* [374], and germline mutations in *PTEN* [375]. Constitutional mismatch repair deficiency syndrome can have features similar to neurofibromatosis type 1 with early onset cancers and gastrointestinal polyps. More recently a 40KB duplication upstream of *GREM1* was found to be the underlying cause of the colorectal polyposis syndrome, hereditary mixed polyposis syndrome, in some Ashkenazi Jewish families [200], although the full phenotypic spectrum, including gastric manifestations is yet to be delineated. Hamartomatous polyps

occur in Tuberous Sclerosis and NF1, although these syndromes do not have reported increases in stomach cancer. Gastric adenocarcinoma and proximal polyposis of the stomach, has been identified as a new autosomal dominant gastrointestinal polyposis syndrome [376], however unlike sporadic fundic gland polyps and those associated with germline mutations in *APC*, there is a significant risk of dysplasia and subsequent gastrointestinal adenocarcinoma in these individuals. To date the underlying genetic susceptibility has not been identified.

To elucidate the underlying germline susceptibility to the GAPPs phenotype we studied the DNA extracted from the germline and a polyp of an individual (II-2) with a personal and family history of GAPPs (Figure 21). The family history (Figure 21) includes three siblings all affected with the GAPPs phenotype. The proband of the family was diagnosed with IGC and cystic fundic gland polyposis at 34 years of age (II-1). Her diagnosis was preceded by symptoms of reflux treated medically for one year. Her *H. pylori* status and colonoscopy status are unknown. The proband's father (I-4) was diagnosed with malignant lymphoma, lymphoplasmacytic type, IgM Lambda and treated for two years with chlorambucil prior to developing a gastroesophageal junction tumor at 50 years of age. The tumor was a moderately differentiated adenocarcinoma described as "polypoid", however the stomach was not assessed as the tumor blocked entry. There was no history of proton pump inhibitor use and his *H. pylori* status is not known. Following the diagnosis in the proband, the siblings, II-2 and II-3 were found to have the phenotype of profuse fundic gland gastric polyposis with antral sparing and no polyposis of the colon. This phenotype with an inheritance pattern consistent with an autosomal dominant Mendelian susceptibility has come to be known as GAPPs. More recently the mother was diagnosed with GC, type and presence of polyposis is not known.

Figure 21 Pedigree of a family with GAPPs
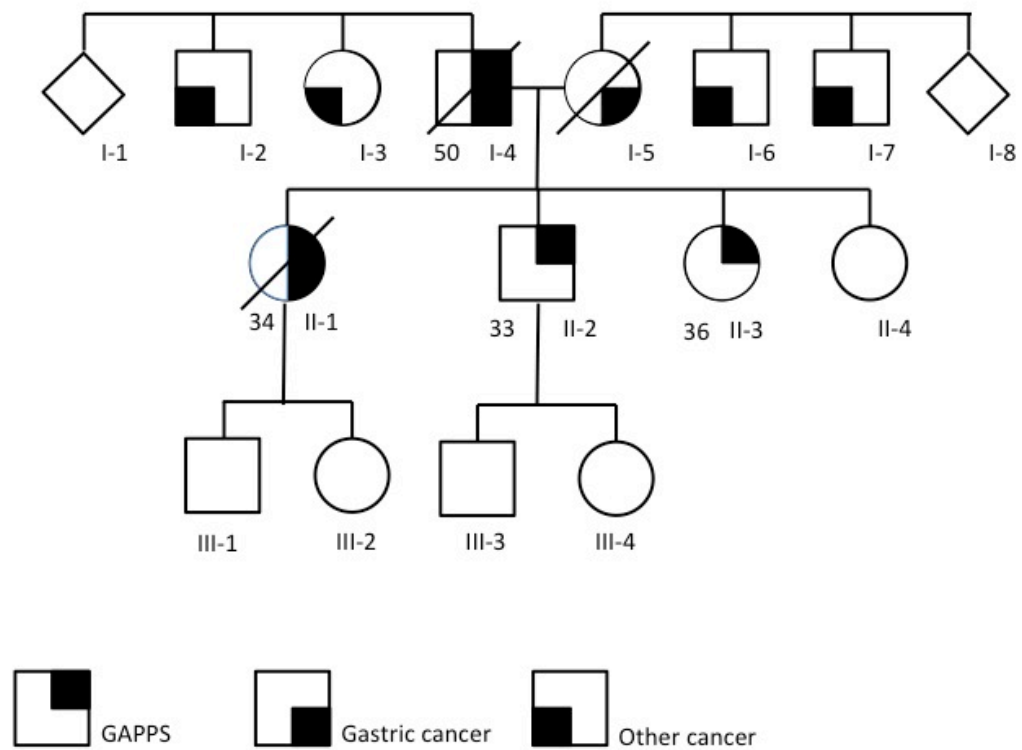
4.2.1 Strategy For Novel Gene Identification In A Proband With GAPPS

Based on the family's autosomal dominant susceptibility pattern, a plausible hypothesis would be that the genetic predisposition is caused by a tumor suppressor gene. Thus exome sequencing of individual, II-2, in addition to exome sequencing of one of their frozen polyps, was undertaken in addition to studying high-density SNPs arrays. Following that reasoning, Knudson's two hit hypothesis would dictate that, an identifiable second hit in the causative gene may occur in the fundic gland polyps, as seen in fundic gland polyps associated with FAP [370]. As demonstrated by Abraham et al, fundic gland polyps are clonal, largely benign neoplasms, whereby; within a patient there can be different molecular lesions acting as second hits seen in different fundic gland polyps [370]. This is important to note, as examining second hits can help define the regions or genes that function as tumor suppressors or drivers for the neoplasms. This principle could be extended such that one could consider discrete clonal fundic gland polyps to be considered as multiple primary neoplasms, each requiring independent second hits meaning that, sequence analysis of discrete fundic gland lesions, can add further evidence to support a potential tumor suppressor role of a germline mutated gene shown to be recurrently somatically mutated in different lesions within in the same patient. This larger experiment, comparing the second hits of multiple fundic gland polyps in an individual could provide a powerful filtering strategy to hone in on recurrently mutated genes in the polyp in a gene already mutated in the germline and could be further refined to first rule out those sharing the same type of second hit (Figure 22). This investigative strategy can be applied to any scenario with multiple primaries in an individual, although in the case of neoplastic polyposis syndromes, the multiplicity of discrete clonal events, provides a unique opportunity to perform multiple comparisons, filtering, at the gene level by cross-referencing germline and recurrently somatically mutated genes or at least narrowing down to the candidate regions through analysis of somatic CNAs and regions of LOH that have retention of the germline variant, thereby increasing the power to find the candidate gene (Figure 22).

Figure 22 Use of multiple discrete clonal neoplasms to narrow candidate germline variant list.

Novel germline variants:
gene a (region A)
gene x (region X)
gene m (region M)

Exome sequencing and high-density SNP array analyses of an individual (blue circle) and their multiple neoplasms (red circles) to compare recurrently somatically mutated genes and germline variants

○ LOH of regions A, B and C
○ LOH of regions A, D and E
○ Frame shift in genes a, f and p
○ Frame shift in gene a
○ Splice site variant in gene a
○ No variants

candidate gene a, within region A, (supported by 5 independent genetic events)

Furthermore this strategy can be applied to a single individual and their multiple neoplasms, as opposed to starting with multiple individuals from a single family whereby the amount of variant sharing increases with the decreasing degree of relatedness. In the situation where there are only first-degree relatives available, it can be extremely difficult to narrow down the shared candidate variants. Ultimately both strategies can be used in parallel, as sequencing of relatives also gives important information regarding segregation of the candidate germline variants with the disease phenotype. With that said, segregation analysis of candidate variants can also be achieved during the validation stage, where it can be performed in a high-throughput manner. Classically, linkage analysis of large families has been utilized to identify Mendelian disease genes. With new sequencing technologies there has been the hope that smaller families with a clear phenotype can also be informative. Although the value of evidence in support of multiple families harboring mutations in the same novel susceptibility genes is undisputed, unless clear linkage data is available, using multiple families to study a phenotype may sometimes confound the analysis due to genetic heterogeneity (Figure 23a). Furthermore even looking for shared variants within a family between multiple affected individuals can also be confounded by phenocopies (situations in which the observed phenotype is not related to the inheritance of an underlying susceptibility) (Figure 23b). Filtering variants seen in unaffected individuals is also sometimes undertaken however can also confound analysis due to non-penetrance. Therefore, studying the second hits on the same genetic background e.g. in the polyps of an individual, avoids issues of genetic heterogeneity and phenocopies. As it is possible that factors producing methylation as a second hit or widespread upregulation of post-transcriptional silencers, may produce a field effect of rapid development (or potential regression of polyps once the instigating factor is no longer present), a targeted examination for e.g. methylation patterns in tumors of patients following this phenotype, may reveal a unified second hit.

Figure 23 Potential confounders to the strategies of multiple individual testing to look for shared variants and genes.

A

Experimental Design

Exome sequencing of multiple affected families to compare shared genes recurrently mutated

B

Exome sequencing of circled individuals to compare shared variants

Potential confounders

Genetic heterogeneity for phenotype caused by genes a, b, and c.

* Actual mutation carriers
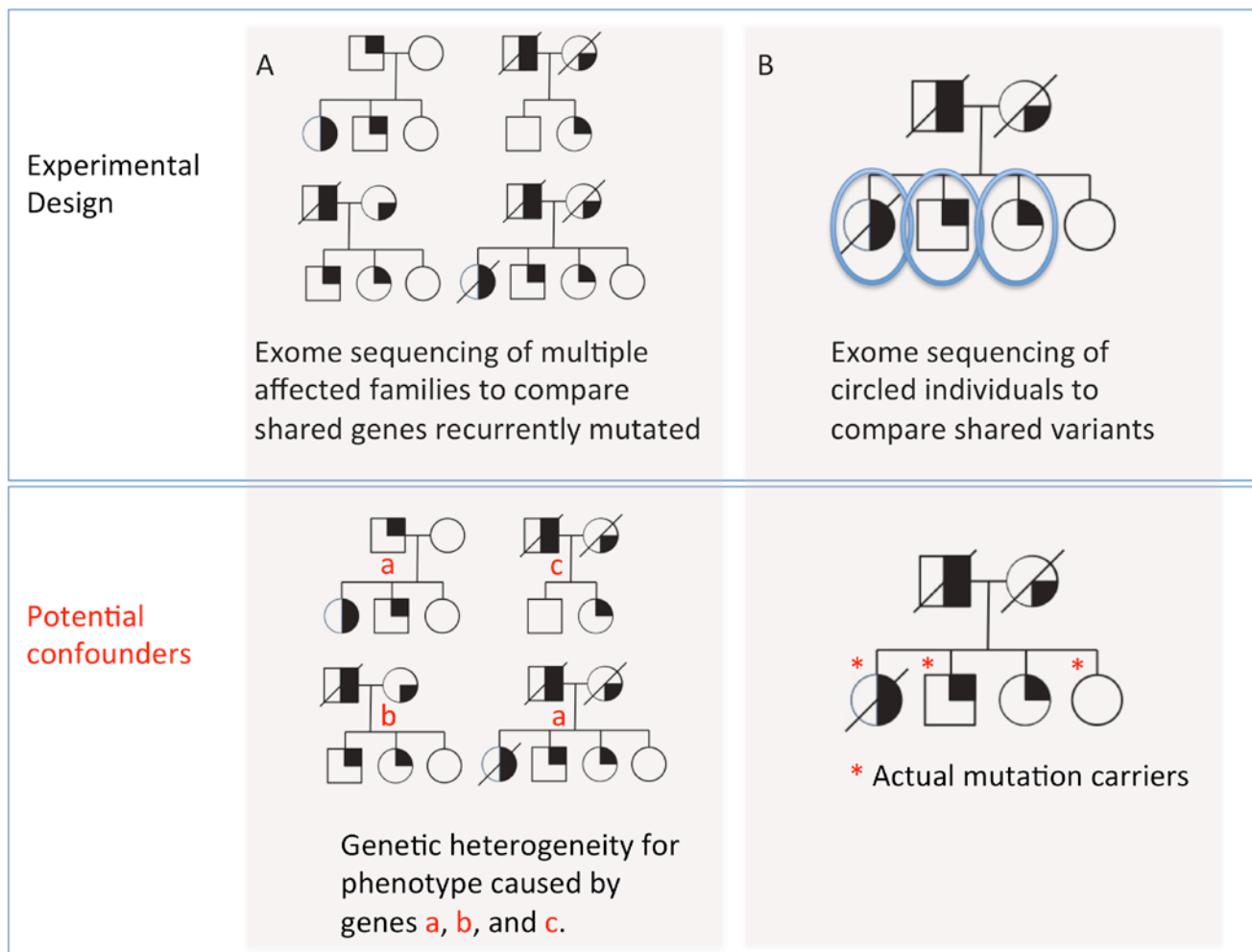
As a start, to look for somatic aberrations indicative of tumor suppressor genes in an unbiased manner, exome sequencing of the polyp was performed to look for canonical somatic mutations (SNVs, indels, and splice-sites) and Affymetrix 6.0 SNP arrays were used on the normal and tumor DNA to assess for CNAs and regions of LOH, indicating the presence of a germline aberration.

DNA samples were obtained from the proband's blood and frozen polyp tissue and extracted in sufficient quantity and quality to be submitted to the Genome Science Center, Vancouver BC, for exome capture and sequencing. Genomic DNA was sonicated, ligated with blunt end adapters, amplified and selected for optimum DNA size fragments and then subjected to targeted whole exome capture using the Agilent SureSelect 50MB system [339]. Targeted exome capture was by RNA-baited streptavidin labeled magnetic beads hybridized with the DNA and eluted using a magnet[339]. RNA was then digested prior to another amplification step[339]. Short paired-read (75bp) sequences obtained from the Illumina Genome Analyzer II were mapped to the reference human genome (NBCI build 36.1, hg18) using Mapping and Assembly with Quality (MAQ) (version 0.7.1) in paired-end mode [340]. The Sequence Alignment/Map (SAM 0.1.7) format was used for downstream processing. Insertion and deletion (indel) information was extracted from the alignment data using the Samtools package [341]. Variant detection was by SNVMix [342]. Variants were enriched for novel non-synonymous, and indels, by filtering with those already present in dbSNP130, The 1000 Genomes Project, and in in-house exomes. Analysis of germline splice sites and copy number variants for the somatic polyp data and splice sites, indels and copy number data inferred from the exome data will need to be investigated as these are critical pieces of information representing canonical categories of mutations..

Germline variants were further filtered by variants seen in the NHLBI Exome Sequencing Project (ESP) ESP6500 dataset, comprised of a set of 2203 African-Americans and 4300 European-Americans unrelated individuals, totaling 13,006 chromosomes, taken from 12 disease cohorts with a range of heart, lung or blood disorders, Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [EVS Release Version: v.0.0.14. (June 20, 2012)]. This resulted in a candidate list of novel germline variants. Germline copy number

variants were also assessed with Affymetrix SNP 6.0 high-density SNP arrays with data analysis using the Affymetrix genotyping console.

To assess for sequence variants, CNAs and areas of LOH potentially indicative of regions containing a tumor suppressor genes, frozen polyp DNA was also subjected to exome sequencing and Affymetrix SNP 6.0.

4.2.1.2 Results

Following bioinformatic filtering with 1000 genomes, dbSNP130, Exome Variant Server and an unrelated in-house exome not known to have gastric polyps. A first pass was taken to identify mutations most likely to be pathogenic secondary to causing protein truncation (Table 9).

Table 9 Novel germline truncating variants

Mutation	Comment	AA variant	Gene	Location	Uniprot	Refseq
1,3321027,A,T	Stop codon	K883	PRDM16	1p36.23-p33	PRD16_HUMAN	NP_071397
12,5931881,G,T	Stop codon	Y2684	VWF	12p13.3	VWF_HUMAN	NP_000543
12,68351562,C,T	Stop codon	W338	VMD2L3	12q14.2-q15	BEST3_HUMAN	NP_116124
16,31067468,C,T	Stop codon	W101	PRSS36	16p11.2	POLS2_HUMAN	NP_775773
2,176666269,C,A	Stop codon	Y135	HOXD13	2q31.1	HXD13_HUMAN	NP_000514
2,196454856,G,A	Stop codon	R1957	DNAH7	2q32.3	DYH7_HUMAN	NP_061720
X,2843605,C,T	Stop codon	W331	ARSD	Xp22.3	ARSD_HUMAN	NP_001660

Copy number (CN) analysis was also undertaken to identify germline copy number variation, such as loss or gain that could potentially disrupt a gene. This revealed areas of increased and decreased copy number (Table 10), none of which were novel as they shared regions previously reported and accessible through the Database of Genomic Variants (http://projects.tcag.ca/variation/).

Table 10 Germline Copy Number Variation (greater than 100 Kb).

Copy Number State	Loss/Gain	Chr	Locus	Size(kb)	#Markers	CNV_Annotation
4	Gain	2	p11.2	365	136	Variation_0525
4	Gain	3	q26.1	113	69	Variation_0563
3	Gain	4	p11	367	18	Variation_0572
3	Gain	4	q35.2	151	28	Variation_104837
3	Gain	7	q11.1	773	33	Variation_105020
4	Gain	7	q11.21	150	10	Variation_105021
1	Loss	7	q35	155	32	Variation_0726
3	Gain	8	p23.3	125	11	Variation_105084
1	Loss	8	p11.22	153	58	Variation_103790
3	Gain	9	p13.1	2256	38	Variation_0302
3	Gain	9	p11.2	1634	97	Variation_0744
3	Gain	9	q13	1521	60	Variation_0750
3	Gain	9	q21.11	430	28	Variation_2144
1	Loss	10	q11.22	282	111	Variation_0307
3	Gain	11	q13.4	176	53	Variation_30559
3	Gain	12	p11.1	127	51	Variation_113307
3	Gain	12	q11	183	39	Variation_103980
3	Gain	14	q11.1	1103	43	Variation_0812
3	Gain	14	q32.33	704	309	Variation_0182
1	Loss	15	q11.1	361	42	Variation_0318
1	Loss	15	q11.1	2001	285	Variation_0318
3	Gain	15	q25.2	264	33	Variation_0460
3	Gain	16	p11.2	1857	202	Variation_0195
3	Gain	17	q21.31	357	109	Variation_0327
3	Gain	17	q24.1	149	22	Variation_2231
1	Loss	19	p12	133	78	Variation_104167
3	Gain	21	p11.1	233	75	Variation_103034
3	Gain	22	q11.22	223	130	Variation_0334
1	Loss	X	q28	232	38	Variation_0825

Somatic variants from the high-density SNP array data and exome sequence data were inferred by subtraction of the germline events.

Affymetrix SNP 6.0 analysis of the polyp DNA revealed three somatic CNAs (>100kb), all resulting in only a single copy number of the region. There were 15 regions of somatic copy number neutral LOH affecting greater than 2 markers, none of these overlapped with germline variants.. Comparisons of the germline variants and somatic regions affected are presented in Table 11.

Table 11 Somatic aberrations were cross-referenced with germline events.

Germline indels	Germline SNVs	Gene	Locus	Somatic SNV (polyp)	Somatic Copy Number Loss
	1,12778409,G,A;A368T 1,12778521,G,A;S405N 1,12778583,A,G;N426D	PRAMEF1	1p36.21		chr1:12,845,851-12,998,268
	1,12841521,T,G;S24A 1,12841533,A,G;M28V 1,12842191,G,T;R115I 1,12844066,C,G;R424G 1,12844072,A,G;N426D	PRAMEF2			
	1,12875419,G,T;P447H 1,12875420,G,T;P447T	PRAMEF10			
1:74810445,*/-CCT		C1orf173	1p31.1	1,74809690,A,C;V1431G	
2:21078316,*/-CTG	2,21080726,A,G;p.S4003P 2,21113552,T,A;p.E206D	APOB	2p24-p23	2,21079402,C,G;A4133P 2,21082588,C,T;E3553K 2,21111213,T,G;N295T	
	2,29771242,C,A;p.E310D	ALK	2p23	2,29793959,C,G;R259P	
	2,71444804,G,A;p.R544Q	ZNF638	2p13.2-p13.1	2,71430249,C,A;Y219	
	2,74513177,T,G;p.N79T	RTKN	2p13.1	2,74508074,G,C;L414V	
	2,220077136,C,T;p.P193S	GMPPA	2q35	2,220078489,C,G;A274G	
2:61214829,*/-G		KIAA1841	2q14	2,61152321,C,G;A76G	
	3,139885254,A,C;p.V794G	PIK3CB	3q22.3	3,139866717,A,C;C841W	
		UGT2B17			chr4:69,375,336-69,489,323
	5,54671629,C,T;p.R184C	SKIV2L2	5q11.2	5,54737161,G,T;C878F	
	5,149735127,A,C;p.K507N	TCOF1	5q32-q33.1	5,149728515,A,C;H141P	
6:152816379,*/-G	6,152512312,C,G;p.R8212P 6,152739385,G,C;p.L3050V	SYNE1	6q25	6,152597570,C,T;E6819K	
	7,7386817,G,C;p.T774R	COL28A1	7p21.3	7,7364852,A,C;C1105W 7,7379433,T,G;T877P	
	7,100289817,T,C;p.L21P	SLC12A9	7q22	7,100294469,T,G;C278W 7,100292618,T,G;V214G	
	8,17961167,G,A;p.T335M	ASAH1	8p22-p21.3	22,21172694,T,G;T344P	
	9,32625510,G,C;p.S23W	TAF1L	9p21.1	9,32623720,G,C;P620A	
	9,125186013,C,G;p.Q526H	DENND1A	9q33.2	9,125454195,A,C;V179G	
	9,129191028,A,G;p.Y851C	GARNL3	9q33.3	9,129159367,C,G;P662A	
	9,138510934,A,G;p.S2360P 9,138524991,C,G;p.R892P	NOTCH1	9q34.3	9,138528936,T,G;N685T	
	10,129794644,C,G;p.R1817T	MKI67	10q25-qter	10,129794674,T,G;D1807A 10,129803431,A,G;S411P	
19:53426829,*/+TT		CARD8	19q13.32	11,46794409,A,C;V125G	
	15,26129754,G,C;p.T1066S	HERC2	15q13	15,26148580,C,G;W333C	
	15,50315243,T,C;p.K960E	MYO5C	15q21	15,50317033,C,G;A936P	
	16,69566555,C,T;p.R1585H	HYDIN	16q22.2-q22.3	16,69483787,T,G;H3131P	
	19,43543293,A,C;p.H617P	C19orf15	19q13.1	19,43543033,T,G;V558G	
	20,40146778,T,C;p.E1387G 20,40261348,C,T;p.A835T	PTPRT	20q12-q13	20,40160530,A,C;V1286G 20,40168951,G,C;C1115W	

Technical validation of candidate variants will be using an orthogonal technology, such as Sanger sequencing. Biological validation of putative causal variants will be by segregation analysis in the related individuals with further validation through sequencing and genotyping other affected families, sporadic cases and controls.

4.2.2 Candidate Gene Identification

Further supportive evidence for causality such as biological plausibility, can be derived from the literature where a candidate gene may already have a defined role in sporadic fundic gland polyposis or intestinal-type stomach adenocarcinoma. Potential candidates for stomach polyposis can be gleaned from somatic mutation data from large sequencing consortiums such as the cancer genome atlas where potential drivers mutations recurrently mutated in numerous samples are identified.

Furthermore mouse models of polyposis may lend insights into potential candidate genes or related pathways [377]. The analysis strategy outlined in this chapter is just one way in which analysis of the tumor in parallel with the germline DNA, can help to narrow the list of potential candidate germline susceptibility genes. Methods to define relevant regions of the genome, like linkage analysis in large families with multiple affected individuals can further aid in analysis, although, this relies on adequate phenotyping of affected family members. As previously mentioned, analysis of a patient's tumor in parallel with their germline DNA, theoretically has multiple benefits over analyses between germline DNA of different affected individuals. These benefits are further amplified by the analysis of multiple primary tumors from a single individual. As technology is continuing to improve, the ability to undertake genome wide analysis on diminishing amounts of DNA is becoming feasible and will more readily enable tumor-normal genomic comparisons from archival tissue specimens including biopsy material. To broaden the applicability of this type of discovery approach, Figure 24 outlines a general strategy for the detection of novel cancer susceptibility genes, taking into account genetic heterogeneity by considering genes in shared biological pathways and drawing support for causality of candidate variants by comparison with mutation data generated from the corresponding type of sporadic cancer .

The presence of genotype-phenotype correlations identified in the germline or sporadic setting, helps direct targeted germline testing of specific subtypes of sporadic disease and

in the recognition of the rare families that segregate inherited germline variants in the corresponding genes. Defining the basis of germline cancer susceptibility in individuals is directly translatable to the clinical care of a patient and their family. Once a germline mutation is identified, new opportunities become available to the patient and family, such as risk-assessment using targeted mutation testing for carrier status, preimplantation genetic diagnosis, the potential for high-risk screening for early diagnosis and targeted treatment strategies.

Figure 24 A general strategy for identification of novel cancer susceptibility genes.

Generation of variant calls:

- Germline familial cases with paired tumor
- Whole genome sequencing of both germline and tumor
- Mapping and alignment of raw reads to reference genome
- Removal of duplicates and poor quality reads
- SNV, indel, structural rearrangement calling

Variant list (germline)

Variant list (somatic)

Selection criteria for candidate variants:

- Not present in 1000G and NHBLI EVS with a minor allele frequency >1%
- Recurrent somatic alteration; variant, gene or pathway

candidate variants
Tier 1: nonsense, splice variants, indels, structural rearrangements
Tier 2: missense

Support for causality:

- Mendelian segregation with affected family members
- Further tumor studies; transcriptome analysis
- Recurrent germline alteration of; variant, gene or pathway in unrelated families and sporadic cases
- Recurrent somatic alteration of; variant, gene or pathway in somatic cancers

CHAPTER 5: Conclusion: Next-Generation Sequencing In Cancer Research And Clinical Care

Cancer is a disease of the genome characterized by the accumulation of somatic genetic and epigenetic alterations. In light of this, oncology and personalized cancer treatment is positioned at the forefront of the clinical application of next-generation DNA sequencing. Research is rapidly advancing because cancer is common, tissue diagnosis (resulting in acquisition of specimens) is requisite and the identification of the tumorigenic properties of a cancer can have immediate treatment implications.

5.1 Targeted Therapy

The ability to therapeutically target the products and pathways of tumor specific mutations has provided the incentive to determine the genetic changes that drive tumor growth and progression.

5.1.2 Targeted Therapy In Sporadic Cancer

Personalized medicine is already practiced in oncology where the decision to use targeted therapies is based on the presence or absence of the targets or resistance mechanisms that indicate whether or not the course of treatment will be of benefit. Examples of targeted therapies in clinical use are numerous and the list is expanding. They include monoclonal antibodies or small molecule inhibitors for directed towards targets such as HER2 overexpression in breast cancer [378], the Philadelphia chromosome in chronic myeloid leukemia and acute lymphoblastic leukemia[379, 380], PDGFR gene rearrangements in myelodysplastic syndrome [381], c-Kit positive gastrointestinal stromal tumors [382], EGFR and ALK overexpression in non-small cell lung cancer [383, 384], EGFR-expressing KRAS-

mutation negative metastatic colorectal carcinoma[385] and unresectable or metastatic BRAF V600E positive melanoma [386]. These advances have not only proven beneficial for the treatment of sporadic cancer but the concept of targeted therapy has implications for chemoprophylaxis and treatment in the hereditary cancer setting.

5.1.3 Targeted Therapy In Hereditary Cancer

Hereditary cancer is generally managed with increased screening and surveillance for disease and prophylactic surgeries. In FAP there has been benefit shown from taking sulindac and selective COX2 inhibitors and in Lynch Syndrome there is evidence to show benefit from taking Aspirin [387].

A relatively recent example of targeted therapy in the hereditary cancer setting is of polyADP-ribose polymerase inhibitors in *BRCA1* and *BRCA2* mutation carriers. PARP inhibition exploits the concept of synthetic lethality [388] by targeting a critical mode of DNA repair in tumor cells of *BRCA1* or *BRCA2* mutation carriers. In cancer cells of these individuals, there is an already compromised DNA repair system that has lost function of both alleles of *BRCA1* or *BRCA2*. By targeting the cancer cells with biallelic loss of *BRCA1* or *BRCA2*, only tumor cells are affected while normal cells with a remaining functional *BRCA1* or *BRCA2* allele are not affected.

Gorlin or basal cell nevus syndrome, caused by mutations in *PTCH*, predisposes individuals to multiple basal cell cancers. The hedgehog pathway is critical for normal patterning and development of the embryo. It responds to the hedgehog ligands that bind PTCH a transmembrane protein. When PTCH is unbound, it inhibits smoothened or SMO, causing downstream regulators of the pathway to inhibit expression of target genes. When PTCH is bound, inhibition of SMO is released which in turn causes downstream regulators of the pathway to induce hedgehog target genes [389]. The small molecule inhibitor of *SMO*, vismodegib, that targets the hedgehog-signaling pathway and

has been trialed in individuals with metastatic or locally advanced basal-cell carcinoma (BCC) [390]. Vismodegib is approved for use in the treatment of advanced BCC and has been shown to be of use in the prevention of BCCs in patients with Gorlin Syndrome [391]

Another example of targeted therapy is that of everolimus, a rapamycin analog, that targets the mTOR pathway and is approved for use in individuals with tuberous sclerosis with pediatric- or adult-onset subependymal giant cell astrocytoma (SEGA) and adult-onset renal angiomyolipoma [392, 393].

A further example of targeted therapy relates to the RET kinase inhibitor, vandetanib which has been approved to treat metastatic medullary thyroid cancer. Trials have been conducted and are ongoing assessing its efficacy in the hereditary setting [394, 395].

5.1.4 Research Use Of Next-Generation Sequencing Identifies Novel Drug Targets

The scope to identify and develop new drug therapies is increased with next-generation sequencing. Genome sequencing of a metastatic urethral neuroendocrine prostate adenocarcinoma and xenograft identified a homozygous deletion containing MTAP-CDKN2A. Treatment of the xenograft with a drug combination specifically targeting the defect in polyamine metabolism conferred by loss of MATP in the tumor cells was able to inhibit tumor growth without host toxicity [396].

5.1.5 Clinical Use Next-Generation Sequencing In The Diagnosis Of Known Drug Targets

In the diagnostic setting, massively parallel sequencing provides the ability to survey a cancer's mutations to identify a targetable genetic lesion that can alter treatment. This was demonstrated by whole genome sequencing of an acute promyelocytic leukemia, that did not have a clinically detectable X-*RARA* fusion, which revealed a cryptic insertion causing the *PML-RARA* fusion gene and allowed for targeted treatment [397]. Currently groups are examining the practicalities with regard to integration of clinical WGS in a timely manner such that results can influence treatment decisions [398]

5.2 The Use In Clinical Diagnostics

Next-generation sequencing in the oncologic diagnostic setting will facilitate the identification of therapeutically targetable genetic lesions, but will also uncover germline disease susceptibility alleles. This will undoubtedly reveal cancer susceptibility variants that will expand the workload of hereditary cancer programs and also general genetics by the discovery of variants in non-cancer related disease susceptibility genes. Therefore the introduction of massively parallel sequencing into state-of-the-art oncology will necessitate a transition period.

5.2.1 Variant Annotation Of Tumor Genomes By Next-Generation Sequencing

Next-generation sequencing studies of the tumor and germline DNA also provide the opportunity to look for and consider both known and novel germline cancer susceptibilities.

The utility of massively parallel sequencing has been demonstrated in the research setting. It has enabled rapid data acquisition regarding the mutational landscapes of many cancers. These data have been generated en mass by large consortium efforts such as The Cancer Genome Atlas Network [399, 400, 400-402] that has undertaken an integrated approach to analyzing genomes, epigenomes and transcriptomes to more fully understand genetic variation, gene expression and development of disease. Other independent efforts have defined the mutational evolution of individual's specific cancers by analysis of multiple tumor specimens biopsied at different clinical stages. These studies have provided valuable insights into the molecular changes that occur from the primary tumor to metastatic or relapse disease [403-405]. The ability to sequence at high depth at multiple sites in the primary and metastatic specimens enables characterization of tumor heterogeneity. Furthermore the ability to analyze single nuclei in this respect have broadened our understanding of the different types of mutational evolution in a tumor [406].

Comparison of mutations and mutation frequencies in the primary and metastatic disease allow for inference of early mutations constituting the clonal population and subclonal populations that survived treatment or developed a growth advantage. These studies have supported the idea of branch evolution within a tumor where different regional areas of the tumor have both common "trunk" and unique mutations indicating a common clonal evolution with branches of different subclones [407]. In some cases the branches demonstrate convergent evolution with different types of mutations occurring in the same genes [408-410]. While there is an emphasis on studying tumor heterogeneity and pattern of tumor evolution within cancers in order to identify the potentially therapeutically targetable driver mutations [407], this type of analysis also provides insight into critical genes and pathways that may act as susceptibility genes when mutated in the germline.

Early and late driver mutations are likely essential for tumorigenesis and cancer autonomy and for the purposes of hereditary cancer research, defining the early recurrently mutated genes in sporadic cancers could give clues to those that confer germline susceptibility to those particular types. One could conceive that genes recurrently mutated early on in a tumor's mutational evolution are the most likely to represent potential germline susceptibility genes seeing as aberration of their function

results in stereotyped disease, similar to that seen in hereditary cancer. This can be postulated based on the results of tumor sequencing which show *VHL*, a known renal cell carcinoma germline susceptibility gene to be frequently mutated as an early event in the mutational evolution of renal cell carcinoma [409]. Germline *VHL* mutations cause the autosomal dominant, Von Hippel Landau Syndrome. By this reasoning, it would not be unexpected to find that germline *PBRM1* mutations cause a subset of familial clear cell renal carcinoma, providing they would be tolerated in the embryo. Another example of a cancer susceptibility gene mutated early in sporadic cancer is that of *CDH1* and LBC; early biallelic inactivation of *CDH1* is demonstrated by the loss of E-cadherin expression in *in situ* LBC neighboring invasive LBC lesions [411].

Thus, a novel approach to look for hereditary cancer genes could be based on a multi-regional tumor sampling strategy in a patient with a suspected hereditary cancer syndrome to determine the mutation evolution and therefore early clonal trunk mutations, most likely to relate to an underlying germline susceptibility. In cases where tumors are not available, potential germline line susceptibility variants could also be considered in relation to early mutations seen in sporadic tumor specific data.

5.2.2 Use Of Next-Generation Sequencing Data To Influence Treatment

Regardless of whether one sequences the germline or the tumor, there remains the potential to reveal medically actionable germline variants. These may be of immediate clinical importance such as those affecting drug metabolism or those predisposing to conditions that when combined with specific drug regimens may lead to adverse events.

Knowledge of whether a person would have less side effects with a certain drug or could potentially have devastating adverse reactions will be critical to impart to the treating doctor in a failsafe manner likely with the addition of this information directly to electronic health records. Therefore, there will need to be a system in place to allow rapid access to relevant genomic information that will enable better health care.

There is also the question of how much information should be extracted from the data that may be relevant to the clinical scenario and what the onus is on the treating doctor and the molecular laboratory to pursue that information. Consider a situation where an individual is diagnosed with non-small cell lung cancer. The tumor genome is sequenced and among other things is found to carry an ALK-translocation. The treating doctor orders commences the patient on the ALK-inhibitor, crizotonib, however two weeks into the treatment the patient suffers a syncopal episode and is found to have ECG abnormalities consistent with long QT syndrome. Review of the patient's genome is positive for a mutation associated with long QT syndrome. Crizotonib's drug labeling has cautions that there may be "QT interval prolongation in patients who have a history of or predisposition for QTc prolongation, or who are taking medications that are known to prolong the QT interval, consider periodic monitoring with electrocardiograms and electrolytes". It is conceivable that in the future a question could potentially arise as to whether there is a duty of care for the laboratory or health care provider to look at potential susceptibility variants for long QT syndrome [412]in a patient about to start a medication associated with exacerbation of this syndrome.

An individual's germline variations and somatic tumor aberrations collectively influence the response to treatment and adverse events (Figure 25). Beyond, the influence of pharmacogenomics on active drug levels, or the possibility of adverse off-target drug effects, germline genetic factors can influence the immune response to cancer [413], the stromal environment supporting cancer growth[414, 415] or resistance to treatment[416, 417], and as demonstrated in high-penetrant hereditary cancer syndromes, the type of clonal somatic genetic abnormality occurring early in tumorigenesis (Figure 26). Bearing in mind the inherent genetic variation from one individual to the next, the overall response to therapy will differ between each patient (Figure 27).

**Figure 25 Consideration of genetic background in relation to treatment**
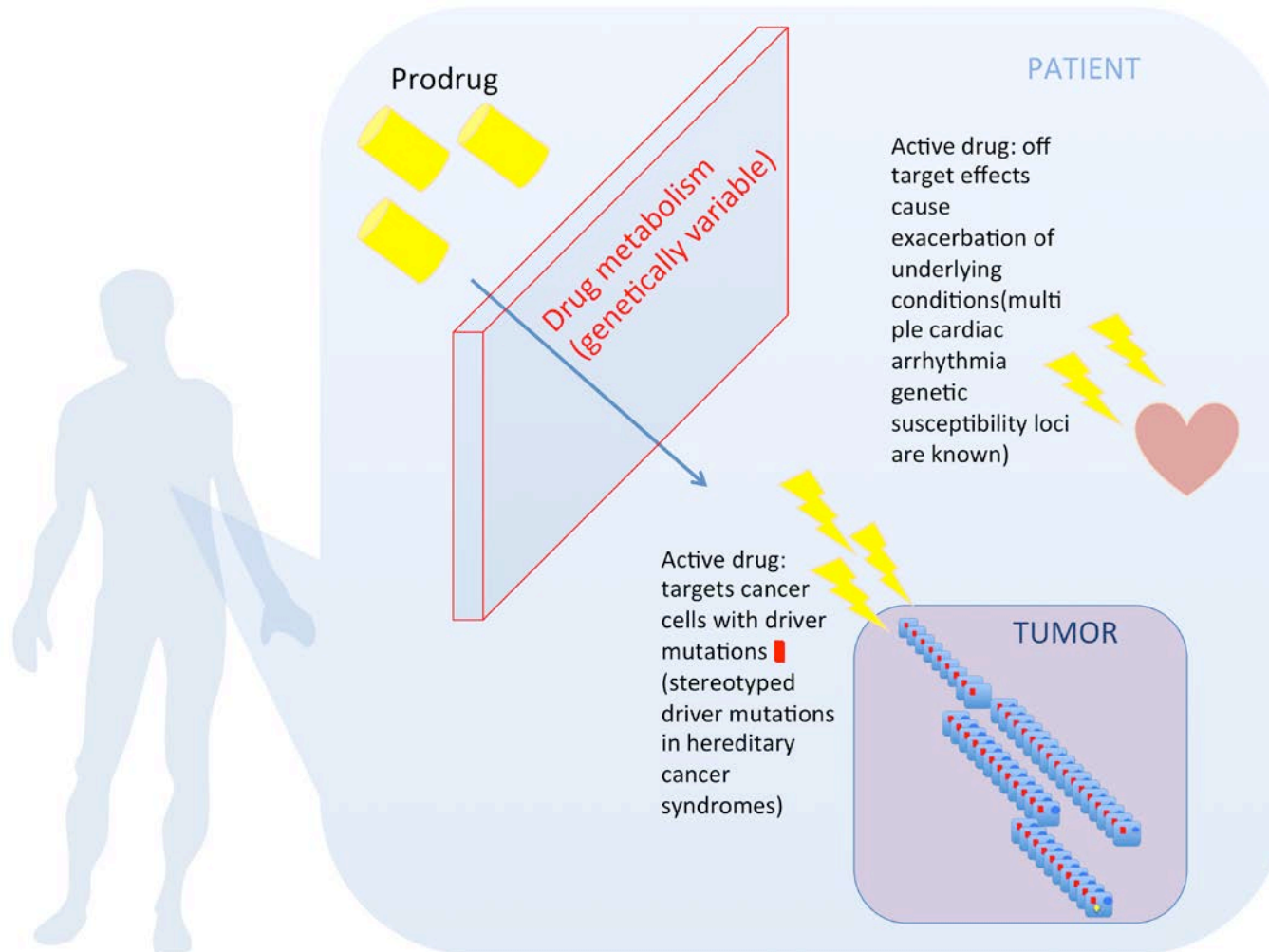
Response to treatment influenced by both somatic and germline genomic variation
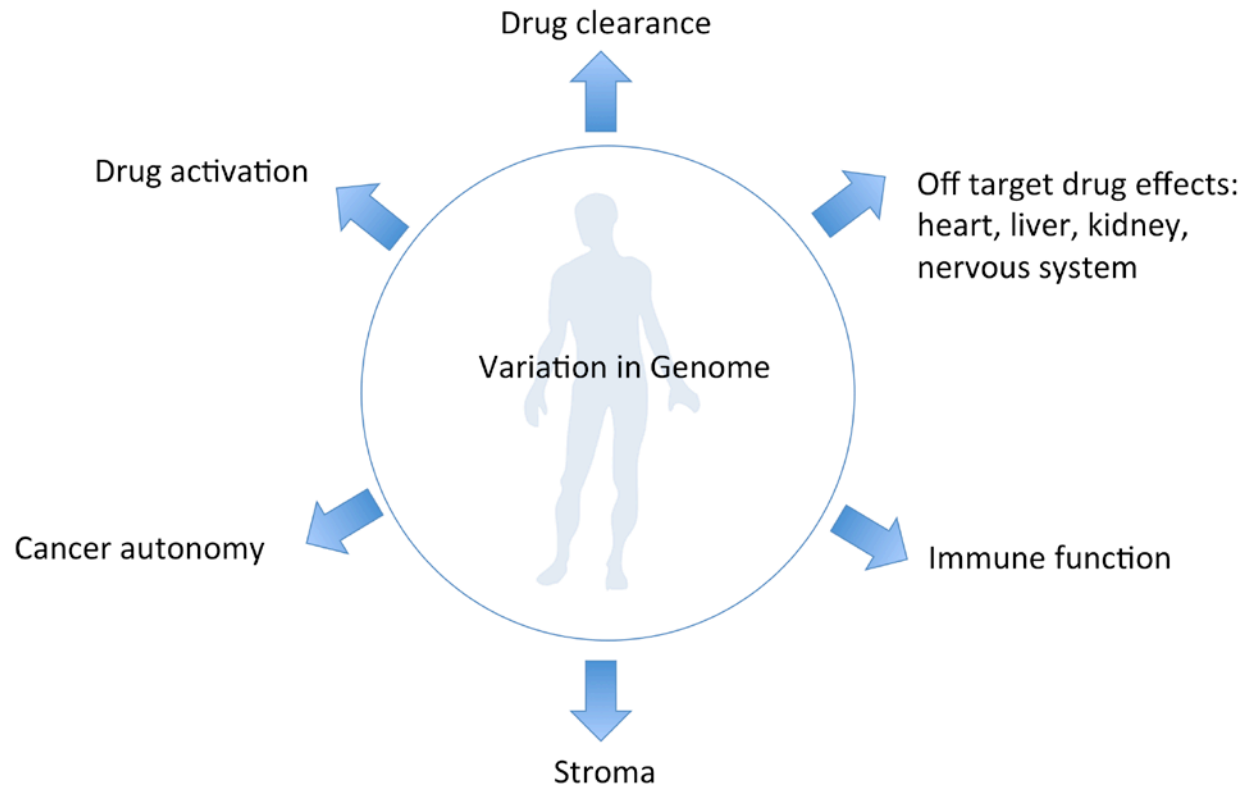
Figure 27 Individual variation affects response to treatment.
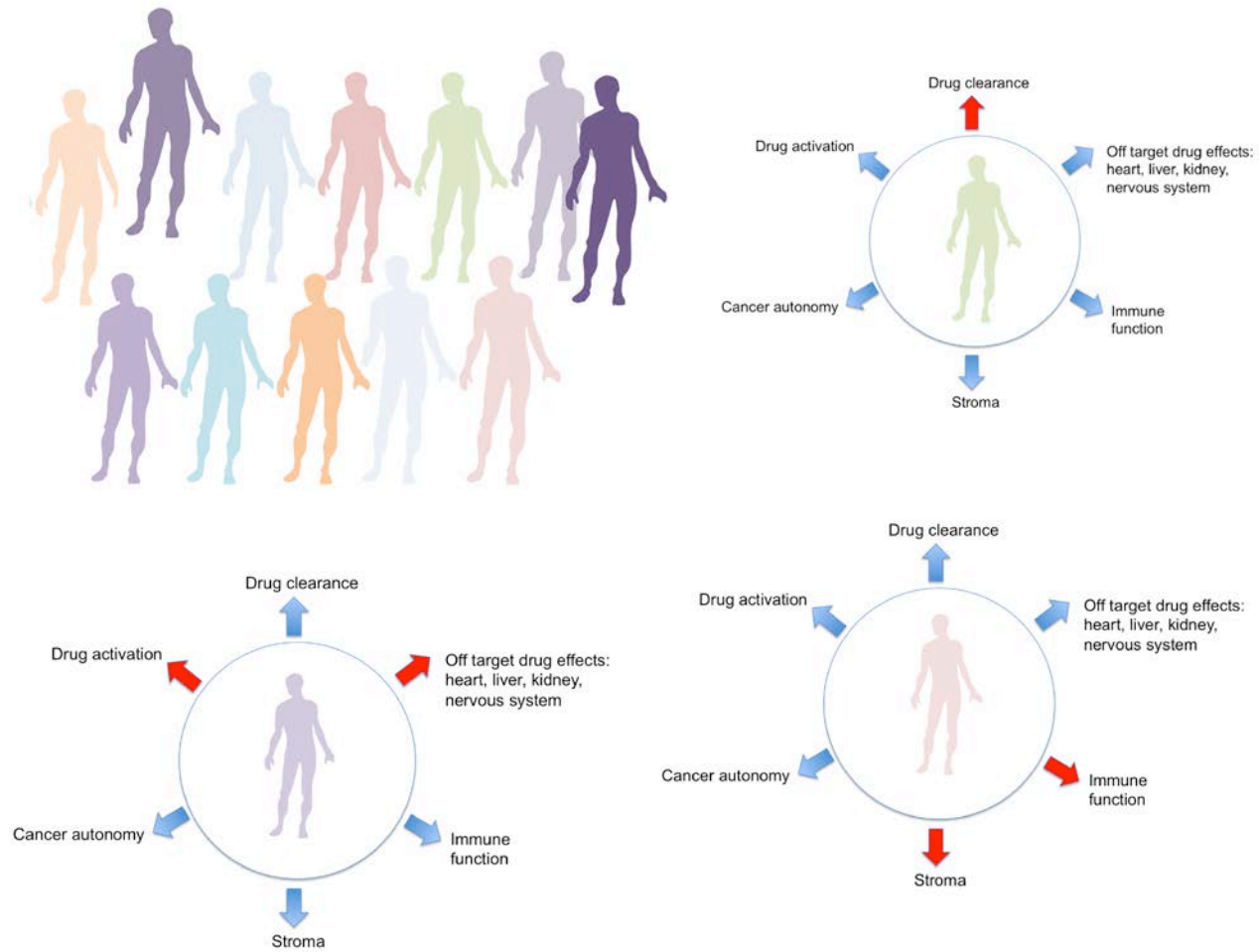
5.2.3 Practical Considerations

It has always been implied that one may learn the germline variants in an individual's genome while investigating those that are somatic, now that the technology has become more accessible, investigators are considering the practicalities involved in consenting patients regarding potential return of clinically relevant results. Of paramount importance is the need for informed consent that covers the possibility of the option for return incidental genetic information of clinical importance potentially arising from research or clinical care. Mardis and colleagues examined the germline and somatic genomes of an individual with multiple primary cancers and discovered a germline mutation in *TP53* [418]. Although the personal and family history was not consistent with the clinical criteria for Li Fraumeni, the need for genetic counseling was discussed and presumably the counseling was based on the known features of Li Fraumeni. In this context, counseling can be difficult as our knowledge and experience with management of many syndromes has been based on a biased ascertainment of the most extreme cases. With the increased sequencing of "sporadic cancers" it is likely more germline variation will be revealed which will give us a greater appreciation of the penetrance and variable expressivity of the known hereditary cancer genes. Determining germline susceptibility variants will be important however invariably there will be difficulties with regard to interpretation of each variant's pathogenicity and clinical relevance. Therefore widespread implementation of next-generation sequencing for clinical diagnostics will require a centralized database of clinically annotated genomic variants.

It could be perceived that the costs of personalized medicine will be significant with regard to widespread sequencing, bioinformatics analysis, interpretation, and validation of findings on an individual basis. Furthermore, other indirect costs related to incidental findings requiring subsequent diagnostic work-up, and cascade genetic counseling and carrier testing for family members, would also need to be considered. A cost-benefit analysis would need to weigh these costs against the benefits of more efficient targeted therapies resulting in decreased mortality (or more person-life years), less off-target side effects causing non-disease related morbidity and the potential for preventative medicine

and early disease intervention for incidentally; identified-disease susceptibility and newly diagnosed-disease in the individual or relevant to their potential offspring. Furthermore, by understanding genotype-phenotype correlations as they relate to pharmacogenomics, there may also be the potential to reinstate previously blacklisted drugs that have serious side effects in specific patient populations, but are otherwise safe and effective in others. These analyses will also need to take into account the transition period as the research and medical community start to build infrastructure and systematic integration sequence data into the medical record, in order to adopt routine large-scale sequencing. Sequencing costs will continue to come down and as the research and medical community streamline the costs of interpretation and validation, through efforts such as the aforementioned centralized clinically annotated centralized databases, it is likely that a better comparison between our current health care practices and a personalized-seqeuncing approach will be able to be made.

Whole genome sequencing and whole exome sequencing will essentially function as giant multiplex assays. Massively parallel strategies will be cheaper and more efficient than sequentially testing individual genes, especially in genetically heterogeneous conditions and circumstances where there is limited materials for genetic testing (Figure 28). If used in clinical diagnosis, these tests would infringe patents such as the one held by Myriad for the *BRCA1* and *BRCA2* diagnosis. Myriad's gene patents have been upheld in a recent ruling by the federal court of appeals and therefore companies that currently offer massively parallel multiplex testing of breast and ovarian cancer susceptibility genes testing in the United States, do not include *BRCA1* and *BRCA2* in their analysis. With that said, the Myriad patent will expire in 2015 and therefore the low cost of performing the sequencing will enable widespread testing. With the ongoing discovery of novel germline susceptibility genes the issue of gene patenting will remain topical.

Figure 28 In most diagnostic cases, current methods are iterative whereas massively parallel testing can be faster and more efficient when used in genetically heterogeneous conditions and in cases where limited material for genetic testing is available.

Proband with high-grade serous ovarian cancer (HGSOC)

Proband with family history ovarian cancer

Ovarian cancer 50 years old *limited tumor available

Breast cancer, no specimens
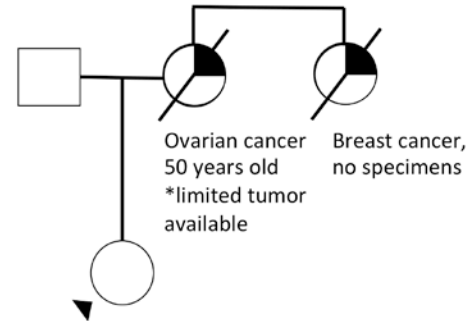
Massively-parallel multiplex testing of serous ovarian, fallopian tube or peritoneal carcinoma (Walsh et al.2011)

- BRCA1
- BRCA2
- CHEK2
- BRIP1
- PALB2
- BARD
- MRE11
- NBN
- RAD50
- TP53
- no germline mutations

6% of germline mutations in HGSOC not in BRCA1 and BRCA2
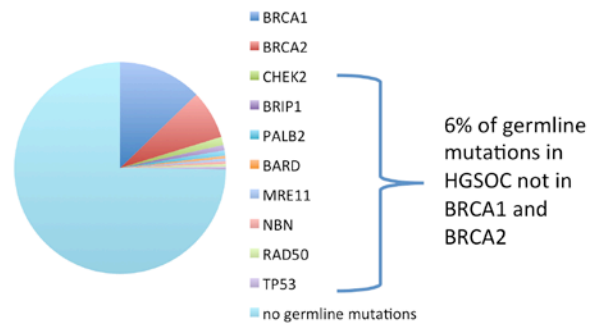
5.2.4 Clinical Practice And The Changing Landscape

What the treating cancer team has to offer their patients will be constantly changing. Accordingly, institutions and oncologists will need to build the capacity to recognize real clinical advances to effectively evaluate the progress in cancer treatments. In order to capture this information, all patients who consider or undertake genome-wide sequencing will need to be on a research protocol. Much of the research data will be generated at the time of the clinical test. There are many questions on the laboratory side as to the depth of coverage required to accurately assess the genomic and somatic genome in a clinical grade test. Questions will also arise with regard to data interpretation and the need for reinterpretation as our understanding of disease changes. Furthermore if the costs of data storage exceed the costs of resequencing using cheaper technologies, will it be simpler to resequence a patient rather than store their terabytes of sequence information. As genomic information will be increasingly be available for every cancer patient and will be relevant to ongoing care, there will be a need for stronger ties between the general clinic and clinical genetics services in order to make sense of the data in the context of hereditary cancer and in the context of the patients' other health issues more broadly.

In the future a typical oncology consultation will take into account a person's germline variants and their tumor variants at the primary diagnosis and at defined stages during the treatment in order to try to identify the driver mutations that require targeted therapy. Free communication between the general oncology service and clinical genetics will be critical as some of these patients will require further evaluation and counseling. There will need to be concerted efforts to understand the risk, benefits, clinical utility and consequences of the information we return to the patient, including their and their family's susceptibility to hereditary cancer. There will also need to be further discussion as to the deposition of results.

5.2.5 Future Implications For Translational Research

This thesis began by examining candidate genes and in particular germline mutations in *CDH1* and the association with LBC. It progressed to include new technology and demonstrated its use in the diagnosis of a Mendelian susceptibility as a proof of principle of how these techniques may be applied to single gene disorders, such as in the investigation of hereditary cancer. The rate at which sequencing costs are diminishing has exceeded expectations and there is excitement as to the broad applicability of this technology to research and clinical medicine. There is also trepidation with regard to the ethical considerations involved in determining potentially pathogenic variants in an individual beyond the primary purpose of the investigation.

With regard to the applicability of high-throughput sequencing to the questions first posed in this thesis, a complete characterization of the exome or genome-wide changes in a collection of individuals with familial LBC would have in addition to answering the question regarding the association of germline variants in *CDH1*, also have potentially highlighted novel susceptibility genes through an examination of rare novel variants occurring in the same genes shared between the individuals. Similarly, if characterization of the protein coding variants was performed in the cohort of women with ovarian cancer, it may have highlighted additional germline susceptibility genes. Furthermore a more, global analysis of the granular cell tumor in the individual with LEOPARD syndrome could have been undertaken and revealed somatic variants in related genes within the MAPkinase pathway that may have been required for tumor initiation. It is clear that there is utility in fully characterizing the genetic changes that are required to drive tumor pathogenicity. Zhang *et al*, showed that retinoblastomas have few somatic mutations in addition to biallelic inactivation of *RB1* [419]. Furthermore, the authors demonstrated the utility of the global analysis of tumors including investigation of the epigenome through chromatin immunoprecipitation, DNA methylation data, and gene expression data, to identify a key driver and novel therapeutic target, *SYK*.

In view of the ability to undertake global analysis of the germline and tumors of an individual, there is scope to move beyond the investigation of single susceptibility genes

to consider epistatic interactions and the modifying genes potentially relevant in determining penetrance. As discussed this is particularly relevant to CDH1 mutation carriers. The somatic genetic and epigenetic changes that are required to transform indolent lesions of DGC in *CDH1* mutation carriers into lethal disease are yet to be explored in depth. Whether progression may relate to inherited modifier genes that harbor germline variants also requiring biallelic inactivation in the tumor or whether it occurs on the basis of purely acquired mechanisms can be investigated through comparison of the variants in the tumors with the germline DNA. As touched upon in the example of retinoblastoma, a survey of gene expression and epigenetic regulation would also be informative. As discussed in the example of GAPPS, a comparison of multiple lesions on the same genetic background and in this case, multiple indolent lesions compared to those that progress would also provide critical insights into the progression of this lesions.

The ability to undertake genome wide analysis will continue to be useful, especially in the hereditary cancer setting. Although hereditary cancer only accounts for a small percentage of cancer cases, the knowledge gained from the study of the extreme cases will be broadly applicable. Many of the questions that have been posed to the scientific community relating to the lack of understanding around certain phenomena in cancer (http://provocativequestions.nci.nih.gov/rfa), can potentially be studied more rapidly in the hereditary cancer setting. By studying families, one can control for potentially confounding variables in light of the shared genetic background and environmental exposures. Furthermore, because germline variants make these individuals more susceptible to cancer, with one example being the increased frequency of GC in germline *APC* mutation carriers from Japan [125], this can highlight exogenous factors that are important for cancer progression.

5.2.6 Importance Of Pathologic Phenotyping In Disease Characterization And Recruitment Of All Patients Into Research Protocols

The strategy of targeted testing of genes known to be associated with particular hereditary malignancies has been proven to be effective and has enabled the implementation of surveillance and life-prolonging risk-reduction surgeries to those at highest risk. However, clinical genetics is changing. With cheaper DNA-sequencing technologies comes the advent of a new era of personalized medicine. The clinic will no longer rely on sequential targeted gene testing for a specific phenotype, instead there will be multiplex genetic testing. By default, this will enable less stringent selection of cases meeting formal testing criteria for each gene that will eventually broaden disease phenotypes by identifying more mildly affected individuals. Furthermore, as somatic tumor sequencing for prognostic and therapeutic purposes becomes mainstream, the germline variants will be revealed in the process. This information will have important implications for the discovery of new cancer predisposition genes and the rediscovery of known genes with variants that form milder alleles or result in different functional roles for the gene. As we gain a higher resolution picture of the rare variants in an individual, the collective impact of all of their variants (rare and common) may in part explain the variable presentations seen within and between families. It will be critical that individuals be enrolled on research protocols to collect phenotypic information and systematically accumulate knowledge regarding new genotype-phenotype correlations. This will be important in countering our current understanding of genotype-phenotype correlations that have been established through the biased ascertainment of extreme phenotypes that have historically defined the syndromes.

5.2.7 Ethical Considerations

Massively parallel DNA-sequencing has the capacity to produce substantial amounts of data within hours. Several thousand variants are seen in a single exome and several million in a whole genome. Despite the reality that the biological implications of the

majority of this variation are not known, for variants found in clinically significant genes that have previously been associated with high risks of disease, there is a general consensus that return of these be offered to the patient. Thus for presymptomatic conditions for which actionable measures exist, there is perceived clinical utility in identifying the associated disease susceptibility variants.

Based on this perceived benefit, in both the research and clinical setting, there is a growing body of literature that addresses the developing notion of "duty to warn" the DNA donor or the ordering clinician, of these variants [420, 421]. An National Institutes of Health committee examining the issue, placed the responsibility of returning incidental findings and research results to genomic research participants, with the biobanks storing their biological materials [422].

This introduces the important point that because of its unique variation, DNA is inherently identifiable. Therefore, DNA cannot be anonymized, but rather it can be "de-identified" (have traditional identifiers such as name, date of birth, address etc. removed), a common practice in research. Thus, the risk for "re-identification" of the individual from which the sample came, can occur in the setting of shared or publically available genomic data which if compared to genomic data from the same individual, can potentially result in the ability to determine whether or not an individual participated in a study [423]. In light of this, there exists the potential for privacy breach and genetic discrimination. Furthermore, these implications extend beyond the individual to their families. Importantly, in recent years, the Genetic Information Nondiscrimination Act of 2008 (GINA) was enacted into the federal legislation of the United States, prohibiting group and individual health insurance plans and employers from discriminating against individuals based on genetic information. GINA does not cover life insurance, disability and long-term-care insurance and does not protect certain groups of individuals.

For the purposes of genomic research, informed consent should clearly state the potential for breach of privacy of data. Measures directed towards protecting privacy e.g. removal of traditional identifiers from biological samples can be taken, although since the ability to link genomic research data with the individual remains, breach of privacy can not be guaranteed. Furthermore, as there is usually a necessity to share sequence data with other researchers, either in a collaboration or though public deposition of the data, there is always the possibility that a research participant may be identifiable. Furthermore, although there is always the option to cease participation in a study which is usually accompanied by destruction of all study related records and biological materials, once data has been disseminated, it is hard or next to impossible to retract. Currently there exist restrictions to data access which try to ensure that only bona fide researchers are able to use it, however this does not address the primary issues of why there need to be these precautions in place. A potential solution is to put in place formal deterrents such as higher level regulations that make it illegal to maliciously re-identify de-identified genomic data without the person or their proxy's consent.

As eluded to, embedded into research and clinical consenting needs to be an understanding that genomic analysis may reveal information that has potential clinical significance. Thus research consents need to state the potential for identification of findings that the research was designed to investigate, and secondary findings, unrelated to the research question. Both of these findings may have clinical significance. The return of which should be optional. Likewise, clinical testing by DNA-sequencing should also address the identification of secondary findings and variants of uncertain significance [American College of Medical Genetics WGS policy statement 2012]. Various models have been proposed to consent individuals for return of results ranging from tiered approaches, to binning, to full return [424-426]. Regardless of the approach used, findings need to be scientifically valid, have high positive predictive value and be tested and reported in compliance with state regulations [422].

Many questions as to the range of impacts of returning high volumes of genetic results to the patient or research participant will be considered in coming years as clinical and research programs undertaking genomic analysis gain experience in the practice. Factors to be monitored include whether exposure to such results induces changes in health behaviors; whether and what kind of utility can be gained from results that are not medically actionable; and whether results cause distress or stimulate demand for psychological care amongst recipients.

Finally, despite the known medical model for return of predictive genetic information and the consideration given to the management of secondary findings, direct-to-consumer genetic testing companies exist and are currently unregulated, operating without the involvement of clinicians or trained genetic specialists interfacing with the patients. Thus in view of the lack of regulation with commercialization of the sequencing of the genome, the public needs further education regarding genetics, the pros and cons of genetic testing and clinical guidance regarding the interpretation of results.

In the succeeding years, genome wide sequencing will generate new information. In addition to novel hereditary cancer genes, it will provide clues to genetic modifiers and epigenetic changes that influence disease status, progression, prognosis and response to therapy.

Bibliography

1. LAUREN P. The two histological main types of gastric carcinoma: Diffuse and so-called intestinal-type carcinoma. an attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 1965;64:31-49.

2. Crew KD, Neugut AI: Epidemiology of gastric cancer. World J Gastroenterol 2006;12:354-362.

3. Correa P. Human gastric carcinogenesis: A multistep and multifactorial process--first american cancer society award lecture on cancer epidemiology and prevention. Cancer Res 1992;52:6735-6740.

4. Suriano G, Yew S, Ferreira P, et al: Characterization of a recurrent germ line mutation of the E-cadherin gene: Implications for genetic testing and clinical management. Clin Cancer Res 2005;11:5401-5409.

5. Huntsman DG, Carneiro F, Lewis FR, et al: Early gastric cancer in young, asymptomatic carriers of germ-line E-cadherin mutations. N Engl J Med 2001;344:1904-1909.

6. Lewis FR, Mellinger JD, Hayashi A, et al: Prophylactic total gastrectomy for familial gastric cancer. Surgery 2001;130:612-7; discussion 617-9.

7. Chun YS, Lindor NM, Smyrk TC, et al: Germline E-cadherin gene mutations: Is prophylactic total gastrectomy indicated? Cancer 2001;92:181-187.

8. Carneiro F, Huntsman DG, Smyrk TC, et al: Model of the early development of diffuse gastric cancer in E-cadherin mutation carriers and its implications for patient screening. J Pathol 2004;203:681-687.

9. Charlton A, Blair V, Shaw D, et al: Hereditary diffuse gastric cancer: Predominance of multiple foci of signet ring cell carcinoma in distal stomach and transitional zone. Gut 2004;53:814-820.

10. Shaw D, Blair V, Framp A, et al: Chromoendoscopic surveillance in hereditary diffuse gastric cancer: An alternative to prophylactic gastrectomy? Gut 2005;54:461-468.

11. Gaya DR, Stuart RC, McKee RF, et al: E-cadherin mutation-associated diffuse gastric adenocarcinoma: Penetrance and non-penetrance. Eur J Gastroenterol Hepatol

2005;17:1425-1428.

12. Blair V, Martin I, Shaw D, et al: Hereditary diffuse gastric cancer: Diagnosis and management. Clin Gastroenterol Hepatol 2006;4:262-275.

13. Francis WP, Rodrigues DM, Perez NE, et al: Prophylactic laparoscopic-assisted total gastrectomy for hereditary diffuse gastric cancer. JSLS 2007;11:142-147.

14. Norton JA, Ham CM, Van Dam J, et al: CDH1 truncating mutations in the E-cadherin gene: An indication for total gastrectomy to treat hereditary diffuse gastric cancer. Ann Surg 2007;245:873-879.

15. Barber ME, Save V, Carneiro F, et al: Histopathological and molecular analysis of gastrectomy specimens from hereditary diffuse gastric cancer patients has implications for endoscopic surveillance of individuals at risk. J Pathol 2008;216:286-294.

16. Rogers WM, Dobo E, Norton JA, et al: Risk-reducing total gastrectomy for germline mutations in E-cadherin (CDH1): Pathologic findings with clinical implications. Am J Surg Pathol 2008;32:799-809.

17. Hebbard PC, Macmillan A, Huntsman D, et al: Prophylactic total gastrectomy (PTG) for hereditary diffuse gastric cancer (HDGC): The newfoundland experience with 23 patients. Ann Surg Oncol 2009;16:1890-1895.

18. Oliveira C, Moreira H, Seruca R, et al: Role of pathology in the identification of hereditary diffuse gastric cancer: Report of a portuguese family. Virchows Arch 2005;446:181-184.

19. Humar B, Fukuzawa R, Blair V, et al: Destabilized adhesion in the gastric proliferative zone and c-src kinase activation mark the development of early diffuse gastric cancer. Cancer Res 2007;67:2480-2489.

20. Goggins M, Schutte M, Lu J, et al: Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas. Cancer Res 1996;56:5360-5364.

21. Guilford P, Hopkins J, Harraway J, et al: E-cadherin germline mutations in familial gastric cancer. Nature 1998;392:402-405.

22. Parkin DM, Bray F, Ferlay J, et al: Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.

23. Suerbaum S, Michetti P: Helicobacter pylori infection. N Engl J Med 2002;347:1175-1186.

24. Uemura N, Okamoto S, Yamamoto S, et al: Helicobacter pylori infection and the development of gastric cancer. N Engl J Med 2001;345:784-789.

25. Helicobacter and Cancer Collaborative Group. Gastric cancer and helicobacter pylori: A combined analysis of 12 case control studies nested within prospective cohorts. Gut 2001;49:347-353.

26. Ferreira AC, Isomoto H, Moriyama M, et al: Helicobacter and gastric malignancies. Helicobacter 2008;13 Suppl 1:28-34.

27. Stein M, Rappuoli R, Covacci A: Tyrosine phosphorylation of the helicobacter pylori CagA antigen after cag-driven host cell translocation. Proc Natl Acad Sci U S A 2000;97:1263-1268.

28. Ohnishi N, Yuasa H, Tanaka S, et al: Transgenic expression of helicobacter pylori CagA induces gastrointestinal and hematopoietic neoplasms in mouse. Proc Natl Acad Sci U S A 2008;105:1003-1008.

29. El-Omar EM, Carrington M, Chow WH, et al: Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature 2000;404:398-402.

30. El-Omar EM, Rabkin CS, Gammon MD, et al: Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. Gastroenterology 2003;124:1193-1201.

31. Canedo P, Corso G, Pereira F, et al: The interferon gamma receptor 1 (IFNGR1) -56C/T gene polymorphism is associated with increased risk of early gastric carcinoma. Gut 2008;57:1504-1508.

32. Forman D, Burley VJ: Gastric cancer: Global pattern of the disease and an overview of environmental risk factors. Best Pract Res Clin Gastroenterol 2006;20:633-649.

33. Liu C, Russell RM: Nutrition and gastric cancer risk: An update. Nutr Rev 2008;66:237-249.

34. McMichael AJ, McCall MG, Hartshorne JM, et al: Patterns of gastro-intestinal cancer in european migrants to australia: The role of dietary change. Int J Cancer 1980;25:431-437.

35. Henson DE, Dittus C, Younes M, et al: Differential trends in the intestinal and diffuse types of gastric carcinoma in the united states, 1973-2000: Increase in the signet ring cell type.

Arch Pathol Lab Med 2004;128:765-770.

36. Roosendaal R, Kuipers EJ, Buitenwerf J, et al: Helicobacter pylori and the birth cohort effect: Evidence of a continuous decrease of infection rates in childhood. Am J Gastroenterol 1997;92:1480-1482.

37. Borch K, Jonsson B, Tarpila E, et al: Changing pattern of histological type, location, stage and outcome of surgical treatment of gastric carcinoma. Br J Surg 2000;87:618-626.

38. Kamangar F, Dawsey SM, Blaser MJ, et al: Opposing risks of gastric cardia and noncardia gastric adenocarcinomas associated with helicobacter pylori seropositivity. J Natl Cancer Inst 2006;98:1445-1452.

39. Tatemichi M, Sasazuki S, Inoue M, et al: Different etiological role of helicobacter pylori (hp) infection in carcinogenesis between differentiated and undifferentiated gastric cancers: A nested case-control study using IgG titer against hp surface antigen. Acta Oncol 2008;47:360-365.

40. Perri F, Cotugno R, Piepoli A, et al: Aberrant DNA methylation in non-neoplastic gastric mucosa of H. pylori infected patients and effect of eradication. Am J Gastroenterol 2007;102:1361-1371.

41. Kaise M, Yamasaki T, Yonezawa J, et al: CpG island hypermethylation of tumor-suppressor genes in H. pylori-infected non-neoplastic gastric mucosa is linked with gastric cancer risk. Helicobacter 2008;13:35-41.

42. Brenner H, Arndt V, Sturmer T, et al: Individual and joint contribution of family history and helicobacter pylori infection to the risk of gastric carcinoma. Cancer 2000;88:274-279.

43. Lee KJ, Inoue M, Otani T, et al: Gastric cancer screening and subsequent risk of gastric cancer: A large-scale population-based cohort study, with a 13-year follow-up in japan. Int J Cancer 2006;118:2315-2321.

44. Sugano K. Gastric cancer: Pathogenesis, screening, and treatment. Gastrointest Endosc Clin N Am 2008;18:513-522.

45. Tsukamoto Y, Uchida T, Karnan S, et al: Genome-wide analysis of DNA copy number alterations and gene expression in gastric cancer. J Pathol 2008;216:471-482.

46. Wang LD, Qin YR, Fan ZM, et al: Comparative genomic hybridization: Comparison between esophageal squamous cell carcinoma and gastric cardia adenocarcinoma from a high-incidence area for both cancers in henan, northern china. Dis Esophagus 2006;19:459-

467.

47. Kang JU, Kang JJ, Kwon KC, et al: Genetic alterations in primary gastric carcinomas correlated with clinicopathological variables by array comparative genomic hybridization. J Korean Med Sci 2006;21:656-665.

48. Kurihara Y, Ghazizadeh M, Bo H, et al: Genome-wide screening of laser capture microdissected gastric signet-ring cell carcinomas. J Nippon Med Sch 2002;69:235-242.

49. Lee SH, Kim HS: Sequence analyses of aberrant FHIT transcripts in gastric cancer cell lines. Korean J Gastroenterol 2003;42:476-483.

50. Rocco A, Schandl L, Chen J, et al: Loss of FHIT protein expression correlates with disease progression and poor differentiation in gastric cancer. J Cancer Res Clin Oncol 2003;129:84-88.

51. Kim JH, Kim MA, Lee HS, et al: Comparative analysis of protein expressions in primary and metastatic gastric carcinomas. Hum Pathol 2009;40:314-322.

52. Barros-Silva JD, Leitao D, Afonso L, et al: Association of ERBB2 gene status with histopathological parameters and disease-specific survival in gastric carcinoma patients. Br J Cancer 2009;100:487-493.

53. Myllykangas S, Junnila S, Kokkola A, et al: Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes. Int J Cancer 2008;123:817-825.

54. Petitjean A, Mathe E, Kato S, et al: Impact of mutant p53 functional properties on *TP53* mutation patterns and tumor phenotype: Lessons from recent developments in the IARC *TP53* database. Hum Mutat 2007;28:622-629.

55. Petitjean A, Achatz MI, Borresen-Dale AL, et al: *TP53* mutations in human cancers: Functional selection and impact on cancer prognosis and outcomes. Oncogene 2007;26:2157-2165.

56. Migliavacca M, Ottini L, Bazan V, et al: *TP53* in gastric cancer: Mutations in the l3 loop and LSH motif DNA-binding domains of *TP53* predict poor outcome. J Cell Physiol 2004;200:476-485.

57. Imyanitov EN. Gene polymorphisms, apoptotic capacity and cancer risk. Hum Genet 2009;3;239-46.

58. Kim JG, Sohn SK, Chae YS, et al: *TP53* codon 72 polymorphism associated with prognosis in patients with advanced gastric cancer treated with paclitaxel and cisplatin. Cancer Chemother Pharmacol 2008;2;355-60.

59. Aparicio SR, Lumsden CE: Light- and electron-microscope studies on the granular cell myoblastoma of the tongue. J Pathol 1969;97:339-355.

60. Falchetti M, Saieva C, Lupi R, et al: Gastric cancer with high-level microsatellite instability: Target gene mutations, clinicopathologic features, and long-term survival. Hum Pathol 2008;39:925-932.

61. Leung SY, Yuen ST, Chung LP, et al: hMLH1 promoter methylation and lack of hMLH1 expression in sporadic gastric carcinomas with high-frequency microsatellite instability. Cancer Res 1999;59:159-164.

62. Wu MS, Lee CW, Shun CT, et al: Distinct clinicopathologic and genetic profiles in sporadic gastric cancer with different mutator phenotypes. Genes Chromosomes Cancer 2000;27:403-411.

63. Lynch HT, Grady W, Suriano G, et al: Gastric cancer: New genetic developments. J Surg Oncol 2005;90:114-33; discussion 133.

64. Mayer B, Johnson JP, Leitl F, et al: E-cadherin expression in primary and metastatic gastric cancer: Down-regulation correlates with cellular dedifferentiation and glandular disintegration. Cancer Res 1993;53:1690-1695.

65. Gamallo C, Palacios J, Suarez A, et al: Correlation of E-cadherin expression with differentiation grade and histological type in breast carcinoma. Am J Pathol 1993;142:987-993.

66. Kadowaki T, Shiozaki H, Inoue M, et al: E-cadherin and alpha-catenin expression in human esophageal cancer. Cancer Res 1994;54:291-296.

67. Winter JM, Ting AH, Vilardell F, et al: Absence of E-cadherin expression distinguishes noncohesive from cohesive pancreatic cancer. Clin Cancer Res 2008;14:412-418.

68. Berx G, Becker KF, Hofler H, et al: Mutations of the human E-cadherin (CDH1) gene. Hum Mutat 1998;12:226-237.

69. Keller R. Shaping the vertebrate body plan by polarized embryonic cell movements. Science 2002;298:1950-1954.

70. Geisbrecht ER, Montell DJ: Myosin VI is required for E-cadherin-mediated border cell migration. Nat Cell Biol 2002;4:616-620.

71. Larue L, Ohsugi M, Hirchenhain J, et al: E-cadherin null mutant embryos fail to form a trophectoderm epithelium. Proc Natl Acad Sci U S A 1994;91:8263-8267.

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

73. Shore EM, Nelson WJ: Biosynthesis of the cell adhesion molecule uvomorulin (E-cadherin) in madin-darby canine kidney epithelial cells. J Biol Chem 1991;266:19672-19680.

74. Caldas C, Carneiro F, Lynch HT, et al: Familial gastric cancer: Overview and guidelines for management. J Med Genet 1999;36:873-880.

75. Boller K, Vestweber D, Kemler R: Cell-adhesion molecule uvomorulin is localized in the intermediate junctions of adult intestinal epithelial cells. J Cell Biol 1985;100:327-332.

76. Thoreson MA, Anastasiadis PZ, Daniel JM, et al: Selective uncoupling of p120(ctn) from E-cadherin disrupts strong adhesion. J Cell Biol 2000;148:189-202.

77. Xiao K, Oas RG, Chiasson CM, et al: Role of p120-catenin in cadherin trafficking. Biochim Biophys Acta 2007;1773:8-16.

78. Van Aelst L, Symons M: Role of rho family GTPases in epithelial morphogenesis. Genes Dev 2002;16:1032-1054.

79. Fujita Y, Hogan C, Braga VM: Regulation of cell-cell adhesion by Rap1. Methods Enzymol 2006;407:359-372.

80. Shen Y, Hirsch DS, Sasiela CA, et al: Cdc42 regulates E-cadherin ubiquitination and degradation through an epidermal growth factor receptor to src-mediated pathway. J Biol Chem 2008;283:5127-5137.

81. Stemmler MP, Hecht A, Kemler R: E-cadherin intron 2 contains cis-regulatory elements essential for gene expression. Development 2005;132:965-976.

82. Gloushankova NA. Changes in regulation of cell-cell adhesion during tumor transformation. Biochemistry (Mosc) 2008;73:742-750.

83. Sorkin BC, Jones FS, Cunningham BA, et al: Identification of the promoter and a transcriptional enhancer of the gene encoding L-CAM, a calcium-dependent cell adhesion molecule. Proc Natl Acad Sci U S A 1993;90:11356-11360.

84. Li LC, Chui RM, Sasaki M, et al: A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. Cancer Res 2000;60:873-876.

85. Jenab M, McKay JD, Ferrari P, et al: CDH1 gene polymorphisms, smoking, helicobacter pylori infection and the risk of gastric cancer in the european prospective investigation into cancer and nutrition (EPIC-EURGAST). Eur J Cancer 2008;44:774-780.

86. Corso G, Berardi A, Marrelli D, et al: CDH1 C-160A promoter polymorphism and gastric cancer risk. Eur J Cancer Prev 2009;18:46-49.

87. Wang GY, Lu CQ, Zhang RM, et al: The E-cadherin gene polymorphism 160C->A and cancer risk: A HuGE review and meta-analysis of 26 case-control studies. Am J Epidemiol 2008;167:7-14.

88. Humar B, Graziano F, Cascinu S, et al: Association of CDH1 haplotypes with susceptibility to sporadic diffuse gastric cancer. Oncogene 2002;21:8192-8195.

89. Nasri S, More H, Graziano F, et al: A novel diffuse gastric cancer susceptibility variant in E-cadherin (CDH1) intron 2: A case control study in an italian population. BMC Cancer 2008;8:138.

90. Oliveira C, Bordin MC, Grehan N, et al: Screening E-cadherin in gastric cancer families reveals germline mutations only in hereditary diffuse gastric cancer kindred. Hum Mutat 2002;19:510-517.

91. Oliveira C, Senz J, Kaurah P, et al: Germline CDH1 deletions in hereditary diffuse gastric cancer families. Hum Mol Genet 2009;9;1545-55.

92. Frebourg T, Oliveira C, Hochain P, et al: Cleft lip/palate and CDH1/E-cadherin mutations in families with hereditary diffuse gastric cancer. J Med Genet 2006;43:138-142.

93. Inoue K, Osaka H, Thurston VC, et al: Genomic rearrangements resulting in PLP1 deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. Am J Hum Genet 2002;71:838-853.

94. Kaurah P, MacMillan A, Boyd N, et al: Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. JAMA 2007;297:2360-2372.

95. Vleminckx K, Vakaet L,Jr, Mareel M, et al: Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 1991;66:107-119.

96. Perl AK, Wilgenbus P, Dahl U, et al: A causal role for E-cadherin in the transition from adenoma to carcinoma. Nature 1998;392:190-193.

97. Bukholm IK, Nesland JM, Borresen-Dale AL: Re-expression of E-cadherin, alpha-catenin and beta-catenin, but not of gamma-catenin, in metastatic tissue from breast cancer patients [seecomments. J Pathol 2000;190:15-19.

98. Handschuh G, Candidus S, Luber B, et al: Tumour-associated E-cadherin mutations alter cellular morphology, decrease cellular adhesion and increase cellular motility. Oncogene 1999;18:4301-4312.

99. Becker KF, Atkinson MJ, Reich U, et al: E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer Res 1994;54:3845-3852.

100. Machado JC, Oliveira C, Carvalho R, et al: E-cadherin gene (CDH1) promoter methylation as the second hit in sporadic diffuse gastric carcinoma. Oncogene 2001;20:1525-1528.

101. Moll R, Mitze M, Frixen UH, et al: Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. Am J Pathol 1993;143:1731-1742.

102. Vos CB, Cleton-Jansen AM, Berx G, et al: E-cadherin inactivation in lobular carcinoma in situ of the breast: An early event in tumorigenesis. Br J Cancer 1997;76:1131-1133.

103. Oliveira C, Sousa S, Pinheiro H, et al: Quantification of epigenetic and genetic second hits in CDH1 during hereditary diffuse gastric cancer syndrome progression. Gastroenterology 2009;7;2137-48.

104. Yang J, Weinberg RA: Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. Dev Cell 2008;14:818-829.

105. Berx G, Cleton-Jansen AM, Nollet F, et al: E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO J 1995;14:6107-6115.

106. Christofori G, Semb H: The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. Trends Biochem Sci 1999;24:73-76.

107. Oda T, Kanai Y, Oyama T, et al: E-cadherin gene mutations in human gastric carcinoma cell lines. Proc Natl Acad Sci U S A 1994;91:1858-1862.

108. Grady WM, Willis J, Guilford PJ, et al: Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. Nat Genet 2000;26:16-17.

109. Barber M, Murrell A, Ito Y, et al: Mechanisms and sequelae of E-cadherin silencing in hereditary diffuse gastric cancer. J Pathol 2008;216:295-306.

110. Knudson AG,Jr. Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971;68:820-823.

111. Orsulic S, Huber O, Aberle H, et al: E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. J Cell Sci 1999;112 (Pt 8):1237-1245.

112. Aoki K, Taketo MM: Adenomatous polyposis coli (*APC*): A multi-functional tumor suppressor gene. J Cell Sci 2007;120:3327-3335.

113. Novak A, Dedhar S: Signaling through beta-catenin and Lef/Tcf. Cell Mol Life Sci 1999;56:523-537.

114. He TC, Sparks AB, Rago C, et al: Identification of c-MYC as a target of the *APC* pathway. Science 1998;281:1509-1512.

115. Gottardi CJ, Gumbiner BM: Distinct molecular forms of beta-catenin are targeted to adhesive or transcriptional complexes. J Cell Biol 2004;167:339-349.

116. Barber M, Fitzgerald RC, Caldas C: Familial gastric cancer - aetiology and pathogenesis. Best Pract Res Clin Gastroenterol 2006;20:721-734.

117. Geary J, Sasieni P, Houlston R, et al: Gene-related cancer spectrum in families with hereditary non-polyposis colorectal cancer (HNPCC). Fam Cancer 2007;2;163-72.

118. Douglas JA, Gruber SB, Meister KA, et al: History and molecular genetics of lynch syndrome in family G: A century later. JAMA 2005;294:2195-2202.

119. Cai SJ, Xu Y, Cai GX, et al: Clinical characteristics and diagnosis of patients with hereditary nonpolyposis colorectal cancer. World J Gastroenterol 2003;9:284-287.

120. Park YJ, Shin KH, Park JG: Risk of gastric cancer in hereditary nonpolyposis colorectal cancer in korea. Clin Cancer Res 2000;6:2994-2998.

121. Giardiello FM, Brensinger JD, Tersmette AC, et al: Very high risk of cancer in familial peutz-jeghers syndrome. Gastroenterology 2000;119:1447-1453.

122. Oliveira C, Ferreira P, Nabais S, et al: E-cadherin (CDH1) and p53 rather than SMAD4 and caspase-10 germline mutations contribute to genetic predisposition in portuguese gastric cancer patients. Eur J Cancer 2004;40:1897-1903.

123. Keller G, Vogelsang H, Becker I, et al: Germline mutations of the E-cadherin(CDH1) and *TP53* genes, rather than of RUNX3 and HPP1, contribute to genetic predisposition in german gastric cancer patients. J Med Genet 2004;41:e89.

124. Jagelman DG, DeCosse JJ, Bussey HJ: Upper gastrointestinal cancer in familial adenomatous polyposis. Lancet 1988;1:1149-1151.

125. Shimoyama S, Aoki F, Kawahara M, et al: Early gastric cancer development in a familial adenomatous polyposis patient. Dig Dis Sci 2004;49:260-265.

126. Brose MS, Rebbeck TR, Calzone KA, et al: Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. J Natl Cancer Inst 2002;94:1365-1372.

127. [Anonymous]. Cancer risks in BRCA2 mutation carriers.the breast cancer linkage consortium. J Natl Cancer Inst 1999;91:1310-1316.

128. Risch HA, McLaughlin JR, Cole DE, et al: Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. Am J Hum Genet 2001;68:700-710.

129. Study Group of Millennium Genome Project for Cancer, Sakamoto H, Yoshimura K, et al: Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. Nat Genet 2008;40:730-740.

130. JONES EG. Familial gastric cancer. N Z Med J 1964;63:287-296.

131. Pharoah PD, Guilford P, Caldas C, et al: Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology 2001;121:1348-1353.

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

133. Guilford PJ, Hopkins JB, Grady WM, et al: E-cadherin germline mutations define an inherited cancer syndrome dominated by diffuse gastric cancer. Hum Mutat 1999;14:249-255.

134. Gayther SA, Gorringe KL, Ramus SJ, et al: Identification of germ-line E-cadherin mutations in gastric cancer families of european origin. Cancer Res 1998;58:4086-4089.

135. Richards FM, McKee SA, Rajpar MH, et al: Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. Hum Mol Genet 1999;8:607-610.

136. Shinmura K, Kohno T, Takahashi M, et al: Familial gastric cancer: Clinicopathological characteristics, RER phenotype and germline p53 and E-cadherin mutations. Carcinogenesis 1999;20:1127-1131.

137. Yoon KA, Ku JL, Yang HK, et al: Germline mutations of E-cadherin gene in korean familial gastric cancer patients. J Hum Genet 1999;44:177-180.

138. Ascano JJ, Frierson H,Jr, Moskaluk CA, et al: Inactivation of the E-cadherin gene in sporadic diffuse-type gastric cancer. Mod Pathol 2001;14:942-949.

139. Dussaulx-Garin L, Blayau M, Pagenault M, et al: A new mutation of E-cadherin gene in familial gastric linitis plastica cancer with extra-digestive dissemination. Eur J Gastroenterol Hepatol 2001;13:711-715.

140. Yabuta T, Shinmura K, Tani M, et al: E-cadherin gene variants in gastric cancer families whose probands are diagnosed with diffuse gastric cancer. Int J Cancer 2002;101:434-441.

141. Humar B, Toro T, Graziano F, et al: Novel germline CDH1 mutations in hereditary diffuse gastric cancer families. Hum Mutat 2002;19:518-525.

142. Jonsson BA, Bergh A, Stattin P, et al: Germline mutations in E-cadherin do not explain association of hereditary prostate cancer, gastric cancer and breast cancer. Int J Cancer 2002;98:838-843.

143. Sarrio D, Moreno-Bueno G, Hardisson D, et al: Epigenetic and genetic alterations of *APC* and CDH1 genes in lobular breast cancer: Relationships with abnormal E-cadherin and catenin expression and microsatellite instability. Int J Cancer 2003;106:208-215.

144. Wang Y, Song JP, Ikeda M, et al: Ile-leu substitution (I415L) in germline E-cadherin gene (CDH1) in japanese familial gastric cancer. Jpn J Clin Oncol 2003;33:17-20.

145. Jiang Y, Wan YL, Wang ZJ, et al: Germline E-cadherin gene mutation screening in familial gastric cancer kindreds. Zhonghua Wai Ke Za Zhi 2004;42:914-917.

146. Karam R, Carvalho J, Bruno I, et al: The NMD mRNA surveillance pathway downregulates aberrant E-cadherin transcripts in gastric cancer cells and in CDH1 mutation carriers. Oncogene 2008;27:4255-4260.

147. Lee JH, Han SU, Cho H, et al: A novel germ line juxtamembrane met mutation in human gastric cancer. Oncogene 2000;19:4947-4953.

148. More H, Humar B, Weber W, et al: Identification of seven novel germline mutations in the human E-cadherin (CDH1) gene. Hum Mutat 2007;28:203.

149. Kim IJ, Park JH, Kang HC, et al: A novel germline mutation in the MET extracellular domain in a korean patient with the diffuse type of familial gastric cancer. J Med Genet 2003;40:e97.

150. Pinheiro H, Bordeira-Carrico R, Seixas S, et al: Allele-specific CDH1 downregulation and hereditary diffuse gastric cancer. Hum Mol Genet 2010;19:943-952.

151. Oliveira C, de Bruin J, Nabais S, et al: Intragenic deletion of CDH1 as the inactivating mechanism of the wild-type allele in an HDGC tumour. Oncogene 2004;23:2236-2240.

152. Fitzgerald RC, Hardwick R, Huntsman D, et al: Hereditary diffuse gastric cancer: Updated consensus guidelines for clinical management and directions for future research. J Med Genet 2010;47:436-444.

153. Vidal-Puig A, Moller DE: Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. BioTechniques 1994;17:490-2, 494, 496.

154. Mullins FM, Dietz L, Lay M, et al: Identification of an intronic single nucleotide polymorphism leading to allele dropout during validation of a CDH1 sequencing assay: Implications for designing polymerase chain reaction-based assays. Genet Med 2007;9:752-760.

155. van Kouwen MC, Drenth JP, Oyen WJ, et al: 18F]fluoro-2-deoxy-D-glucose positron emission tomography detects gastric carcinoma in an early stage in an asymptomatic E-cadherin mutation carrier. Clin Cancer Res 2004;10:6456-6459.

156. Pacelli F, Papa V, Rosa F, et al: Four hundred consecutive total gastrectomies for gastric cancer: A single-institution experience. Arch Surg 2008;143:769-75; discussion 775.

157. Lynch HT, Kaurah P, Wirtzfeld D, et al: Hereditary diffuse gastric cancer: Diagnosis, genetic counseling, and prophylactic total gastrectomy. Cancer 2008;112:2655-2663.

158. Kaurah P, Fitzgerald R, Dwerryhouse S, et al: Pregnancy after prophylactic total gastrectomy. Fam Cancer 2010;3;331-4.

159. Masciari S, Larsson N, Senz J, et al: Germline E-cadherin mutations in familial lobular breast cancer. J Med Genet 2007;11;726-31.

160. Schrader KA, Masciari S, Boyd N, et al: Hereditary diffuse gastric cancer: Association with lobular breast cancer. Fam Cancer 2008;7:73-82.

161. Schrader KA, Masciari S, Boyd N, et al: Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers. J Med Genet 2011;48:64-68.

162. Arpino G, Bardou VJ, Clark GM, et al: Infiltrating lobular carcinoma of the breast: Tumor characteristics and clinical outcome. Breast Cancer Res 2004;6:R149-56.

163. Berx G, Van Roy F: The E-cadherin/catenin complex: An important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 2001;3:289-293.

164. Berx G, Cleton-Jansen AM, Strumane K, et al: E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13:1919-1925.

165. Droufakou S, Deshmane V, Roylance R, et al: Multiple ways of silencing E-cadherin gene expression in lobular carcinoma of the breast. Int J Cancer 2001;92:404-408.

166. Thull DL, Vogel VG: Recognition and management of hereditary breast cancer syndromes. Oncologist 2004;9:13-24.

167. Rosman DS, Kaklamani V, Pasche B: New insights into breast cancer genetics and impact on patient management. Curr Treat Options Oncol 2007;8:61-73.

168. Byrnes GB, Southey MC, Hopper JL: Are the so-called low penetrance breast cancer genes, ATM, BRIP1, PALB2 and CHEK2, high risk for women with strong family histories? Breast Cancer Res 2008;10:208.

169. Keller G, Vogelsang H, Becker I, et al: Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. Am J Pathol 1999;155:337-342.

170. Smith RA, Saslow D, Sawyer KA, et al: American cancer society guidelines for breast cancer screening: Update 2003. CA Cancer J Clin 2003;53:141-169.

171. Saslow D, Boetes C, Burke W, et al: American cancer society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 2007;57:75-89.

172. Port ER, Park A, Borgen PI, et al: Results of MRI screening for breast cancer in high-risk patients with LCIS and atypical hyperplasia. Ann Surg Oncol 2007;14:1051-1057.

173. Fisher B, Costantino J, Redmond C, et al: A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor-positive tumors. N Engl J Med 1989;320:479-484.

174. Land SR, Wickerham DL, Costantino JP, et al: Patient-reported symptoms and quality of life during treatment with tamoxifen or raloxifene for breast cancer prevention: The NSABP study of tamoxifen and raloxifene (STAR) P-2 trial. JAMA 2006;295:2742-2751.

175. Wolmark N, Dunn BK: The role of tamoxifen in breast cancer prevention: Issues sparked by the NSABP breast cancer prevention trial (P-1). Ann N Y Acad Sci 2001;949:99-108.

176. Rebbeck TR, Friebel T, Lynch HT, et al: Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: The PROSE study group. J Clin Oncol 2004;22:1055-1062.

177. Barton MB, West CN, Liu IL, et al: Complications following bilateral prophylactic mastectomy. J Natl Cancer Inst Monogr 2005;(35):61-66.

178. Lodder LN, Frets PG, Trijsburg RW, et al: One year follow-up of women opting for presymptomatic testing for BRCA1 and BRCA2: Emotional impact of the test outcome and decisions on risk management (surveillance or prophylactic surgery). Breast Cancer Res Treat 2002;73:97-112.

179. Qiu LX, Li RT, Zhang JB, et al: The E-cadherin (CDH1)--160 C/A polymorphism and prostate cancer risk: A meta-analysis. Eur J Hum Genet 2009;17:244-249.

180. Tomlinson IP, Webb E, Carvajal-Carmona L, et al: A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet

2008;40:623-630.

181. Tenesa A, Farrington SM, Prendergast JG, et al: Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 2008;40:631-637.

182. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet 2010;11:31-46.

183. Ng SB, Turner EH, Robertson PD, et al: Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009;461:272-276.

184. Ng SB, Buckingham KJ, Lee C, et al: Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010;42:30-35.

185. Roach JC, Glusman G, Smit AF, et al: Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science 2010;328:636-639.

186. Choi M, Scholl UI, Ji W, et al: Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009;106:19096-19101.

187. Hoischen A, van Bon BW, Gilissen C, et al: De novo mutations of SETBP1 cause schinzel-giedion syndrome. Nat Genet 2010;42:483-485.

188. Bertolotto C, Lesueur F, Giuliano S, et al: A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma. Nature 2011;480:94-98.

189. Comino-Mendez I, Gracia-Aznarez FJ, Schiavi F, et al: Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma. Nat Genet 2011;43:663-667.

190. He ML, Chen Y, Chen Q, et al: Multiple gene dysfunctions lead to high cancer-susceptibility: Evidences from a whole-exome sequencing study. Am J Cancer Res 2011;1:562-573.

191. Lindhurst MJ, Sapp JC, Teer JK, et al: A mosaic activating mutation in AKT1 associated with the proteus syndrome. N Engl J Med 2011;365:611-619.

192. Loveday C, Turnbull C, Ramsay E, et al: Germline mutations in RAD51D confer susceptibility to ovarian cancer. Nat Genet 2011;43:879-882.

193. Ostergaard P, Simpson MA, Connell FC, et al: Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (emberger syndrome). Nat Genet 2011;43:929-931.

194. Park DJ, Odefrey FA, Hammet F, et al: FAN1 variants identified in multiple-case early-onset breast cancer families via exome sequencing: No evidence for association with risk for breast cancer. Breast Cancer Res Treat 2011;130:1043-1049.

195. Rafnar T, Gudbjartsson DF, Sulem P, et al: Mutations in BRIP1 confer high risk of ovarian cancer. Nat Genet 2011;43:1104-1107.

196. Saarinen S, Aavikko M, Aittomaki K, et al: Exome sequencing reveals germline NPAT mutation as a candidate risk factor for hodgkin lymphoma. Blood 2011;118:493-498.

197. Snape K, Hanks S, Ruark E, et al: Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. Nat Genet 2011;43:527-529.

198. Testa JR, Cheung M, Pei J, et al: Germline BAP1 mutations predispose to malignant mesothelioma. Nat Genet 2011;43:1022-1025.

199. Yokoyama S, Woods SL, Boyle GM, et al: A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. Nature 2011;480:99-103.

200. Jaeger E, Leedham S, Lewis A, et al: Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. Nat Genet 2012;44:699-703.

201. Park DJ, Lesueur F, Nguyen-Dumont T, et al: Rare mutations in XRCC2 increase the risk of breast cancer. Am J Hum Genet 2012;90:734-739.

202. Tatton-Brown K, Hanks S, Ruark E, et al: Germline mutations in the oncogene EZH2 cause weaver syndrome and increased human height. Oncotarget 2011;12;1127-33.

203. Allen-Brady K, Camp NJ, Ward JH, et al: Lobular breast cancer: Excess familiality observed in the utah population database. Int J Cancer 2005;117:655-661.

204. Lakhani SR, Gusterson BA, Jacquemier J, et al: The pathology of familial breast cancer: Histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. Clin Cancer Res 2000;6:782-789.

205. Machado JC, Oliveira C, Carvalho R, et al: E-cadherin gene (CDH1) promoter methylation as the second hit in sporadic diffuse gastric carcinoma. Oncogene 2001;20:1525-1528.

206. Sarrio D, Moreno-Bueno G, Hardisson D, et al: Epigenetic and genetic alterations of *APC* and CDH1 genes in lobular breast cancer: Relationships with abnormal E-cadherin and catenin expression and microsatellite instability. Int J Cancer 2003;106:208-215.

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

208. Rahman N, Stone JG, Coleman G, et al: Lobular carcinoma in situ of the breast is not caused by constitutional mutations in the E-cadherin gene. Br J Cancer 2000;82:568-570.

209. Salahshor S, Haixin L, Huo H, et al: Low frequency of E-cadherin alterations in familial breast cancer. Breast Cancer Res 2001;3:199-207.

210. Lei H, Sjoberg-Margolin S, Salahshor S, et al: CDH1 mutations are present in both ductal and lobular breast cancer, but promoter allelic variants show no detectable breast cancer risk. Int J Cancer 2002;98:199-204.

211. Graziano F, Humar B, Guilford P: The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: From the laboratory to clinical practice. Ann Oncol 2003;14:1705-1713.

212. Robson M, Svahn T, McCormick B, et al: Appropriateness of breast-conserving treatment of breast carcinoma in women with germline mutations in BRCA1 or BRCA2: A clinic-based series. Cancer 2005;103:44-51.

213. Sjoblom T, Jones S, Wood LD, et al: The consensus coding sequences of human breast and colorectal cancers. Science 2006;314:268-274.

214. Antoniou A, Pharoah PD, Narod S, et al: Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. Am J Hum Genet 2003;72:1117-1130.

215. Jakubowska A, Nej K, Huzarski T, et al: BRCA2 gene mutations in families with aggregations of breast and stomach cancers. Br J Cancer 2002;87:888-891.

216. Figer A, Irmin L, Geva R, et al: The rate of the 6174delT founder Jewish mutation in BRCA2 in patients with non-colonic gastrointestinal tract tumours in Israel. Br J Cancer

2001;84:478-481.

217. Risch HA, McLaughlin JR, Cole DE, et al: Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. Am J Hum Genet 2001;68:700-710.

218. Starink TM, van der Veen JP, Arwert F, et al: The cowden syndrome: A clinical and genetic study in 21 patients. Clin Genet 1986;29:222-233.

219. Brownstein MH, Wolf M, Bikowski JB: Cowden's disease: A cutaneous marker of breast cancer. Cancer 1978;41:2393-2398.

220. Hamby LS, Lee EY, Schwartz RW: Parathyroid adenoma and gastric carcinoma as manifestations of cowden's disease. Surgery 1995;118:115-117.

221. Birch JM, Alston RD, McNally RJ, et al: Relative frequency and morphology of cancers in carriers of germline *TP53* mutations. Oncogene 2001;20:4621-4628.

222. Chompret A, Abel A, Stoppa-Lyonnet D, et al: Sensitivity and predictive value of criteria for p53 germline mutation screening. J Med Genet 2001;38:43-47.

223. Kim IJ, Kang HC, Shin Y, et al: A TP53-truncating germline mutation (E287X) in a family with characteristics of both hereditary diffuse gastric cancer and li-fraumeni syndrome. J Hum Genet 2004;49:591-595.

224. CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: A collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. Am J Hum Genet 2004;74:1175-1182.

225. Kimura K, Shinmura K, Yoshimura K, et al: Absence of germline CHK2 mutations in familial gastric cancer. Jpn J Cancer Res 2000;91:875-879.

226. Dulaimi E, Hillinck J, Ibanez de Caceres I, et al: Tumor suppressor gene promoter hypermethylation in serum of breast cancer patients. Clin Cancer Res 2004;10:6189-6193.

227. Shinozaki M, Hoon DS, Giuliano AE, et al: Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis. Clin Cancer Res 2005;11:2156-2162.

228. Risinger JI, Barrett JC, Watson P, et al: Molecular genetic evidence of the occurrence of breast cancer as an integral tumor in patients with the hereditary nonpolyposis colorectal

carcinoma syndrome. Cancer 1996;77:1836-1843.

229. Scott RJ, McPhillips M, Meldrum CJ, et al: Hereditary nonpolyposis colorectal cancer in 95 families: Differences and similarities between mutation-positive and mutation-negative kindreds. Am J Hum Genet 2001;68:118-127.

230. Westenend PJ, Schutte R, Hoogmans MM, et al: Breast cancer in an MSH2 gene mutation carrier. Hum Pathol 2005;36:1322-1326.

231. Spagnoletti I, Pizzi C, Galietta A, et al: Loss of hMSH2 expression in primary breast cancer with p53 alterations. Oncol Rep 2004;11:845-851.

232. Ahmed M, Rahman N: ATM and breast cancer susceptibility. Oncogene 2006;25:5906-5911.

233. Haerer AF, Jackson JF, Evers CG: Ataxia-telangiectasia with gastric adenocarcinoma. JAMA 1969;210:1884-1887.

234. Watanabe A, Hanazono H, Sogawa H, et al: Stomach cancer of a 14-year-old boy with ataxia-telangiectasia. Tohoku J Exp Med 1977;121:127-131.

235. Bigbee WL, Langlois RG, Swift M, et al: Evidence for an elevated frequency of in vivo somatic cell mutations in ataxia telangiectasia. Am J Hum Genet 1989;44:402-408.

236. Thompson D, Duedal S, Kirner J, et al: Cancer risks and mortality in heterozygous ATM mutation carriers. J Natl Cancer Inst 2005;97:813-822.

237. Milo Y, Deutsch AA, Zahavi S, et al: Xeroderma pigmentosum with recurrent infiltrating ductal carcinoma of breast. Postgrad Med J 1994;70:240-241.

238. Puig L, Marti R, Matias-Guiu X, et al: Gastric adenocarcinoma in a patient with xeroderma pigmentosum. Br J Dermatol 1985;113:632-633.

239. Tsuchiya H, Tomita K, Ohno M, et al: Werner's syndrome combined with quintuplicate malignant tumors: A case report and review of literature data. Jpn J Clin Oncol 1991;21:135-142.

240. Wirtenberger M, Frank B, Hemminki K, et al: Interaction of werner and bloom syndrome genes with p53 in familial breast cancer. Carcinogenesis 2006;27:1655-1660.

241. Palka Bayard de Volo C, Alfonsi M, Gatta V, et al: 16q22.1 microdeletion detected by array-CGH in a family with mental retardation and lobular breast cancer. Gene 2012;498:328-331.

242. John EM, Hopper JL, Beck JC, et al: The breast cancer family registry: An infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 2004;6:R375-89.

243. Mann GJ, Thorne H, Balleine RL, et al: Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. Breast Cancer Res 2006;8:R12.

244. Mullins FM, Dietz L, Lay M, et al: Identification of an intronic single nucleotide polymorphism leading to allele dropout during validation of a CDH1 sequencing assay: Implications for designing polymerase chain reaction-based assays. Genet Med 2007;9:752-760.

245. Boggon TJ, Murray J, Chappuis-Flament S, et al: C-cadherin ectodomain structure and implications for cell adhesion mechanisms. Science 2002;296:1308-1313.

246. Suriano G, Mulholland D, de Wever O, et al: The intracellular E-cadherin germline mutation V832 M lacks the ability to mediate cell-cell adhesion and to suppress invasion. Oncogene 2003;22:5716-5719.

247. Wu C, Cipollone J, Maines-Bandiera S, et al: The morphogenic function of E-cadherin-mediated adherens junctions in epithelial ovarian carcinoma formation and progression. Differentiation 2008;76:193-205.

248. Salahshor S, Haixin L, Huo H, et al: Low frequency of E-cadherin alterations in familial breast cancer. Breast Cancer Res 2001;3:199-207.

249. Suriano G, Mulholland D, de Wever O, et al: The intracellular E-cadherin germline mutation V832 M lacks the ability to mediate cell-cell adhesion and to suppress invasion. Oncogene 2003;22:5716-5719.

250. Neuhausen SL, Ozcelik H, Southey MC, et al: BRCA1 and BRCA2 mutation carriers in the breast cancer family registry: An open resource for collaborative research. Breast Cancer Res Treat 2009;116:379-386.

251. Curtis MW, Ly QP, Wheelock MJ, et al: Evidence that the V832M E-cadherin germ-line missense mutation does not influence the affinity of alpha -catenin for the cadherin/catenin complex. Cell Commun Adhes 2007;14:45-55.

252. Ling K, Bairstow SF, Carbonara C, et al: Type igamma phosphatidylinositol phosphate kinase modulates adherens junction and E-cadherin trafficking via a direct interaction with mu 1B adaptin. J Cell Biol 2007;176:343-353.

253. Schubbert S, Shannon K, Bollag G: Hyperactive ras in developmental disorders and cancer. Nat Rev Cancer 2007;7:295-308.

254. Kontaridis MI, Swanson KD, David FS, et al: *PTPN11* (Shp2) mutations in LEOPARD syndrome have dominant negative, not activating, effects. J Biol Chem 2006;281:6785-6792.

255. Pandit B, Sarkozy A, Pennacchio LA, et al: Gain-of-function *RAF1* mutations cause noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. Nat Genet 2007;8;1007-12.

256. Bowen ME, Boyden ED, Holm IA, et al: Loss-of-function mutations in *PTPN11* cause metachondromatosis, but not ollier disease or maffucci syndrome. PLoS Genet 2011;7:e1002050.

257. Tartaglia M, Niemeyer CM, Fragale A, et al: Somatic mutations in *PTPN11* in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet 2003;34:148-150.

258. FISHER ER, WECHSLER H: Granular cell myoblastoma--a misnomer. electron microscopic and histochemical evidence concerning its schwann cell derivation and nature (granular cell schwannoma). Cancer 1962;15:936-954.

259. Khansur T, Balducci L, Tavassoli M: Granular cell tumor. clinical spectrum of the benign and malignant entity. Cancer 1987;60:220-222.

260. Gorlin RJ, Anderson RC, Blaw M: Multiple lentigenes syndrome. Am J Dis Child 1969;117:652-662.

261. Digilio MC, Sarkozy A, de Zorzi A, et al: LEOPARD syndrome: Clinical diagnosis in the first year of life. Am J Med Genet A 2006;140:740-746.

262. Digilio MC, Conti E, Sarkozy A, et al: Grouping of multiple-lentigines/LEOPARD and noonan syndromes on the *PTPN11* gene. Am J Hum Genet 2002;71:389-394.

263. Sarkozy A, Conti E, Digilio MC, et al: Clinical and molecular analysis of 30 patients with multiple lentigines LEOPARD syndrome. J Med Genet 2004;41:e68.

264. Merks JH, Caron HN, Hennekam RC: High incidence of malformation syndromes in a series of 1,073 children with cancer. Am J Med Genet A 2005;134:132-143.

265. Ucar C, Calyskan U, Martini S, et al: Acute myelomonocytic leukemia in a boy with LEOPARD syndrome (*PTPN11* gene mutation positive). J Pediatr Hematol Oncol 2006;28:123-125.

266. Heney D, Lockwood L, Allibone EB, et al: Nasopharyngeal rhabdomyosarcoma and multiple lentigines syndrome: A case report. Med Pediatr Oncol 1992;20:227-228.

267. Choi WW, Yoo JY, Park KC, et al: LEOPARD syndrome with a new association of congenital corneal tumor, choristoma. Pediatr Dermatol 2003;20:158-160.

268. Kontaridis MI, Swanson KD, David FS, et al: *PTPN11* (Shp2) mutations in LEOPARD syndrome have dominant negative, not activating, effects. J Biol Chem 2006;281:6785-6792.

269. Hanna N, Montagner A, Lee WH, et al: Reduced phosphatase activity of SHP-2 in LEOPARD syndrome: Consequences for PI3K binding on Gab1. FEBS Lett 2006;580:2477-2482.

270. Tartaglia M, Mehler EL, Goldberg R, et al: Mutations in *PTPN11*, encoding the protein tyrosine phosphatase SHP-2, cause noonan syndrome. Nat Genet 2001;29:465-468.

271. Razzaque MA, Nishizawa T, Komoike Y, et al: Germline gain-of-function mutations in *RAF1* cause noonan syndrome. Nat Genet 2007;39:1013-1017.

272. Voron DA, Hatfield HH, Kalkhoff RK: Multiple lentigines syndrome. case report and review of the literature. Am J Med 1976;60:447-456.

273. Marchese C, Montera M, Torrini M, et al: Granular cell tumor in a PHTS patient with a novel germline *PTEN* mutation. Am J Med Genet A 2003;120:286-288.

274. Selmanowitz VJ, Orentreich N: Lentiginosis profusa in daughter and mother: Multiple granular cell "myoblastomas" in the former. Arch Dermatol 1970;101:615-616.

275. Martin RW,3rd, Neldner KH, Boyd AS, et al: Multiple cutaneous granular cell tumors and neurofibromatosis in childhood. A case report and review of the literature. Arch Dermatol 1990;126:1051-1056.

276. Sahn EE, Dunlavey ES, Parsons JL: Multiple cutaneous granular cell tumors in a child with possible neurofibromatosis. J Am Acad Dermatol 1997;36:327-330.

277. Lohmann DR, Gillessen-Kaesbach G: Multiple subcutaneous granular-cell tumours in a patient with noonan syndrome. Clin Dysmorphol 2000;9:301-302.

278. Fryssira H, Leventopoulos G, Psoni S, et al: Tumor development in three patients with noonan syndrome. Eur J Pediatr 2007;9;1025-31.

279. Muscardin LM, Paradisi M, Provini A, et al: Multiple cutaneous granular cell tumors, joint hypermobility and mild facial dysmorphism in a child. Int J Dermatol 2006;45:847-850.

280. Tomson N, Abdullah A, Tan CY: Multiple granular cell tumors in a child with growth retardation. report of a case and review of the literature. Int J Dermatol 2006;45:1358-1361.

281. Bentires-Alj M, Kontaridis MI, Neel BG: Stops along the RAS pathway in human genetic disease. Nat Med 2006;12:283-285.

282. Nurse DS. Granular cell myoblastoma lentiginosis profusa. Australas J Dermatol 1974;15:33-34.

283. KESSLER M. On a case of multiple granular cell myoblastic myomas in a child. Dermatologica 1963;126:167-181.

284. Apted JH. Multiple granular-cell myoblastoma (schwannoma) in a child. Br J Dermatol 1968;80:257-260.

285. Papageorgiou S, Litt JZ, Pomeranz JR: Multiple granular cell myoblastomas in children. Arch Dermatol 1967;96:168-171.

286. Bakos L. Multiple cutaneous granular cell tumors with systemic defects: A distinct entity? Int J Dermatol 1993;32:432-435.

287. Savin JA, Wallace HJ: Granular cell myoblastoma. Proc R Soc Med 1971;64:863-864.

288. Seiter S, Ugurel S, Tilgen W, et al: Multiple granular cell tumors and growth hormone deficiency in a child. Pediatr Dermatol 1999;16:308-310.

289. Zhu Y, Ghosh P, Charnay P, et al: Neurofibromas in NF1: Schwann cell origin and role of tumor environment. Science 2002;296:920-922.

290. Koukourakis GV, Kouloulias VE, Koukourakis MJ, et al: Granulosa cell tumor of the ovary: Tumor review. Integr Cancer Ther 2008;7:204-215.

291. Shah SP, Kobel M, Senz J, et al: Mutation of *FOXL2* in granulosa-cell tumors of the ovary. N Engl J Med 2009;26:2719-29.

292. Schmidt D, Ovitt CE, Anlag K, et al: The murine winged-helix transcription factor *FOXL2* is required for granulosa cell differentiation and ovary maintenance. Development 2004;131:933-942.

293. Kalfa N, Philibert P, Patte C, et al: Juvenile granulosa-cell tumor: Clinical and molecular expression. Gynecol Obstet Fertil 2009;37:33-44.

294. Nishi Y, Yanase T, Mu Y, et al: Establishment and characterization of a steroidogenic human granulosa-like tumor cell line, KGN, that expresses functional follicle-stimulating hormone receptor. Endocrinology 2001;142:437-445.

295. Reed GH, Wittwer CT: Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clin Chem 2004;50:1748-1754.

296. Herrmann MG, Durtschi JD, Wittwer CT, et al: Expanded instrument comparison of amplicon DNA melting analysis for mutation scanning and genotyping. Clin Chem 2007;53:1544-1548.

297. Li D, Zeng W, Tao J, et al: Mutations of the transcription factor *FOXL2* gene in chinese patients with blepharophimosis-ptosis-epicanthus inversus syndrome. Genet Test Mol Biomarkers 2009;13:257-268.

298. Mayr D, Hirschmann A, Marlow S, et al: Analysis of selected oncogenes (AKT1, FOS, BCL2L2, TGFbeta) on chromosome 14 in granulosa cell tumors (GCTs): A comprehensive study on 30 GCTs combining comparative genomic hybridization (CGH) and fluorescence-in situ-hybridization (FISH). Pathol Res Pract 2008;204:823-830.

299. Crisponi L, Deiana M, Loi A, et al: The putative forkhead transcription factor *FOXL2* is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. Nat Genet 2001;27:159-166.

300. Beysen D, Moumne L, Veitia R, et al: Missense mutations in the forkhead domain of *FOXL2* lead to subcellular mislocalization, protein aggregation and impaired transactivation. Hum

Mol Genet 2008;17:2030-2038.

301. Batista F, Vaiman D, Dausset J, et al: Potential targets of *FOXL2*, a transcription factor involved in craniofacial and follicular development, identified by transcriptomics. Proc Natl Acad Sci U S A 2007;104:3330-3335.

302. Benayoun BA, Batista F, Auer J, et al: Positive and negative feedback regulates the transcription factor *FOXL2* in response to cell stress: Evidence for a regulatory imbalance induced by disease-causing mutations. Hum Mol Genet 2009;18:632-644.

303. Beysen D, De Jaegere S, Amor D, et al: Identification of 34 novel and 56 known *FOXL2* mutations in patients with blepharophimosis syndrome. Hum Mutat 2008;29:E205-19.

304. Moumne L, Dipietromaria A, Batista F, et al: Differential aggregation and functional impairment induced by polyalanine expansions in *FOXL2*, a transcription factor involved in cranio-facial and ovarian development. Hum Mol Genet 2008;17:1010-1019.

305. Kobel M, Kalloger SE, Boyd N, et al: Ovarian carcinoma subtypes are different diseases: Implications for biomarker studies. PLoS Med 2008;5:e232.

306. Gilks CB, Prat J: Ovarian carcinoma pathology and genetics: Recent advances. Hum Pathol 2009;40:1213-1223.

307. Kalloger SE, Kobel M, Leung S, et al: Calculator for ovarian carcinoma subtype prediction. Mod Pathol 2011;24:512-521.

308. Gilks CB. Molecular abnormalities in ovarian cancer subtypes other than high-grade serous carcinoma. J Oncol 2010;2010:740968.

309. Risch HA, McLaughlin JR, Cole DE, et al: Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: A kin-cohort study in ontario, canada. J Natl Cancer Inst 2006;98:1694-1706.

310. Pal T, Permuth-Wey J, Betts JA, et al: BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. Cancer 2005;104:2807-2816.

311. Pharoah PD, Easton DF, Stockton DL, et al: Survival in familial, BRCA1-associated, and BRCA2-associated epithelial ovarian cancer. united kingdom coordinating committee for cancer research (UKCCCR) familial ovarian cancer study group. Cancer Res 1999;59:868-871.

312. Werness BA, Ramus SJ, DiCioccio RA, et al: Histopathology, FIGO stage, and BRCA mutation status of ovarian cancers from the gilda radner familial ovarian cancer registry. Int J Gynecol Pathol 2004;23:29-34.

313. Lakhani SR, Reis-Filho JS, Fulford L, et al: Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer Res 2005;11:5175-5180.

314. Zhang S, Royer R, Li S, et al: Frequencies of BRCA1 and BRCA2 mutations among 1,342 unselected patients with invasive ovarian cancer. Gynecol Oncol 2011;121:353-357.

315. Alsop K, Fereday S, Meldrum C et al. (2011) Germ-line BRCA mutations in high-grade ovarian cancer: A case for routine BRCA mutation screening after a diagnosis of invasive ovarian cancer.. J Clin Oncol 29: suppl; abstr 5026.

316. Press JZ, De Luca A, Boyd N, et al: Ovarian carcinomas with genetic and epigenetic BRCA1 loss have distinct molecular abnormalities. BMC Cancer 2008;8:17.

317. McAlpine JN, Porter H, Kobel M, et al: BRCA1 and BRCA2 mutations correlate with *TP53* abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma. Mod Pathol 2012;5;740-50.

318. Hennessy BT, Timms KM, Carey MS, et al: Somatic mutations in BRCA1 and BRCA2 could expand the number of patients that benefit from poly (ADP ribose) polymerase inhibitors in ovarian cancer. J Clin Oncol 2010;28:3570-3576.

319. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474:609-615.

320. Arnold A. G., Bhatia J., Phillips M. F. et al. (2011) Feasibility and yield of offering genetic counseling to all patients with newly diagnosed high-grade epithelial ovarian cancer.. J Clin Oncol 29: suppl; abstr 1509.

321. Walsh T, Casadei S, Lee MK, et al: Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci U S A 2011;108:18032-18037.

322. Gilks CB, Ionescu DN, Kalloger SE, et al: Tumor cell type can be reproducibly diagnosed and is of independent prognostic significance in patients with maximally debulked ovarian carcinoma. Hum Pathol 2008;39:1239-1251.

323. Kobel M, Kalloger SE, Baker PM, et al: Diagnosis of ovarian carcinoma cell type is highly reproducible: A transcanadian study. Am J Surg Pathol 2010;34:984-993.

324. American College of Obstetricians and Gynecologists, ACOG Committee on Practice Bulletins--Gynecology, ACOG Committee on Genetics, et al: ACOG practice bulletin no. 103: Hereditary breast and ovarian cancer syndrome. Obstet Gynecol 2009;113:957-966.

325. Whyte MP, Gottesman GS, Eddy MC, et al: X-linked recessive spondyloepiphyseal dysplasia tarda. clinical and radiographic evolution in a 6-generation kindred and review of the literature. Medicine (Baltimore) 1999;78:9-25.

326. Bal S, Kocyigit H, Turan Y, et al: Spondyloepiphyseal dysplasia tarda: Four cases from two families. Rheumatol Int 2009;29:699-702.

327. Van Camp G, Snoeckx RL, Hilgert N, et al: A new autosomal recessive form of stickler syndrome is caused by a mutation in the COL9A1 gene. Am J Hum Genet 2006;79:449-457.

328. Farrar GJ, Kenna PF, Humphries P: On the genetics of retinitis pigmentosa and on mutation-independent approaches to therapeutic intervention. EMBO J 2002;21:857-864.

329. Boyer A, Paquet M, Lague MN, et al: Dysregulation of WNT/CTNNB1 and PI3K/AKT signaling in testicular stromal cells causes granulosa cell tumor of the testis. Carcinogenesis 2009;30:869-878.

330. Eichers ER, Green JS, Stockton DW, et al: Newfoundland rod-cone dystrophy, an early-onset retinal dystrophy, is caused by splice-junction mutations in RLBP1. Am J Hum Genet 2002;70:955-964.

331. Rahman P, Jones A, Curtis J, et al: The newfoundland population: A unique resource for genetic investigation of complex diseases. Hum Mol Genet 2003;12 Spec No 2:R167-72.

332. Green JS, Parfrey PS, Harnett JD, et al: The cardinal manifestations of bardet-biedl syndrome, a form of laurence-moon-biedl syndrome. N Engl J Med 1989;321:1002-1009.

333. Young TL, Woods MO, Parfrey PS, et al: Canadian bardet-biedl syndrome family reduces the critical region of BBS3 (3p) and presents with a variable phenotype. Am J Med Genet 1998;78:461-467.

334. Young TL, Penney L, Woods MO, et al: A fifth locus for bardet-biedl syndrome maps to chromosome 2q31. Am J Hum Genet 1999;64:900-904.

335. Young TL, Woods MO, Parfrey PS, et al: A founder effect in the newfoundland population reduces the bardet-biedl syndrome I (BBS1) interval to 1 cM. Am J Hum Genet 1999;65:1680-1687.

336. Katsanis N, Beales PL, Woods MO, et al: Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with bardet-biedl syndrome. Nat Genet 2000;26:67-70.

337. Fan Y, Esmail MA, Ansley SJ, et al: Mutations in a member of the ras superfamily of small GTP-binding proteins causes bardet-biedl syndrome. Nat Genet 2004;36:989-993.

338. Ng SB, Buckingham KJ, Lee C, et al: Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010;42:30-35.

339. Gnirke A, Melnikov A, Maguire J, et al: Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol 2009;27:182-189.

340. Li H, Ruan J, Durbin R: Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 2008;18:1851-1858.

341. Li H, Handsaker B, Wysoker A, et al: The sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-2079.

342. Goya R, Sun MG, Morin RD, et al: SNVMix: Predicting single nucleotide variants from next-generation sequencing of tumors. Bioinformatics 2010;26:730-736.

343. Raas-Rothschild A, Cormier-Daire V, Bao M, et al: Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). J Clin Invest 2000;105:673-681.

344. Varki AP, Reitman ML, Kornfeld S: Identification of a variant of mucolipidosis III (pseudo-hurler polydystrophy): A catalytically active N-acetylglucosaminylphosphotransferase that fails to phosphorylate lysosomal enzymes. Proc Natl Acad Sci U S A 1981;78:7773-7777.

345. Tiede S, Cantz M, Raas-Rothschild A, et al: A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site. Hum Mutat 2004;24:535.

346. Raas-Rothschild A, Bargal R, Goldman O, et al: Genomic organisation of the UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) and its mutations in mucolipidosis III. J Med Genet 2004;41:e52.

347. Pohl S, Tiede S, Castrichini M, et al: Compensatory expression of human N-acetylglucosaminyl-1-phosphotransferase subunits in mucolipidosis type III gamma. Biochim Biophys Acta 2009;3;221-5.

348. Encarnacao M, Lacerda L, Costa R, et al: Molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations. Clin Genet 2009;76:76-84.

349. Persichetti E, Chuzhanova NA, Dardis A, et al: Identification and molecular characterization of six novel mutations in the UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (*GNPTG*) gene in patients with mucolipidosis III gamma. Hum Mutat 2009;30:978-984.

350. Pohl S, Encarnacao M, Castrichini M, et al: Loss of N-acetylglucosamine-1-phosphotransferase gamma subunit due to intronic mutation in *GNPTG* causes mucolipidosis type III gamma: Implications for molecular and cellular diagnostics. Am J Med Genet A 2010;152A:124-132.

351. Reitman ML, Varki A, Kornfeld S: Fibroblasts from patients with I-cell disease and pseudo-hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: Glycoprotein N-acetylglucosaminylphosphotransferase activity. J Clin Invest 1981;67:1574-1579.

352. Vogel P, Payne BJ, Read R, et al: Comparative pathology of murine mucolipidosis types II and IIIC. Vet Pathol 2009;46:313-324.

353. Traboulsi EI, Maumenee IH: Ophthalmologic findings in mucolipidosis III (pseudo-hurler polydystrophy). Am J Ophthalmol 1986;102:592-597.

354. Kang C, Riazuddin S, Mundorff J, et al: Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering. N Engl J Med 2010;362:677-685.

355. Guilford P, Hopkins J, Harraway J, et al: E-cadherin germline mutations in familial gastric cancer. Nature 1998;392:402-405.

356. Gelfman CM, Vogel P, Issa TM, et al: Mice lacking alpha/beta subunits of GlcNAc-1-phosphotransferase exhibit growth retardation, retinal degeneration, and secretory cell lesions. Invest Ophthalmol Vis Sci 2007;48:5221-5228.

357. Isidor B, Le Merrer M, Ramos E, et al: Cone-rod dystrophy, growth hormone deficiency and spondyloepiphyseal dysplasia: Report of a new case without nephronophtisis. Am J Med Genet A 2009;149A:788-792.

358. Walters BA, Raff ML, Hoeve JV, et al: Spondylometaphyseal dysplasia with cone-rod dystrophy. Am J Med Genet A 2004;129A:265-276.

359. Isidor B, Baron S, Khau van Kien P, et al: Axial spondylometaphyseal dysplasia: Confirmation and further delineation of a new SMD with retinal dystrophy. Am J Med Genet A 2010;152A:1550-1554.

360. Sousa SB, Russell-Eggitt I, Hall C, et al: Further delineation of spondylometaphyseal dysplasia with cone-rod dystrophy. Am J Med Genet A 2008;146A:3186-3194.

361. Kuriyama M, Miyatake T, Owada M, et al: Neuraminidase activities in sialidosis and mucolipidosis. J Neurol Sci 1982;54:181-187.

362. Abraham SC, Nobukawa B, Giardiello FM, et al: Sporadic fundic gland polyps: Common gastric polyps arising through activating mutations in the beta-catenin gene. Am J Pathol 2001;158:1005-1010.

363. Burt RW. Gastric fundic gland polyps. Gastroenterology 2003;125:1462-1469.

364. Ally MR, Veerappan GR, Maydonovitch CL, et al: Chronic proton pump inhibitor therapy associated with increased development of fundic gland polyps. Dig Dis Sci 2009;54:2617-2622.

365. Kadmon M, Tandara A, Herfarth C: Duodenal adenomatosis in familial adenomatous polyposis coli. A review of the literature and results from the heidelberg polyposis register. Int J Colorectal Dis 2001;16:63-75.

366. Lynch HT, Snyder C, Davies JM, et al: FAP, gastric cancer, and genetic counseling featuring children and young adults: A family study and review. Fam Cancer 2010;9:581-588.

367. Bianchi LK, Burke CA, Bennett AE, et al: Fundic gland polyp dysplasia is common in familial adenomatous polyposis. Clin Gastroenterol Hepatol 2008;6:180-185.

368. Garrean S, Hering J, Saied A, et al: Gastric adenocarcinoma arising from fundic gland polyps in a patient with familial adenomatous polyposis syndrome. Am Surg 2008;74:79-83.

369. Hofgartner WT, Thorp M, Ramus MW, et al: Gastric adenocarcinoma associated with fundic gland polyps in a patient with attenuated familial adenomatous polyposis. Am J Gastroenterol 1999;94:2275-2281.

370. Abraham SC, Nobukawa B, Giardiello FM, et al: Fundic gland polyps in familial adenomatous polyposis: Neoplasms with frequent somatic adenomatous polyposis coli gene alterations. Am J Pathol 2000;157:747-754.

371. Vogt S, Jones N, Christian D, et al: Expanded extracolonic tumor spectrum in MUTYH-associated polyposis. Gastroenterology 2009;137:1976-10.

372. Amos CI, Keitheri-Cheteri MB, Sabripour M, et al: Genotype-phenotype correlations in peutz-jeghers syndrome. J Med Genet 2004;41:327-333.

373. Friedl W, Uhlhaas S, Schulmann K, et al: Juvenile polyposis: Massive gastric polyposis is more common in MADH4 mutation carriers than in BMPR1A mutation carriers. Hum Genet 2002;111:108-111.

374. Rio Frio T, Lavoie J, Hamel N, et al: Homozygous BUB1B mutation and susceptibility to gastrointestinal neoplasia. N Engl J Med 2010;363:2628-2637.

375. Ha M, Chung JW, Hahm KB, et al: A case of cowden syndrome diagnosed from multiple gastric polyposis. World J Gastroenterol 2012;18:861-864.

376. Worthley DL, Phillips KD, Wayte N, et al: Gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS): A new autosomal dominant syndrome. Gut 2012;61:774-779.

377. Jenkins BJ, Grail D, Nheu T, et al: Hyperactivation of Stat3 in gp130 mutant mice promotes gastric hyperproliferation and desensitizes TGF-beta signaling. Nat Med 2005;11:845-852.

378. Baselga J, Norton L, Albanell J, et al: Recombinant humanized anti-HER2 antibody (herceptin) enhances the antitumor activity of paclitaxel and doxorubicin against HER2/neu overexpressing human breast cancer xenografts. Cancer Res 1998;58:2825-2831.

379. Peggs K, Mackinnon S: Imatinib mesylate--the new gold standard for treatment of chronic myeloid leukemia. N Engl J Med 2003;348:1048-1050.

380. Thomas DA, Faderl S, Cortes J, et al: Treatment of philadelphia chromosome-positive acute lymphocytic leukemia with hyper-CVAD and imatinib mesylate. Blood 2004;103:4396-4407.

381. Apperley JF, Gardembas M, Melo JV, et al: Response to imatinib mesylate in patients with chronic myeloproliferative diseases with rearrangements of the platelet-derived growth factor receptor beta. N Engl J Med 2002;347:481-487.

382. Demetri GD, von Mehren M, Blanke CD, et al: Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. N Engl J Med 2002;347:472-480.

383. Lynch TJ, Bell DW, Sordella R, et al: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129-2139.

384. Kwak EL, Bang YJ, Camidge DR, et al: Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 2010;363:1693-1703.

385. Karapetis CS, Khambata-Ford S, Jonker DJ, et al: K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med 2008;359:1757-1765.

386. Flaherty KT, Puzanov I, Kim KB, et al: Inhibition of mutated, activated BRAF in metastatic melanoma. N Engl J Med 2010;363:809-819.

387. Burn J, Mathers J, Bishop DT: Genetics, inheritance and strategies for prevention in populations at high risk of colorectal cancer (CRC). Recent Results Cancer Res 2012;191:157-183.

388. Fong PC, Boss DS, Yap TA, et al: Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 2009;361:123-134.

389. Svard J, Heby-Henricson K, Persson-Lek M, et al: Genetic elimination of suppressor of fused reveals an essential repressor function in the mammalian hedgehog signaling pathway. Dev Cell 2006;10:187-197.

390. Von Hoff DD, LoRusso PM, Rudin CM, et al: Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. N Engl J Med 2009;361:1164-1172.

391. Tang JY, Mackay-Wiggan JM, Aszterbaum M, et al: Inhibiting the hedgehog pathway in patients with the basal-cell nevus syndrome. N Engl J Med 2012;366:2180-2188.

392. Bissler J.
Everolimus therapy for angiomyolipoma in patients with tuberous sclerosis complex or sporadic lymphangioleiomyomatosis: Results from EXIST-2.. J Clin Oncol 2012;30:.

393. Krueger DA, Care MM, Holland K, et al: Everolimus for subependymal giant-cell astrocytomas in tuberous sclerosis. N Engl J Med 2010;363:1801-1811.

394. Robinson BG, Paz-Ares L, Krebs A, et al: Vandetanib (100 mg) in patients with locally advanced or metastatic hereditary medullary thyroid cancer. J Clin Endocrinol Metab 2010;95:2664-2671.

395. Thornton K, Kim G, Maher VE, et al: Vandetanib for the treatment of symptomatic or progressive medullary thyroid cancer in patients with unresectable locally advanced or metastatic disease: U.S. food and drug administration drug approval summary. Clin Cancer Res 2012;18:3722-3730.

396. Collins CC, Volik SV, Lapuk AV, et al: Next generation sequencing of prostate cancer from a patient identifies a deficiency of methylthioadenosine phosphorylase, an exploitable tumor target. Mol Cancer Ther 2012;11:775-783.

397. Welch JS, Westervelt P, Ding L, et al: Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. JAMA 2011;305:1577-1584.

398. Roychowdhury S, Iyer MK, Robinson DR, et al: Personalized oncology through integrative high-throughput sequencing: A pilot study. Sci Transl Med 2011;3:111ra121.

399. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330-337.

400. Cancer Genome Atlas Research Network, Hammerman PS, Hayes DN, et al: Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519-525.

401. The Cancer Genome Atlas Network, Genome sequencing centres: Washington University in St Louis, Koboldt DC, et al: Comprehensive molecular portraits of human breast tumours. Nature 2012;7418;61-70.

402. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 2008;455:1061-1068.

403. Shah SP, Morin RD, Khattra J, et al: Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature 2009;461:809-813.

404. Ding L, Ley TJ, Larson DE, et al: Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 2012;481:506-510.

405. Turajlic S, Furney SJ, Lambros MB, et al: Whole genome sequencing of matched primary and metastatic acral melanomas. Genome Res 2012;22:196-207.

406. Navin N, Kendall J, Troge J, et al: Tumour evolution inferred by single-cell sequencing. Nature 2011;472:90-94.

407. Swanton C. Intratumor heterogeneity: Evolution through space and time. Cancer Res 2012;19;4875-82.

408. Shah SP, Roth A, Goya R, et al: The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012;486:395-399.

409. Gerlinger M, Rowan AJ, Horswell S, et al: Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883-892.

410. Anderson K, Lutz C, van Delft FW, et al: Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature 2011;469:356-361.

411. De Leeuw WJ, Berx G, Vos CB, et al: Simultaneous loss of E-cadherin and catenins in invasive lobular breast cancer and lobular carcinoma in situ. J Pathol 1997;183:404-411.

412. Ware JS, Walsh R, Cunningham F, et al: Paralogous annotation of disease-causing variants in long QT syndrome genes. Hum Mutat 2012;33:1188-1191.

413. de Visser KE, Eichten A, Coussens LM: Paradoxical roles of the immune system during cancer development. Nat Rev Cancer 2006;6:24-37.

414. Bhowmick NA, Chytil A, Plieth D, et al: TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. Science 2004;303:848-851.

415. Tanwar PS, Kaneko-Tarui T, Zhang L, et al: Stromal liver kinase B1 [STK11] signaling loss induces oviductal adenomas and endometrial cancer by activating mammalian target of rapamycin complex 1. PLoS Genet 2012;8:e1002906.

416. Straussman R, Morikawa T, Shee K, et al: Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature 2012;487:500-504.

417. Landsberg J, Kohlmeyer J, Renn M, et al: Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. Nature 2012;7420;412-6.

418. Link DC, Schuettpelz LG, Shen D, et al: Identification of a novel *TP53* cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML. JAMA 2011;305:1568-1576.

419. Zhang J, Benavente CA, McEvoy J, et al: A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature 2012;481:329-334.

420. Caulfield T, McGuire AL, Cho M, et al: Research ethics recommendations for whole-genome research: Consensus statement. PLoS Biol 2008;6:e73.

421. Bredenoord AL, Onland-Moret NC, Van Delden JJ: Feedback of individual genetic results to research participants: In favor of a qualified disclosure policy. Hum Mutat 2011;32:861-867.

422. Wolf SM, Crock BN, Van Ness B, et al: Managing incidental findings and research results in genomic research involving biobanks and archived data sets. Genet Med 2012;14:361-384.

423. Homer N, Szelinger S, Redman M, et al: Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS Genet 2008;4:e1000167.

424. Biesecker LG. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: Lessons from the ClinSeq project. Genet Med 2012;14:393-398.

425. Berg JS, Khoury MJ, Evans JP: Deploying whole genome sequencing in clinical practice and public health: Meeting the challenge one bin at a time. Genet Med 2011;13:499-504.

426. Lunshof JE, Bobe J, Aach J, et al: Personal genomes in progress: From the human genome project to the personal genome project. Dialogues Clin Neurosci 2010;12:47-60.

427. Adzhubei IA, Schmidt S, Peshkin L, et al: A method and server for predicting damaging missense mutations. Nat Methods 2010;4:248-249.

448. Tucker T, Montpetit A, Chai D,et al: Comparison of genome-wide array genomic hybridization platforms for the detection of copy number variants in idiopathic mental retardation. BMC Med Genomics. 2011; 4: 25.

449. Cleary SP, Cotterchio M, Jenkins MA, et al: Germline MutY Human Homologue Mutations and Colorectal Cancer: A Multisite Case-Control Study, Gastro 2009; 136; 1251-1260.

450. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. 1990; 20:6097-100.

451. Robinson JT, Thorvaldsdóttir H, Winckler W, et al: Integrative genomics viewer. Nat Biotechnol. 2011; 29:24-6.

452. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Altekruse SF, Kosary CL, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). *SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations)*, National

Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2009_pops09/, based on November 2011 SEER data submission, posted to the SEER web site, 2012.