MOLECULAR INVESTIGATION OF MINOR GENOMIC POPULATIONS AND BIOLOGICAL EXPOSURES IN HUMAN HEALTH

Brandon Ned Johnson

MOLECULAR INVESTIGATION OF MINOR GENOMIC POPULATIONS AND BIOLOGICAL EXPOSURES IN HUMAN HEALTH

Brandon Ned Johnson

Molecular investigation of minor genomic populations and biological exposures in human health

Brandon Ned Johnson

ISBN: 978-94-6458-878-1

Printing: Ridderprint, www.ridderprint.nl
Layout and cover design: Harma Makken, persoonlijkproefschrift.nl

VRIJE UNIVERSITEIT

MOLECULAR INVESTIGATION OF MINOR GENOMIC POPULATIONS AND BIOLOGICAL EXPOSURES IN HUMAN HEALTH

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op donderdag 16 februari 2023 om 15.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Brandon Ned Johnson

geboren te Maplewood, Minnesota, Verenigde Staten

promotoren:	prof.dr. D.I. Boomsma prof.dr. A.H.M. Willemsen
copromotor:	dr. E.A. Ehli
promotiecommissie:	prof.dr. J.C.N. de Geus prof.dr. M.J. Daly prof.dr. A.G. Uitterlinde prof.dr. P.E. Slagboom dr. J.L. Petersen dr. E.A. Sistermans

paranymphs:	Austin J. Van Asselt Jeffrey J. Beck

ACKNOWLEDGEMENTS

The body of work herein would not have been possible without the support and contributions from the extraordinary group of individuals that I have been surrounded by.

First and foremost, I want to extend my eternal gratitude to my wife, Katie. None of this would have been possible without her steadfast support, patience, and continuous encouragement throughout the entirety of my graduate studies. Additionally, my children, Aiden and Cade, have been a constant source of joy and they continue to inspire me every day.

I gratefully acknowledge my parents, Bill and Karen, my brother, Brice, and his wife Alyssa. I am sincerely appreciative of their unwavering support throughout my life and especially during my graduate studies. I am immensely grateful for their spirit and compassion which has provided the motivation to overcome countless challenges.

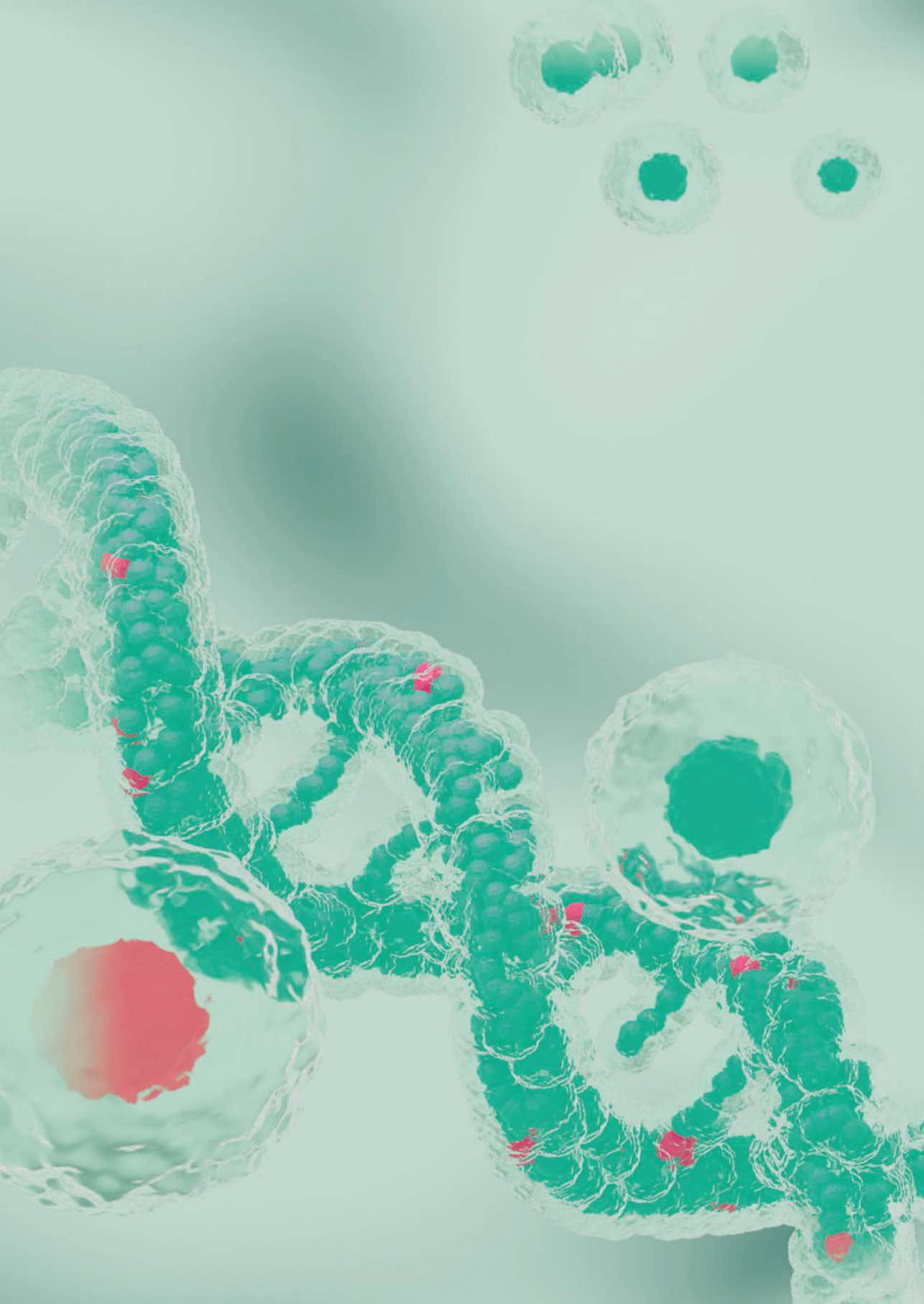
I would like to extend my appreciation to the support of numerous family members and friends throughout my life who have been instrumental in my personal and professional development. I truly appreciate every one of their contributions which made this work possible.

Thank you to my promotors Dorret Boomsma and Gonneke Willemsen, as well as my co-promotor Erik Ehli. Each has provided invaluable guidance which has shaped my development as a scientist. I am grateful for their tremendous support and mentorship while navigating the various challenges throughout the course of my PhD.

I would also like to recognize all of my colleagues whose expertise and collaboration contributed to the work in this thesis. I would especially like to thank Cornelis Lambalk and Henrike Peters for their inspiring discussions which fostered exceptional collaboration.

Last but not least, I want to recognize the Avera Institute for Human Genetics, Vrije Universiteit, and the University of South Dakota for providing the educational environment to develop as an aspiring scientist. Throughout my PhD, these institutions have provided supportive and collaborative relationships which made it possible to produce the work presented in this thesis.

TABLE OF CONTENTS



GENERAL INTRODUCTION

GENERAL INTRODUCTION

Since Gregor Mendel's initial observations of patterns of color and shape among generations of plants in the 19th century, scientists have sought to understand these patterns of inheritance and their influence in every field of biology [1]. At the core of biological heredity is the genome, a molecular blueprint of proteins and regulatory mechanisms stored in nucleic acids and written in a quaternary code of nucleotides. In humans, this data is stored in deoxyribonucleic acid (DNA), a double helix of two polynucleotide chains containing sequences of four nitrogenous bases: Adenine (A), Thymine (T), Guanine (G), and Cytosine (C). In the early 21st-century, the Human Genome Project produced the first draft of the human genome. Across 23 pairs of chromosomes and 3.3 billion nucleotide base pairs (A paired with T and G with C), this genome contains the fundamental code of every protein for human development, reproduction, cellular function and the numerous mechanisms to control their expression. Heredity is achieved through reproduction where a child receives one set of 23 chromosomes from each parent, which includes half of the genetic variants referred to as inherited alleles. Studies of genetics continue to present new discoveries which further expand the scope of scientific inquiry.

Examination of heredity among humans provides a basis for understanding the genetic contribution to numerous traits and in turn medical genetics applies the knowledge gained on the genetic basis for disease to improve clinical care. Gregor Mendel first described rules of inheritance through observable traits (phenotypes) in pea plants and later in bees [2]. Thomas Hunt Morgan further explained separation of linked genetic traits within pedigrees on recombination events [3] which Alfred Henry Sturtevant used to draft the first gene map of the *Drosophila* X chromosome [4]. The discovery of restriction-fragment length polymorphisms [5] presented the opportunity to use variation in genome between individuals to develop linkage-maps for the human genome and further map disease-associated genes [6]. Like Mendel's observations, many of the first genetic diseases in humans were identified due to a clear and traceable inheritance within familial pedigrees. Labeled as Mendelian diseases, these traits are influenced by single genes (monogenic), can be distinctly categorized based on the trait (discrete), and feature recognizable Mendelian patterns of inheritance across multiple generations in pedigrees through genetic characteristics like dominant or recessive alleles and autosomal or sex-linked genes.

Generally, rare Mendelian diseases are caused by equally rare variants with high penetrance, with a large proportion of individuals with a particular genetic variant presenting with the phenotype, sparking findings for conditions like cystic fibrosis and Huntington's disease [7].

However, there are also many diseases that are not monogenic and therefore do not follow clear Mendelian patterns of inheritance. For common disorders like heart disease, diabetes, hypertension, and schizophrenia there are strong genetic contributors demonstrated by associations with genomic variants, although they do not follow clear patterns of inheritance due in part to the underlying genetic complexity. These conditions, called complex traits or quantitative traits, are themselves polygenic and a result of contributions from numerous genes found across the human genome [8]. Thus, the genetic contribution of these traits results in small amounts of variation along a spectrum or continuum. However, in most cases genetics itself only comprises a portion of the overall contribution to disease. The variance not explained by heritability is due to additional non-genetic factors, commonly described as environment. While an individual's genetics are known to play a part in the overall variation for any given complex trait, the field of genomics is faced with the challenge of untangling the complicated relationship between genetics and environment. While genetics and environment each contribute to the overall variance associated with complex traits, understanding the role and contribution of each has significant value in our understanding of overall health while also promoting better approaches to disease treatment and prevention.

In addition, recent studies have demonstrated that the distinction between genetics and environment is not as clear as it may initially appear. For example, the non-inherited alleles of parents, the half of genetic variants in the parents that were not inherited by the offspring, have been shown to influence educational attainment in offspring [9]. The association of non-inherited alleles of parents with traits in the offspring demonstrates a contribution of parent genetics to the environment of the child. Similarly, the genomes of millions of microbes both on and in our bodies that comprise the microbiome play an essential role in maintaining our health. While the composition of the microbiome is shaped by the host genetics and environment [10, 11], the genetics of specific microbes also dictate their contribution to the microbiome both inside and outside the body. Further, exposure to cells from other individuals is of immunological importance for individuals receiving an organ or tissue transplant

[12] but is also a common phenomenon among women during pregnancy [13]. These areas of study have introduced a new dimension to genetics, promoting inquiry into not only if there is a genetic contribution to a particular trait but from whose genome the genetic contribution originates. At the core of this thesis are characteristics of environmental exposure to foreign genetic material and the associated implications among individuals, pedigrees, and communities.

GESTATION ENVIRONMENT AND TRANSPLACENTAL EXCHANGE

Pregnancy is a complex biological process that results in many physiological changes to support the growing fetus. The mode of reproduction in humans is placental viviparity, thus the developing fetus is dependent on the development of the placenta which facilitates the exchange of nutrients and waste with the maternal circulation, ultimately supporting fetal bioenergetic processes [14]. This mechanism is critical for fetal survival and therefore the placenta is the first complete organ to develop following conception. The placental trophoblasts also facilitate remodeling of the uterine structure to infiltrate the maternal vasculature [15]. As a result, these cells directly interface with the maternal circulation and are involved in maternal immunoregulation. Furthermore, while the placenta serves as a barrier to separate the maternal and fetal circulatory systems, it is imperfect and frequently results in exposure to foreign cellular and genomic material [13, 16].

Transplacental exchange has been known to have serious clinical implications in the pathogenesis of pregnancy related complications. For example, it has been well-described that maternal exposure to rhesus (Rh) D, a red blood cell (RBC) antigen, can result in a strong allosensitization response. Throughout pregnancy small populations of maternal and fetal blood cells are transfused across the placental barrier, generally increasing in concentration with gestational age with the largest prevalence found immediately following delivery [17]. When Rh D-negative mothers give birth to Rh D-positive offspring, who inherited the Rh D allele from the father, this small population of Rh D-positive RBCs enters the maternal circulation and are subject to maternal immune system scrutiny. Exposure to the highly immunogenic Rh D antigen, most commonly, facilitates the production of Rh D alloantibodies [18]. Prior to routine administration of prophylactic Rh immunoglobulin, anti-D IgG

immunoglobulins resulting from exposure to Rh D-positive RBCs in a Rh D-negative patient caused the most prevalent and severe form of hemolytic disease of the newborn in subsequent pregnancies [18]. Interestingly, this phenomenon has not only been implicated in mothers but also in their offspring. The so called "grandmother theory" suggests that in specific cases anti-D alloantibody production may originate from exposure to Rh-positive maternal cells *in utero* [13, 18-20].

The knowledge derived from studies of hemolytic disease of the newborn have furthered the understanding of interaction at the placental interface. However, much is still unknown regarding the consequences of this incomplete barrier. Similar to the exchange between mother and fetus, twins present a unique intrauterine scenario where blood may not only be exchanged with the mother but also between co-twins. Twin-to-twin transfusion syndrome (TTTS) has been known to be a primary concern for monochorionic pregnancies, most commonly associated with monozygotic (MZ) twinning, and without treatment is the leading contributor to fetal loss prior to viability [21]. With incidental findings of monochorionic dizygotic (DZ) twinning [22], the possibility for additional exposure to allogeneic cell populations has raised additional questions regarding the consequences of sustained interactions due to vascular anastomoses (that is, interconnections formed between blood vessels). Regardless of whether cell populations are of child, twin, or maternal origin, the persistent exposure to foreign cells as a result of transplacental exchange has been described as the leading cause for persistent foreign cell populations, labeled chimerism.

CHIMERISM

The term chimera originates from a creature of the same name in Greek mythology, which features a lion's head, the body of a goat and tail of a snake. Representing the amalgamation of several unique parts to create a whole, the term chimera has been used to describe similar phenomena in biology. One such example is chimeric fusion proteins, which are single proteins that comprise of functional properties of different proteins. Chimeric fusion proteins are routinely used by researchers to create unique tools in modern molecular biology, such as adding fluorescence from green fluorescent protein to study the expression and localization of a protein of interest.

Chimerism as a condition describes the occurrence of an individual organism comprised of an amalgamation of genetically distinct cell populations that originate from two or more unique zygotes. Different forms of chimerism are broadly defined by the mechanism of donor cell acquisition, generally producing different presentations of acquired cell localization within the organism. The most familiar category is surgical or artificial chimerism, produced by transplantation of donor organs or tissues into a host patient. In this scenario it is clear, that the tissues that originate from the donor are genetically distinct from the host and will present within the tissue of the graft. Second is the most extreme form of chimerism, known as tetragametic chimerism. Normally, when two eggs are fertilized this would produce DZ twins. However, in the case of a tetragametic chimera, the two embryos fuse together producing a single offspring. The resulting individual organism will possess tissues and organs comprised of cells that originate from one or both zygotes at various concentrations.

In this thesis I will primarily focus on transplacental chimerism, another form of naturally occurring chimerism. Transplacental chimerism can arise from cellular exchange across the placenta between mother and child. During pregnancy but also postpartum, low levels of fetal cells can be identified in maternal peripheral blood. When the concentration of donor cells is less than 1%, this is called microchimerism. Additionally, twin pairs present a unique scenario for transplacental chimerism, where co-twins directly exchange cells via the placental vasculature. Early efforts to understand the chimerism phenomenon have examined twins as a model for exposure to foreign cell populations which has improved with new methodologies and technological advances [23, 24].

When an individual has acquired allogeneic cells (either through surgical transplantation or natural occurrence) this chimerism presents a complex immunological challenge. In modern medicine, the use of donated organs and tissues has provided valuable treatment options for improving patient outcomes. While replacing damaged and failing organs with healthy counterparts or replenishing low blood volume with donated blood products provides the body with the essential components for normal functioning, this must be balanced with the high sensitivity of the adaptive immune system for discriminating self and non-self cells. As is well described in the field of transplantation medicine, introducing genetically distinct material to an individual is complex with immunological challenges leading to host

rejection or graft versus host disease. To prevent this, the characteristics of the tissue, organ, or blood product in question are examined closely to find a suitable match with the least likely chance of subsequent immune complications via crossmatching [25, 26]. Responses to organ transplantation such as graft versus host disease or graft rejection can produce a rapid and robust immune response that is especially dangerous for the host, and therefore these patients are generally placed on long-term immunosuppressant regimens [27].

While the immunological response to RBC alloantigens has been well established, the complex relationship between fetal and maternal microchimerism in other diseases is an area for further exploration [28]. The female predilection for autoimmune disease has made fetal microchimerism an attractive subject for investigation as a contributor to disease pathogenesis. Studies of microchimerism have suggested associations between detectable allogeneic cell populations and autoimmune disease in the host [29-37]. Additionally, autoimmune diseases have also been found to affect children with unknown etiology. Like the “grandmother theory” in production of anti-D alloantibodies, the exposure to maternal cells in maternal microchimerism has been theorized to be involved in disease pathogenesis [34, 35, 38]. Exposure to cell populations which originate from other zygotes may have numerous biological implications in both health and disease, while also indicating prior environmental exposures to other cell populations which may provide insight into disease etiology. Therefore, the chimerism phenomenon presents a source of human exposure to foreign genetic material and there are imperative questions of who in a population is most affected by chimerism and which diseases are influenced by exposure to these foreign materials. In this thesis, I investigate these questions to further understand the underlying associations in foreign cell acquisition resulting in chimerism and contributions of chimerism to disease etiology.

MAYER-ROKITANSKY-KÜSTER-HAUSER SYNDROME

Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome is a disorder of sexual development affecting 1 in 5000 females [39]. The clinical presentation of MRKH syndrome includes females with a normal (46,XX) karyotype and normal female secondary sex characteristics along with congenital aplasia of Müllerian structures including the uterus and upper two-thirds of the vagina. These absent or

underdeveloped structures have significant biological and psychosocial implications for the patient as a they result in primary amenorrhea and the inability to carry a pregnancy, which has only recently been able to be addressed in clinical trials via uterus transplantation [40]. This disorder is further categorized into two types: type I MRKH syndrome which features the aforementioned aplasia of the Müllerian structures, and type II (atypical) MRKH syndrome which presents with additional renal or skeletal malformations [41]. While classified as a rare disease, an estimate of prevalence based on a large population study in Denmark revealed no significant variation in annual live births diagnosed with MRKH, implying a common frequency of occurrence in the population [39]. Despite this classification as a rare disease there has been no causal mutation or genomic association that has been identified with relation to MRKH.

The MRKH syndrome phenotype is not unique to humans. A remarkably similar phenotype to MRKH has been long documented in cattle, called freemartinism, which has also been recorded in other mammalian species [42]. The freemartin is a sterile female calf, presenting with varying degrees of masculinization of the female reproductive tract, including vaginal agenesis and absence of the Müllerian structures. Freemartinism in cattle is a common occurrence in the female co-twin of opposite-sex dizygotic twins [43]. Further investigation of the condition has revealed XX/XY (two populations of cells, each with different sex chromosomes) microchimerism in the peripheral blood of freemartins [44, 45] and implied that exposure of the female to anti-Mullerian hormone produced by the testes of the male via twin-to-twin transfusion as the cause of the freemartin phenotype [46]. Freemartin syndrome is a burden for livestock farmers where female animals with functional reproduction are valued for lactation and future breeding in the herd [42].

The freemartin condition illustrates the significant contribution of the gestational environment and timing of exposure to foreign material which may indicate a similar phenomenon in the etiology of MRKH syndrome. Based on the observation of XX/XY microchimerism among freemartins, I inquire if this similarity between MRKH syndrome and freemartin syndrome can be demonstrated in females with MRKH syndrome by measuring detectable male microchimerism as a result of exposure to male blood from a co-twin *in utero*.

TWINNING AND TWIN DESIGN

The twinning phenomenon has long fascinated humanity. Twins present a captivating concept for artists and scientists alike, individuals with considerable commonality yet each uniquely independent. As such, many philosophers and writers over human history have included both identical and fraternal twins as literary devices to represent remarkable similarities as in William Shakespeare's *The Comedy of Errors*, symbolize natural duality such as Artemis and Apollo of Greek mythology, or illustrate divergence and reunion such as Luke Skywalker and Princess Leia in the Star Wars saga. Further, Saint Scholastica, regarded as the founder and patron saint of the Benedictine nuns, is said to be the twin sister of Saint Benedict of Nursia and the two had parallel devotion in their faith. Regardless of the narrative, twins always demonstrate some exceptional commonality and close bond that binds them.

In early human history, twins were seen as an enigma explicitly explained by the supernatural or divinity. However, scientific research on the genetics of twins has provided biological explanations for the two different categories of multiples. Twins that genetically are (nearly) identical are termed monozygotic (MZ) as their genetic similarity is explained by originating from a common zygote that split in early embryo development. Alternatively, twins that on average share half of their genetic variation are called dizygotic (DZ) twins and are explained by two or more separate fertilization events (i.e. analogous to singleton siblings) producing as many unique embryos. As DZ twins each originate from a unique zygote, they may present as DZ same-sex twins (DZss) or DZ opposite-sex twins (DOS). The genetic differences between twin types presents an invaluable resource for investigating heritability and further calculating the contributors of variance for numerous traits [47].

To understand the contribution of genetic variance on a particular trait, researchers first needed to develop a study design that would address the relative contribution of different sources of variation. In 1875, Francis Galton published the article "The history of twins, as a criterion of the relative powers of nature and nurture" which is regarded as the first recorded work that details how similar and dissimilar twins could enable the investigation of inheritance [48], although he did not know about the distinction of MZ and DZ twins. Beginning in the early 20th Century, researchers recruited twins to study contributors to phenotypic variance among humans, as a type of "natural experiment" featuring control of genetic similarity and early-life

environmental factors. It was Hermann Werner Siemens, a dermatologist, who in 1924 first applied the classical twin design by examining the correlation of moles or birthmarks as a phenotype between identical and non-identical twins [49]. This study found that identical twins have a greater correlation of mole counts compared to non-identical twins, leading to an association between genetic and phenotype resemblance.

Since then, twin studies have been fundamental to the understanding of heritability among numerous measurable phenotypes including anthropometric traits, biomarkers, behavior, and disease. As previously described, by incorporating both MZ and DZ twins the classical twin design examines the correlation between pairs and is used to ascertain the contribution of different components of phenotypic variance including both genetic and environmental sources.

THE NETHERLANDS TWIN REGISTER AT VRIJE UNIVERSITEIT

Given the well-established scientific merits of twin studies in examining the nature-nurture debate, study designs incorporating twins have only become more invaluable for understanding human biology and psychology. Exploring the nuances of phenotypic variation due to genetics and environment requires a considerable amount of meticulously documented genetic and longitudinal phenotypic data on twins and their families. To address this, regional and national twin registries continue to be established to provide an invaluable resource for twin research. One such example is the Netherlands Twin Register (NTR) which was founded in 1987 [50] and is maintained by the Department of Biological Psychology at the Vrije Universiteit (VU) Amsterdam. For over three decades the NTR, has collected and maintained a repository of biological and phenotypic data from twins and their families [51].

With over 120,000 multiples and a nearly equivalent number of non-twin family members [51], the resource curated by the NTR at the Vrije Universiteit has proven to provide a robust dataset for investigation of human complex traits. After beginning systematic recruitment of twin-pairs there has been an excellent representation of Dutch twins in the birth years of 1970-1981 and 1987-2017 at 29% and 52% respectively [51]. Surveys provide ongoing collection of longitudinal data for investigation on core

research interests including development, health, personality, as well as emotional and behavioral problems.

Additional efforts by the NTR have included the collection of biological samples to match with all collected phenotypic data (Table 1). Biological samples have been collected to produce a comprehensive dataset for future multi-omics and biomarker studies [52, 53]. Biobanking efforts have utilized whole blood collection for measurement of several biomarkers and is used as the primary source of genetic material (e.g. DNA and RNA) for adult twins. Additionally, buccal epithelial cell samples are routinely collected via a less invasive oral swab method and are the source of genetic material [54]. These samples can be assessed via polymerase chain reaction (PCR; described later), DNA microarrays and DNA sequencing for direct interrogation of hundreds of thousands of genetic variants across the genome.

Table 1. *Netherlands Twin Register data as of April 2019 reported by Ligthart et al. (2019) [51]*

	Total N
NTR Pedigrees	
Total N	255785
Multiples	123157
Twins	119040
Triplets	2613
Quadruplets	52
Quintuplets	10
Biological Samples	
DNA (blood or buccal)	29138
Genomewide SNPs	26146
DNA methylation whole blood 450K	3025
DNA methylayion buccal 850K	22
DNA sequencing whole blood	1526
RNA expression U219 array	3370
RNA sequencing	1606
RNA sequencing Lipopolysaccharide-stimulated samples	390
Metabolomics in serum or plasma	6485
Metabolomics in urine	1564
Microbiome	568
Hair sample	222

Through the generous contribution of their phenotypic data via surveys and biological samples the twin-participants and their family members have provided an invaluable resource for enabling studies of the interplay between the human genome and environment. The extensive database curated and maintained by the NTR provides a resource for both individual population-based studies and contributions to larger global consortia to investigate variance in complex traits. In this thesis, I examined biological materials from twin pedigrees in the NTR to investigate sources of microchimerism through examination of patterns and characteristics within families.

THE NETHERLANDS TWIN REGISTER AND AVERA INSTITUTE FOR HUMAN GENETICS COLLABORATION

To better explore the association between complex traits and the human genome, there is need for large-scale genetic studies. To this end, the successful execution of these studies requires the integration of highly specialized skill sets for both the wet-lab generation of high-quality genetic datasets and dry-lab bioinformatics data analyses. It is uncommon for institutions to possess both the necessary laboratory instrumentation for high-throughput measures of genetic variation and sufficient computational resources, experience, and expertise for downstream analyses. It is for this reason that institutions may leverage the expertise of others equipped with the knowledge and tools to collectively achieve what they alone cannot. As a solution for accomplishing these complex research goals, the Avera Institute for Human Genetics (AIHG) and NTR began collaboration in 2008 followed by a formal agreement beginning in 2015.

The foundation of the NTR-AIHG collaboration is the collective goal to aid the improvement of both physical and mental health through the collection and investigation of genetic data. As a part of this collaboration, AIHG provides technical laboratory expertise for conducting a multitude of molecular genetics experiments for data generation, whereas the NTR provides expertise in biobanking practices, collection of phenotypic data and the statistical methods for conducting twin-family research in addition to a large repository of collected samples. Among the many molecular genetics laboratory services provided by the AIHG, for the purpose of establishing genetic variation between individuals, advances in DNA microarray

technology have allowed for cost-effective genome-wide interrogation of select markers with known variation in the population [55]. Together, the NTR and AIHG have applied their collective expertise in the design of custom DNA microarrays, the Affymetrix Axiom-NL array [56] and the Illumina Global Screening Array (GSA) [57]. The collective genotyping initiative by NTR-AIHG has provided data to support a diverse array of genetic association studies in human traits and disorders, including among others behavior and emotion [58-75], female fertility and twinning [76, 77].

In addition, AIHG has employed numerous molecular genetics methodologies and instrumentation to facilitate research initiatives within the collaboration with the NTR to better understand genetic contributions to complex traits. Epigenetic studies to investigate environmental influences on genetic expression through methylation patterns across the genome have been captured at the AIHG using bisulfite conversion and specially designed Illumina HumanMethylation450K and MethylationEPIC BeadChip DNA methylation arrays [78-81]. Studies of human gut microbiome have utilized next-generation (massively parallel) sequencing (NGS) at AIHG for 16S rRNA sequencing to examine the composition and diversity of microbial communities [10, 11, 82]. In addition, real-time quantitative polymerase chain reaction has previously been applied for the study of telomere repeat mass of blood and buccal derived DNA [83]. The numerous molecular genetics methodologies supported by AIHG provide a variety of tools promoting current and future research interests of the collaboration to better understand human genetic variation and its role in health and disease. Similarly, throughout this thesis I apply molecular genetic methodologies at AIHG to investigate the frequency, concentration, and composition of biological samples for minor populations of genetic material.

POLYMERASE CHAIN REACTION

Central to the molecular genetic techniques I apply in this thesis is the polymerase chain reaction. The polymerase chain reaction (PCR) has been fundamental in establishing the modern molecular genetics laboratory by allowing for the efficient production of a large quantity of replicates of the target sequence with as little as a single copy of the template in the primary sample. Originally developed by Kary B. Mullis, PCR is an *in vitro* laboratory technique for the specific enzymatic amplification of a template DNA sequence [84, 85]. The methodology exploits the

natural characteristics of nucleic acids for replication of genetic information which allows for reliable duplication of genomic information in either semi-conservative genome replication or transcription via base-pair matching. The PCR reaction is able to produce specific amplification of a region of DNA selected by the annealing of two primer sequences, one complementary to the positive-sense (sense) strand and the other complementary to the negative-sense (antisense) strand of the DNA double helix.

The fundamental reaction is comprised of a DNA polymerase I (Taq polymerase) originating from *Thermus aquaticus*, a thermophilic bacterium, in addition to nucleic acids and target-specific primers. Combined with template DNA these reagents provide the necessary molecular components for DNA replication. The reaction is controlled by thermocycling, a series of temperature changes that are repeated over a specified number of cycles, to achieve efficient replication (Figure 1A). This process of thermal cycling is repeated over many cycles, each duplicating the amount of genomic material between the two primers, as long as there is efficient primer binding and sufficient resources (Figure 1B).

The PCR technique further provides the essential tools for investigation of rare genomic targets in complex samples. The products of PCR have provided the means for geneticists to have sufficient material to study the genome through conventional laboratory methods and develop innovative solutions to further life sciences and genetics-based research. Further, these techniques have been able to produce revolutionary methodologies that have reshaped diagnostic medicine [86]. The base-pair matching of primer sequences provides an advantage in test sensitivity and specificity through precise amplification of genomic target sequences, adding to utility in detection of microorganisms. As PCR thermocycling produces results within a few hours, the advantage of PCR in molecular diagnostics is a reduction in sample preparation time and elimination of microorganism culture for analysis, thus reducing the turnaround time for clinical diagnosis and treatment. The advantages of PCR in clinical diagnostics are of critical importance for providing patients, physicians, and public health officials with fast and accurate information for infectious disease control and prevention [86-88].

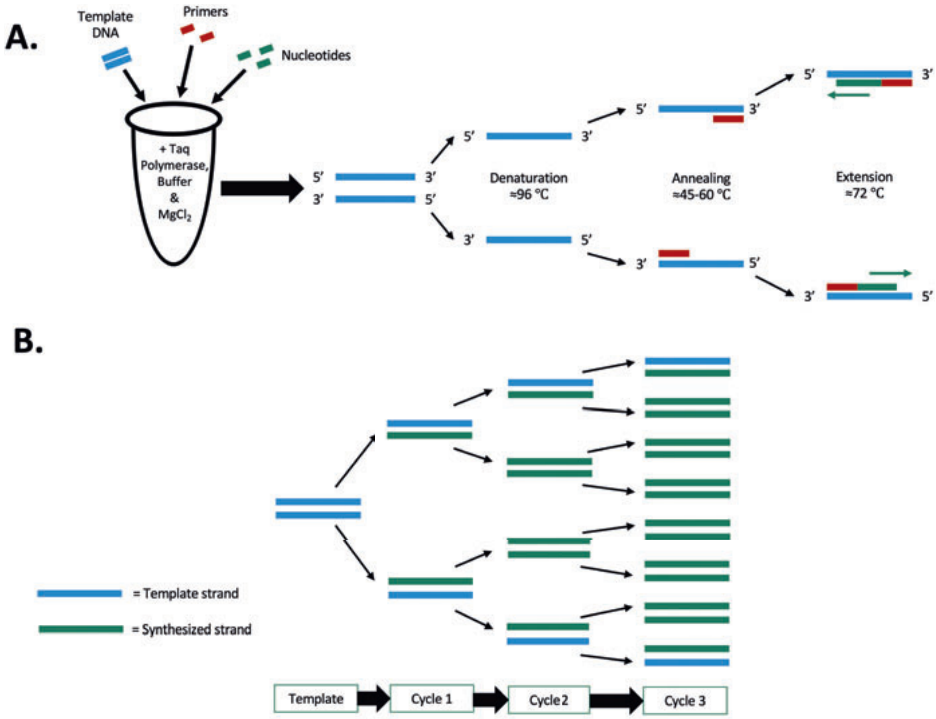


Figure 1: *Diagrams of PCR amplification: (A) Template DNA (blue), primers (red), and nucleotides (green) are combined with PCR master mix and proceed through one cycle of PCR thermal cycling. Thermal cycling includes denaturation of the DNA, annealing of the primers, and extension to synthesize the new strand. (B) PCR amplification of a single template DNA molecule (blue) is able to produce many copies of newly synthesized DNA (green) through duplication during each cycle.*

QUANTITATIVE PCR

While the advent of PCR changed the landscape of molecular biology by providing a valuable tool for the amplification of genomic material for numerous applications, the use of end-point analysis provides only one of two possible answers for a particular sample, target detection or no-detection. To measure the amount of material produced during the PCR reaction two methods have become available: through detection of double-stranded DNA [89] and through fluorescent signal produced by molecular probes following degradation by Taq polymerase exonuclease activity [90]. Both these methods result in a direct relationship between the amount of DNA amplicons produced in the reaction and measure of detectable fluorescence. While each methodology provides the ability to examine reaction kinetics, the added

specificity of molecular probes which require base-pair matching of a specific target sequence made these preferrable for application in the high sensitivity testing developed and implemented in this thesis.

Advanced technology allowing for the reading of fluorescence following each PCR cycle in each individual reaction has made it possible to observe the kinetics of PCR amplification in real-time. This real-time PCR introduced a new measure, the cycle-threshold (Ct), which is the number of the cycle during thermocycling which the amplification growth curve of the sample becomes distinguishable from background signal and crosses a fluorescence threshold defined by exponential amplification. The Ct value produced for a sample has an inverse relationship with the amount of starting material in the reaction. Parallel testing a series of known standard concentrations allows researchers to produce a linear regression model to provide an accurate measure of the amount of starting material in an unknown sample. The Ct metric provides an additional quantitative dimension to PCR data and resulted in the quantitative PCR (qPCR) technique.

The simple detection of genomic targets through traditional end-point PCR have value in providing results for clinical diagnostics where detection alone is the only relevant measure for establishing a treatment plan. However, in this thesis I investigate if there is measurable variation in the genomic concentration of foreign genomes, related to both chimerism and viral infection, in samples between different groups or timepoints. Simply, is the concentration of material derived by qPCR or variation in Ct values between timepoints informative for explaining differences in presentation of foreign genetic material within populations?

SARS-COV-2

The application of real-time PCR in molecular diagnostics provides highly sensitive and specific detection of microbe genomes in patient samples which is invaluable for managing the spread of emerging pathogens. In December of 2019 an unexplained outbreak of acute respiratory disease was reported in Wuhan, China. The condition, coronavirus disease 2019 (COVID-19), was discovered to be caused by a novel *Betacoronavirus* of the *Cononaviridae* family; the virus was later named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [91]. The COVID-19 condition

presented with symptoms largely similar to other respiratory diseases, including fever, cough, sore throat, headache, and malaise in addition to anosmia (loss of smell) and ageusia (loss of taste) [92, 93] with severity ranging from asymptomatic to hospitalization and death.

Over the following months SARS-CoV-2 rapidly spread across the globe resulted in a pandemic declaration by the World Health Organization in March 2020. The seemingly remarkable efficiency of viral transmission within populations resulted in sudden rapid increases in disease prevalence within geographic regions, most notably in areas of high population density [94-96]. Transmission of the virus via respiratory droplets allows for rapid spread of viral particles to new hosts [97]. Among infected individuals, the SARS-CoV-2 virus is transmissible while asymptomatic or presymptomatic [98, 99], increasing the importance of early detection of infected individuals.

With the goal of mitigating spread and reducing community exposure, among the leading strategies for controlling the outbreak was the test, trace and isolate model [100]. To employ this tactic, local governments and public health agencies were dependent on real-time diagnostic laboratory testing to identify cases and trace potential contacts. The application of PCR-based molecular diagnostic assays had been well established for use in confirmatory testing of numerous microbes and therefore PCR testing of samples extracted from upper-respiratory samples became the most accessible methodology early in the pandemic response. The expertise and capabilities at AIHG for developing and implementing molecular genetic testing were directly translated during the COVID-19 pandemic to provide real-time diagnostics and surveillance for Avera Health and the surrounding community. Applying nucleic acid extraction techniques and high-sensitivity real-time PCR for the detection of foreign viral genome in upper respiratory samples, the laboratory provided real-time diagnostics to patients, physicians, and public health officials for monitoring the disease in the local community.

In contrast to humans and other organisms that use DNA to store their genomic information, the genome of Coronaviridae family of viruses is contained in single-stranded ribonucleic acid (RNA). The coronavirus genome has four structural proteins: spike (S), envelope (E), membrane (M), and nucleoprotein (N). Among the SARS-CoV-2 structural proteins, the spike protein is responsible for direct interaction

with angiotensin-converting enzyme 2 (ACE2), the receptor on human cells. To gain entry to the human host cell, the spike protein is cleaved into S1 and S2 subunits where the interaction of the receptor binding domain (RBD) on the S1 protein and the ACE2 receptor promotes virus entry [101-103].

In late 2020, the emergence of the first variant of concern, B.1.1.7 or Alpha, was identified with key mutations present in the RBD, resulting in enhanced virus transmissibility and antibody evasion [104-106]. Further monitoring of the virus genome for variants of concern is maintained through genomic surveillance of the SARS-CoV-2 genome via genomic sequencing of a randomized selection of PCR confirmed positive samples representative of the population. In response to news of novel variants, AIHG proceeded with the development and implementation of a genomic surveillance program providing up-to-date information on the prevalence of emerging SARS-CoV-2 variants in the local community. The data produced from genomic surveillance equips epidemiologists and public health officials with the necessary information to effectively monitor the unique dynamics of virus variants, evaluate policy changes in response to changes in community spread, and allocate the necessary resources to address anticipated changes in case numbers.

During the pandemic, there was significant interest in applying the Ct value of real-time PCR diagnostic tests as a proxy for viral load in clinical patients. Indeed, samples tested via PCR for other viruses such as HIV [107], hepatitis B [108] and C [109] which are collected via blood samples use Ct measures to quantify viral load in individual patients. However, due to upper respiratory swab collection methods, the variable nature of the collection process may influence Ct value in SARS-CoV-2 tests resulting in inaccurate measurement and flawed interpretation of the data. As a result, many scientists and leading professional organizations like the American Association for Clinical Chemistry (AACC), Association for Molecular Pathology (AMP), and Infectious Diseases Society of America (IDSA) issued statements expressing caution for application of Ct values to inform clinical decisions [110, 111]. However, real-time PCR data may have value as a relative measure of magnitude within a population. Through examination of the SARS-CoV-2 testing conducted at AIHG I inquire if the average Ct value in the AIHG testing population may reflect differences in the characterization of clinical presentation and virus dynamics over time.

HYPOTHESIS AND OBJECTIVES OF THE DISSERTATION

This dissertation examines the application of molecular genomics techniques for the investigation of small populations of foreign genetic material from biological samples, i.e. detection of chimerism and viral pathogens. Here I hypothesized that measure of small populations of genetic material will provide insight into the prevalence and propagation of health conditions. Throughout this dissertation I test hypotheses of prior exposures to male blood via measures of chimerism in clinical cases and controls for a rare disorder and further investigate potential risk factors of chimerism through prevalence and concordance within twin pedigrees. In addition, I examine the high sensitivity assays that were developed for the specific detection of SARS-CoV-2 in respiratory samples to investigate quantitative variation in relation to disease epidemiology.

Chapter 2 provides a comprehensive systematic review of chimerism as it pertains to human health and behavior. This chapter details the chimerism phenomenon and provides distinction between disparate forms of the condition with a central focus on naturally occurring chimerism. Available methodologies employed in the literature for the detection of small concentrations of unique cells are discussed for use in studies of chimerism and evaluated for sensitivity relevant to application in study of microchimerism. Prior studies are summarized as the findings pertain to both animal studies and investigations on human diseases. A substantial amount of research in chimerism is focused on the potential pathogenic implications due to the inherent immune complexities of persistent exposure to allogeneic cell populations. Additional examinations of chimerism studies related to behavior are discussed including current hypotheses and areas of further investigation.

The complex environment presented by chimerism has led to many proposals for pathogenic, protective, or neutral roles in the course of disease and there is significant need for further research to understand the underlying mechanisms of this phenomenon. Both chapters 3 and 4 address specific scientific questions regarding the prevalence and implications of human microchimerism.

Chapter 3 addresses a hypothesis of exposure to male blood in utero as a contributor to the etiology of Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome. The known

phenotypic similarity to the freemartinism condition in cattle has led to the hypothesis that MRKH syndrome is in part caused by *in utero* exposure during fetal development to Anti-Mullerian Hormone (AMH), potentially via blood transfusion through placental vascular anastomoses from a male co-twin. The objective of this study was to explore this hypothesis in adult women diagnosed with MRKH, using a Y chromosome specific assay for the detection of male genome due to any resultant persistent male microchimerism. Investigation of microchimerism quantitation and prevalence were examined and compared in both case and control populations. Furthermore, additional sources of male cells including male offspring and older brothers (through maternal exposure) were considered. This study further determined if persistent chimerism among the population of cases indicates prior exposure to male genetic material support

Chapter 4 expands upon the findings presented in the previous chapter and examines patterns of microchimerism among female members of twin pedigrees. With a notable prevalence of detectable microchimerism in the population and detection in approximately 13% of women and adolescents with no history of pregnancy [112, 113], there is significant interest in understanding the sources of exposure to foreign cells and genetic material. As previous research has indicated blood chimerism to occur among human multiples [24], the objective of this study was to investigate the numerous relationships within pedigrees for patterns of male microchimerism. We hypothesized that pedigree features such as having a male co-twin, older brother, or male offspring would increase the risk of persistent male microchimerism detectable in female members. For this study, families of female monozygotic (MZ) twins, female dizygotic same-sex (DZss) twins and dizygotic opposite sex (DOS) twins were included if the female twin(s), a non-twin sister, and the mother of twins had previously donated peripheral blood samples in the Netherlands Twin Register (NTR). This study specifically examines the prevalence of male microchimerism between the types of family members and concordance for pairs of family members as well as potential associations with having an older brother, male offspring, or both.

Chapter 5 extends these real-time PCR measures of minor genomic populations to clinical laboratory molecular diagnostics and explores if real-time measures of human exposure to foreign material of a viral pathogen may provide a valuable measure for monitoring dynamics of viral infections within a community. This chapter describes the development, validation, implementation, and evolution of a diagnostic

laboratory developed test and workflow to address the dynamics of the ongoing COVID-19 pandemic at AIHG. In addition, SARS-CoV-2 variants that emerged at the end of 2020 presented an additional challenge addressed by genetic testing and analysis. Details on the development of laboratory workflows and analysis pipelines for both targeted sanger sequencing and whole genome sequencing of SARS-CoV-2 for genomic surveillance are also presented. Analyses pertaining to changes in quantitative testing metrics are examined in relation to clinical characteristics and changes in local epidemiology through future demand for healthcare services. Categorized SARS-CoV-2 variants via genomic surveillance are further examined for changes in real-time PCR to examine the effects of variant transition. Ultimately, the information presented in this chapter illustrates the application of diverse sensitive molecular genomics techniques and analyses for providing epidemiological data for management of emerging pathogens at the local level.

Chapter 6 presents a summary of the main findings of the prior chapters. This is followed by general conclusions of this work and discussion of broad perspectives on the field and future for investigation of foreign genetic material from mixed samples in human health. Chapter 7 concludes the dissertation with a succinct summary of the thesis.

REFERENCES

1. Strachan, T. and A. Read, *Human Molecular Genetics*. 2018: CRC Press.
2. van Dijk, P.J., F.J. Weissing, and T.H.N. Ellis, *How Mendel's Interest in Inheritance Grew out of Plant Improvement.* Genetics, 2018. **210**(2): p. 347-355.
3. Morgan, T.H., *Random Segregation Versus Coupling in Mendelian Inheritance.* Science, 1911. **34**(873): p. 384.
4. Sturtevant, A.H., *The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association.* Journal of Experimental Zoology, 1913. **14**(1): p. 43-59.
5. Kan, Y.W. and A.M. Dozy, *Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation.* Proc Natl Acad Sci U S A, 1978. **75**(11): p. 5631-5.
6. Bates, G.P., *History of genetic disease: the molecular genetics of Huntington disease - a history.* Nat Rev Genet, 2005. **6**(10): p. 766-73.
7. McCarthy, M.I., et al., *Genome-wide association studies for complex traits: consensus, uncertainty and challenges.* Nat Rev Genet, 2008. **9**(5): p. 356-69.
8. Slunecka, J.L., et al., *Implementation and implications for polygenic risk scores in healthcare.* Hum Genomics, 2021. **15**(1): p. 46.
9. Kong, A., et al., *The nature of nurture: Effects of parental genotypes.* Science, 2018. **359**(6374): p. 424-428.
10. Finnicum, C.T., et al., *Cohabitation is associated with a greater resemblance in gut microbiota which can impact cardiometabolic and inflammatory risk.* BMC Microbiol, 2019. **19**(1): p. 230.
11. Finnicum, C.T., et al., *Metataxonomic Analysis of Individuals at BMI Extremes and Monozygotic Twins Discordant for BMI.* Twin Res Hum Genet, 2018. **21**(3): p. 203-213.
12. Lee, T.H., et al., *Survival of donor leukocyte subpopulations in immunocompetent transfusion recipients: frequent long-term microchimerism in severe trauma patients.* Blood, 1999. **93**(9): p. 3127-39.
13. Lo, Y.M., et al., *Two-way cell traffic between mother and fetus: biologic and clinical implications.* Blood, 1996. **88**(11): p. 4390-5.
14. Maltepe, E. and S.J. Fisher, *Placenta: the forgotten organ.* Annu Rev Cell Dev Biol, 2015. **31**: p. 523-52.
15. Bracamonte-Baran, W. and W. Burlingham, *Non-inherited maternal antigens, pregnancy, and allotolerance.* Biomed J, 2015. **38**(1): p. 39-51.
16. Ariga, H., et al., *Kinetics of fetal cellular and cell-free DNA in the maternal circulation during and after pregnancy: implications for noninvasive prenatal diagnosis.* Transfusion, 2001. **41**(12): p. 1524-30.
17. Zipursky, A. and L.G. Israels, *The pathogenesis and prevention of Rh immunization.* Can Med Assoc J, 1967. **97**(21): p. 1245-57.
18. Avent, N.D. and M.E. Reid, *The Rh blood group system: a review.* Blood, 2000. **95**(2): p. 375-87.
19. Hattevig, G., et al., *Screening of Rh-antibodies in Rh-negative female infants with Rh-positive mothers.* Acta Paediatr Scand, 1981. **70**(4): p. 541-5.

20. Taylor, J.F., *Sensitization of Rh-negative daughters by their Rh-positive mothers.* N Engl J Med, 1967. **276**(10): p. 547-51.

21. Miller, J.L., *Twin to twin transfusion syndrome.* Transl Pediatr, 2021. **10**(5): p. 1518-1529.

22. Peters, H.E., et al., *Unusual Twinning Resulting in Chimerism: A Systematic Review on Monochorionic Dizygotic Twins.* Twin Research and Human Genetics, 2017. **20**(2): p. 161-168.

23. Race, R.R. and R. Sanger, *Blood groups in man.* 6th ed. 1975, Philadelphia: Oxford : Blackwell Scientific Publications.

24. vanDijk, B.A., D.I. Boomsma, and A.J.M. deMan, *Blood group chimerism in human multiple births is not rare.* American Journal of Medical Genetics, 1996. **61**(3): p. 264-268.

25. Fuchs, E.J., *Transplantation tolerance: from theory to clinic.* Immunol Rev, 2014. **258**(1): p. 64-79.

26. Lakkis, F.G. and R.I. Lechler, *Origin and biology of the allogeneic response.* Cold Spring Harb Perspect Med, 2013. **3**(8).

27. Enderby, C. and C.A. Keller, *An overview of immunosuppression in solid organ transplantation.* Am J Manag Care, 2015. **21**(1 Suppl): p. s12-23.

28. Nelson, J.L., *Microchimerism: incidental byproduct of pregnancy or active participant in human health?* Trends in Molecular Medicine, 2002. **8**(3): p. 109-113.

29. Lambert, N.C., et al., *Male microchimerism in healthy women and women with scleroderma: cells or circulating DNA? A quantitative answer.* Blood, 2002. **100**(8): p. 2845-51.

30. Johnson, K.L., et al., *Fetal cell microchimerism in tissue from multiple sites in women with systemic sclerosis.* Arthritis Rheum, 2001. **44**(8): p. 1848-54.

31. Rak, J.M., et al., *Male microchimerism and HLA compatibility in French women with sclerodema: a different profile in limited and diffuse subset.* Rheumatology (Oxford), 2009. **48**(4): p. 363-6.

32. Klintschar, M., et al., *Evidence of fetal microchimerism in Hashimoto's thyroiditis.* J Clin Endocrinol Metab, 2001. **86**(6): p. 2494-8.

33. Renne, C., et al., *Thyroid fetal male microchimerisms in mothers with thyroid disorders: presence of Y-chromosomal immunofluorescence in thyroid-infiltrating lymphocytes is more prevalent in Hashimoto's thyroiditis and Graves' disease than in follicular adenomas.* J Clin Endocrinol Metab, 2004. **89**(11): p. 5810-4.

34. Ye, J., M. Vives-Pi, and K.M. Gillespie, *Maternal microchimerism: increased in the insulin positive compartment of type 1 diabetes pancreas but not in infiltrating immune cells or replicating islet cells.* PLoS One, 2014. **9**(1): p. e86985.

35. Nelson, J.L., et al., *Maternal microchimerism in peripheral blood in type 1 diabetes and pancreatic islet beta cell microchimerism.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1637-42.

36. Nelson, J.L., *The otherness of self: microchimerism in health and disease.* Trends Immunol, 2012. **33**(8): p. 421-7.

37. Lambert, N.C., *Microchimerism and rheumatic diseases.* Joint Bone Spine, 2012. **79**(6): p. 563-4.

38. Stevens, A.M., *Maternal microchimerism in health and disease.* Best Pract Res Clin Obstet Gynaecol, 2016. **31**: p. 121-30.

39. Herlin, M., et al., *Prevalence and patient characteristics of Mayer-Rokitansky-Kuster-Hauser syndrome: a nationwide registry-based study.* Hum Reprod, 2016. **31**(10): p. 2384-90.

40. Herlin, M.K., M.B. Petersen, and M. Brannstrom, *Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome: a comprehensive update.* Orphanet J Rare Dis, 2020. **15**(1): p. 214.

41. Oppelt, P., et al., *Clinical aspects of Mayer-Rokitansky-Kuester-Hauser syndrome: recommendations for clinical diagnosis and staging.* Hum Reprod, 2006. **21**(3): p. 792-7.

42. Padula, A.M., *The freemartin syndrome: an update.* Anim Reprod Sci, 2005. **87**(1-2): p. 93-109.

43. Lillie, F.R., *The Theory of the Free-Martin.* Science, 1916. **43**(1113): p. 611-3.

44. Owen, R.D., *Immunogenetic Consequences of Vascular Anastomoses between Bovine Twins.* Science, 1945. **102**(2651): p. 400-1.

45. Peretti, V., et al., *XX/XY chimerism in cattle: clinical and cytogenetic studies.* Sex Dev, 2008. **2**(1): p. 24-30.

46. Vigier, B., et al., *Origin of anti-Mullerian hormone in bovine freemartin fetuses.* J Reprod Fertil, 1984. **70**(2): p. 473-9.

47. van Dongen, J., et al., *The continuing value of twin studies in the omics era.* Nat Rev Genet, 2012. **13**(9): p. 640-53.

48. Galton, F., *The history of twins, as a criterion of the relative powers of nature and nurture.* Fraser's Magazine, 1875. **12**(71): p. 566-576.

49. Siemens, H.W., *Die Zwillingspathologie: Ihre Bedeutung· Ihre Methodik· Ihre Bisherigen Ergebnisse.* 1924: Springer.

50. Boomsma, D.I., J.F. Orlebeke, and G.C. van Baal, *The Dutch Twin Register: growth data on weight and height.* Behav Genet, 1992. **22**(2): p. 247-51.

51. Ligthart, L., et al., *The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs.* Twin Res Hum Genet, 2019. **22**(6): p. 623-636.

52. Willemsen, G., et al., *The Netherlands Twin Register biobank: a resource for genetic epidemiological studies.* Twin Res Hum Genet, 2010. **13**(3): p. 231-45.

53. van Dongen, J., et al., *Longitudinal weight differences, gene expression and blood biomarkers in BMI-discordant identical twins.* Int J Obes (Lond), 2015. **39**(6): p. 899-909.

54. van Beijsterveldt, C.E., et al., *The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children.* Twin Res Hum Genet, 2013. **16**(1): p. 252-67.

55. Van Asselt, A.J. and E.A. Ehli, *Whole-Genome Genotyping Using DNA Microarrays for Population Genetics.* Methods Mol Biol, 2022. **2418**: p. 269-287.

56. Ehli, E.A., et al., *A method to customize population-specific arrays for genome-wide association testing.* Eur J Hum Genet, 2017. **25**(2): p. 267-270.

57. Beck, J.J., et al., *Genetic Similarity Assessment of Twin-Family Populations by Custom-Designed Genotyping Array.* Twin Res Hum Genet, 2019. **22**(4): p. 210-219.

58. Ip, H.F., et al., *Genetic association study of childhood aggression across raters, instruments, and age.* Transl Psychiatry, 2021. **11**(1): p. 413.

59. Pappa, I., et al., *A genome-wide approach to children's aggressive behavior: The EAGLE consortium.* Am J Med Genet B Neuropsychiatr Genet, 2016. **171**(5): p. 562-72.

60. Abdellaoui, A., et al., *CNV Concordance in 1,097 MZ Twin Pairs.* Twin Res Hum Genet, 2015. **18**(1): p. 1-12.

61. de Zeeuw, E.L., et al., *Polygenic scores associated with educational attainment in adults predict educational achievement and ADHD symptoms in children.* Am J Med Genet B Neuropsychiatr Genet, 2014. **165B**(6): p. 510-20.

62. Demontis, D., et al., *Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder.* Nat Genet, 2019. **51**(1): p. 63-75.

63. Ehli, E.A., et al., *De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on Attention Problems.* Eur J Hum Genet, 2012. **20**(10): p. 1037-43.

64. Groen-Blokhuis, M.M., et al., *A prospective study of the effects of breastfeeding and FADS2 polymorphisms on cognition and hyperactivity/attention problems.* Am J Med Genet B Neuropsychiatr Genet, 2013. **162B**(5): p. 457-65.

65. Middeldorp, C.M., et al., *A Genome-Wide Association Meta-Analysis of Attention-Deficit/Hyperactivity Disorder Symptoms in Population-Based Pediatric Cohorts.* J Am Acad Child Adolesc Psychiatry, 2016. **55**(10): p. 896-905 e6.

66. Baumert, J., et al., *No evidence for genome-wide interactions on plasma fibrinogen by smoking, alcohol consumption and body mass index: results from meta-analyses of 80,607 subjects.* PLoS One, 2014. **9**(12): p. e111156.

67. Liu, M., et al., *Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.* Nat Genet, 2019. **51**(2): p. 237-244.

68. Minica, C.C., et al., *Genome-wide association meta-analysis of age at first cannabis use.* Addiction, 2018. **113**(11): p. 2073-2086.

69. Benke, K.S., et al., *A genome-wide association meta-analysis of preschool internalizing problems.* J Am Acad Child Adolesc Psychiatry, 2014. **53**(6): p. 667-676 e7.

70. Mbarek, H., et al., *Genome-Wide Significance for PCLO as a Gene for Major Depressive Disorder.* Twin Res Hum Genet, 2017. **20**(4): p. 267-270.

71. Middeldorp, C.M., et al., *The genetic association between personality and major depression or bipolar disorder. A polygenic score analysis using genome-wide association data.* Transl Psychiatry, 2011. **1**: p. e50.

72. Okbay, A., et al., *Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses.* Nat Genet, 2016. **48**(6): p. 624-33.

73. Wray, N.R., et al., *Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression.* Nat Genet, 2018. **50**(5): p. 668-681.

74. DJ, V.D.M., et al., *Dopaminergic Genetic Variants and Voluntary Externally Paced Exercise Behavior.* Med Sci Sports Exerc, 2018. **50**(4): p. 700-708.

75. Huppertz, C., et al., *A twin-sibling study on the relationship between exercise attitudes and exercise behavior.* Behav Genet, 2014. **44**(1): p. 45-55.

76. Barban, N., et al., *Genome-wide analysis identifies 12 loci influencing human reproductive behavior.* Nat Genet, 2016. **48**(12): p. 1462-1472.

77. Mbarek, H., et al., *Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility.* Am J Hum Genet, 2016. **98**(5): p. 898-908.

78. van Dongen, J., et al., *Genome-wide analysis of DNA methylation in buccal cells: a study of monozygotic twins and mQTLs.* Epigenetics Chromatin, 2018. **11**(1): p. 54.

79. van Dongen, J., et al., *Epigenetic variation in monozygotic twins: a genome-wide analysis of DNA methylation in buccal cells.* Genes (Basel), 2014. **5**(2): p. 347-65.

80. Odintsova, V.V., et al., *Predicting Complex Traits and Exposures From Polygenic Scores and Blood and Buccal DNA Methylation Profiles.* Front Psychiatry, 2021. **12**: p. 688464.

81. van Dongen, J., et al., *Identical twins carry a persistent epigenetic signature of early genome programming.* Nat Commun, 2021. **12**(1): p. 5618.

82. Kurilshikov, A., et al., *Large-scale association analyses identify host factors influencing human gut microbiome composition.* Nat Genet, 2021. **53**(2): p. 156-165.

83. Finnicum, C.T., et al., *Relative Telomere Repeat Mass in Buccal and Leukocyte-Derived DNA*. PLoS One, 2017. **12**(1): p. e0170765.

84. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harb Symp Quant Biol, 1986. **51 Pt 1**: p. 263-73.

85. Mullis, K.B., *The unusual origin of the polymerase chain reaction.* Sci Am, 1990. **262**(4): p. 56-61, 64-5.

86. Post, J.C. and G.D. Ehrlich, *The impact of the polymerase chain reaction in clinical medicine.* JAMA, 2000. **283**(12): p. 1544-6.

87. van Seventer, J.M. and N.S. Hochberg, *Principles of Infectious Diseases: Transmission, Diagnosis, Prevention, and Control.* International Encyclopedia of Public Health, 2017: p. 22-39.

88. Louie, M., L. Louie, and A.E. Simor, *The role of DNA amplification technology in the diagnosis of infectious diseases.* CMAJ, 2000. **163**(3): p. 301-9.

89. Higuchi, R., et al., *Simultaneous amplification and detection of specific DNA sequences.* Biotechnology (N Y), 1992. **10**(4): p. 413-7.

90. Holland, P.M., et al., *Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase.* Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7276-80.

91. Coronaviridae Study Group of the International Committee on Taxonomy of, V., *The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.* Nat Microbiol, 2020. **5**(4): p. 536-544.

92. Gandhi, R.T., J.B. Lynch, and C. Del Rio, *Mild or Moderate Covid-19.* N Engl J Med, 2020. **383**(18): p. 1757-1766.

93. Berlin, D.A., R.M. Gulick, and F.J. Martinez, *Severe Covid-19.* N Engl J Med, 2020. **383**(25): p. 2451-2460.

94. Smith, T.P., et al., *Temperature and population density influence SARS-CoV-2 transmission in the absence of nonpharmaceutical interventions.* Proc Natl Acad Sci U S A, 2021. **118**(25).

95. Wong, D.W.S. and Y. Li, *Spreading of COVID-19: Density matters.* PLoS One, 2020. **15**(12): p. e0242398.

96. Rubin, D., et al., *Association of Social Distancing, Population Density, and Temperature With the Instantaneous Reproduction Number of SARS-CoV-2 in Counties Across the United States.* JAMA Netw Open, 2020. **3**(7): p. e2016099.

97. Cevik, M., et al., *Virology, transmission, and pathogenesis of SARS-CoV-2.* BMJ, 2020. **371**: p. m3862.

98. Arons, M.M., et al., *Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility.* N Engl J Med, 2020. **382**(22): p. 2081-2090.

99. Rothe, C., et al., *Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany.* N Engl J Med, 2020. **382**(10): p. 970-971.

100. Rajan, S., J. Cylus, and M. McKee, *Successful find-test-trace-isolate-support systems: how to win at snakes and ladders.* Eurohealth, 2020. **26**(2).

101. Samavati, L. and B.D. Uhal, *ACE2, Much More Than Just a Receptor for SARS-COV-2.* Front Cell Infect Microbiol, 2020. **10**: p. 317.

102. Shang, J., et al., *Cell entry mechanisms of SARS-CoV-2.* Proc Natl Acad Sci U S A, 2020. **117**(21): p. 11727-11734.

103. Jackson, C.B., et al., *Mechanisms of SARS-CoV-2 entry into cells.* Nat Rev Mol Cell Biol, 2022. **23**(1): p. 3-20.

104. Galloway, S.E., et al., *Emergence of SARS-CoV-2 B.1.1.7 Lineage - United States, December 29, 2020-January 12, 2021.* MMWR Morb Mortal Wkly Rep, 2021. **70**(3): p. 95-99.

105. Volz, E., et al., *Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England.* Nature, 2021. **593**(7858): p. 266-269.

106. Collier, D.A., et al., *Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies.* Nature, 2021. **593**(7857): p. 136-141.

107. Drain, P.K., et al., *Point-of-Care HIV Viral Load Testing: an Essential Tool for a Sustainable Global HIV/AIDS Response.* Clin Microbiol Rev, 2019. **32**(3).

108. Poiteau, L., et al., *Evaluation of the Xpert HBV Viral Load for hepatitis B virus molecular testing.* J Clin Virol, 2020. **129**: p. 104481.

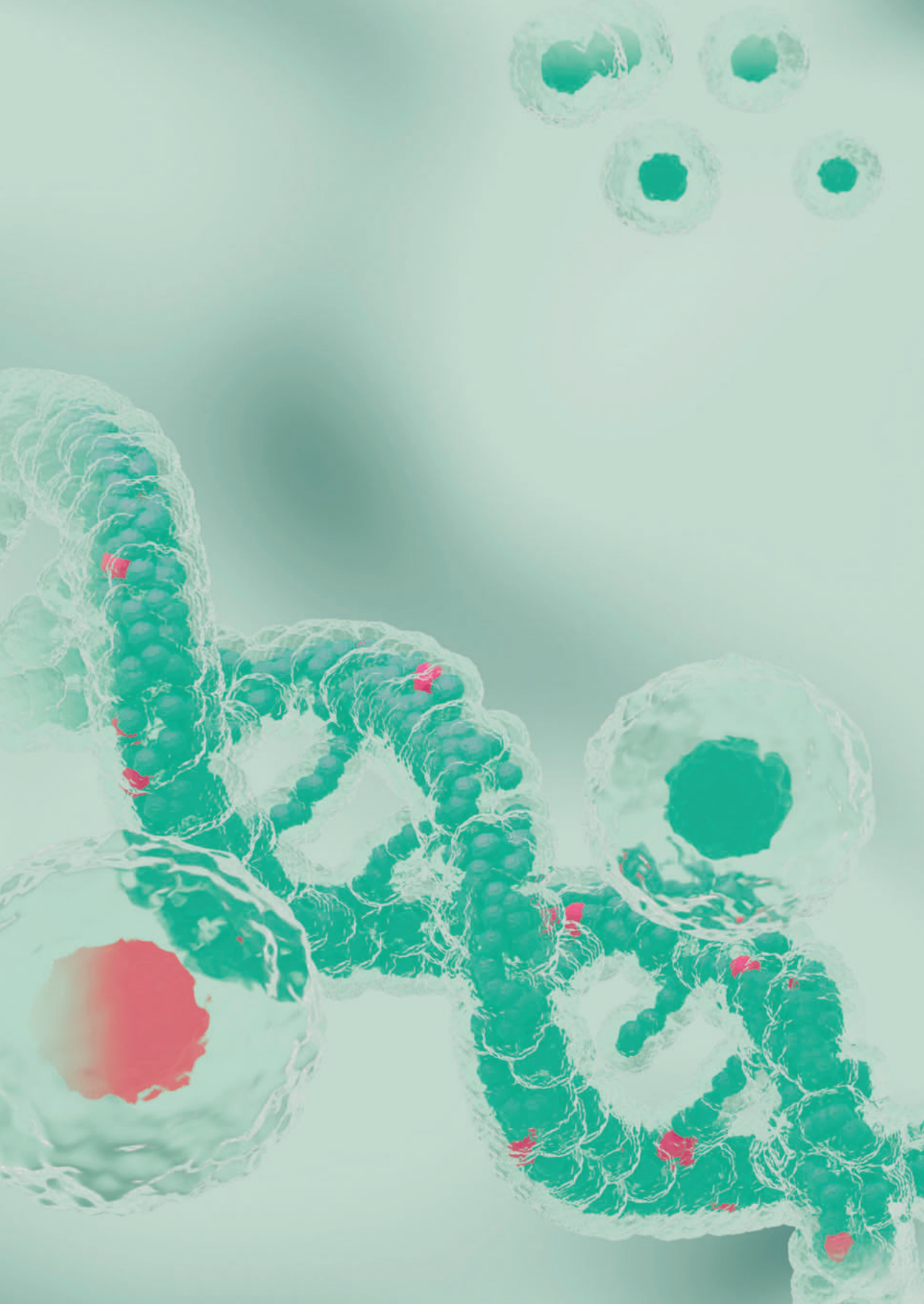
109. Grebely, J., et al., *Time to Detection of Hepatitis C Virus Infection With the Xpert HCV Viral Load Fingerstick Point-of-Care Assay: Facilitating a More Rapid Time to Diagnosis.* J Infect Dis, 2020. **221**(12): p. 2043-2049.

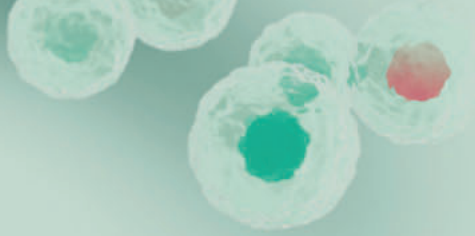
110. *AACC Recommendation for Reporting SARS-CoV-2 Cycle Threshold (CT) Values*. 2021 July 7; Available from: https://www.aacc.org/science-and-research/covid-19-resources/statements-on-covid-19-testing/aacc-recommendation-for-reporting-sars-cov-2-cycle-threshold-ct-values.

111. *IDSA and AMP joint statement on the use of SARS-CoV-2 PCR cycle threshold (Ct) values for clinical decision-making.* 2021 March 12; Available from: https://www.idsociety.org/globalassets/idsa/public-health/covid-19/idsa-amp-statement.pdf.

112. Muller, A.C., et al., *Microchimerism of male origin in a cohort of Danish girls.* Chimerism, 2015. **6**(4): p. 65-71.

113. Yan, Z., et al., *Mate microchimerism in women without sons: Quantitative assessment and correlation with pregnancy history.* American Journal of Medicine, 2005. **118**(8): p. 899-906.





2

CHIMERISM IN HEALTH AND POTENTIAL IMPLICATIONS ON BEHAVIOR: A SYSTEMATIC REVIEW

Published as:

Johnson, B. N., Ehli, E. A., Davies, G. E., & Boomsma, D. I. (2020).

Chimerism in health and potential implications on behavior: A systematic review. *American Journal of Medical Genetics: Part A*, 182(6), 1513-1529.

https://doi.org/10.1002/ajmg.a.61565

ABSTRACT

In this review, we focus on the phenomenon of chimerism and especially microchimerism as one of the currently underexplored explanations for differences in health and behavior. Chimerism is an amalgamation of cells from two or more unique zygotes within a single organism, with microchimerism defined by a minor cell population of <1%. This article first presents an overview of the primary techniques employed to detect and quantify the presence of microchimerism and then reviews empirical studies of chimerism in mammals including primates and humans. In women, male microchimerism, a condition suggested to be the result of fetomaternal exchange *in utero*, is relatively easily detected by polymerase chain reaction molecular techniques targeting Y chromosomal markers. Consequently, studies of chimerism in human diseases have largely focused on diseases with a predilection for females including autoimmune diseases, and female cancers. We detail studies of chimerism in human diseases and also discuss some potential implications in behavior. Understanding the prevalence of chimerism and the associated health outcomes will provide invaluable knowledge of human biology and guide novel approaches for treating diseases.

Keywords: Microchimerism, women, autoimmune diseases, cancer

1. INTRODUCTION

Originating from Greek mythology, the chimera was known as a creature with a lion's head, goat's body and a serpent for its tail [1]. The term chimera is commonly used in many scientific disciplines to describe an entity that is comprised of parts from multiple sources to make a whole. For instance, chimeric fusion proteins are made from joining parts of separate genes to produce a fusion protein with domain regions or the function of the original independent proteins. Similarly, a genetic chimera is an individual organism comprised of cell populations from two or more unique zygotes [2]. Sometimes chimerism and genetic mosaicism are interpreted as describing similar phenomena, however mosaicism refers to unique cell populations that originate from the same zygote. In medicine, the creation of man-made chimeras is the foundation of the lifesaving techniques in transplant and transfusion medicine. Research has explored chimeras in many areas of biology such as genetics, molecular biology, and the generation of model organisms [3]. The studies of spontaneous, natural chimerism are of increasing interest for the consequences that chimerism may have for health, behavior, fertility and disease.

Known cases of chimerism can be predominantly classified in three categories; 1) artificial, i.e. arising after blood transfusion or bone marrow transplantation; 2) tetragametic, i.e. the case of fertilization of two oocytes by two spermatozoa and fusion of the resulting embryos leading to development of one organism; 3) transplacental, as a result of the passage of blood between mother and child, with twin chimerism as a special case. Transplacental microchimerism is bidirectional with potential for mothers to carry their children's cells (fetal microchimerism) and children to carry their mother's cells (maternal microchimerism). Fetal cells have been discovered in maternal peripheral blood samples during pregnancy or postpartum, including mesenchymal stem cells, leukocytes, nucleated erythrocytes, trophoblasts and hematopoietic progenitor cells [4-9]. Studies in several mammalian species have documented that blood exchange during pregnancy can facilitate the transfer of a small number of cells, further discussed in section 3 [5, 10, 11]. These specific cases encompass the majority of identified cases of chimerism and are labeled as "microchimerism", defined as a minor cell population contributing to less than 1% of the total cell population. As there is no lower limit that produces microchimerism in an individual organism, the ability to identify a mixed chimerism within an individual

sample is ultimately dependent on the sensitivity of the techniques available, as presented in Table 1. In the study of chimerism, there are four primary techniques which have been utilized to detect, quantify and identify unique cell populations within individual samples.

Table 1: *Summary of common techniques used in identifying or researching chimerism.*

Lowest detection limit is dependent on specific target and application; examples presented from associated literature. [15, 27, 28, 33, 148-152]

Technique/ Criteria	Primary Principle	Target	Detection limit per assay (%)	Advantages	Disadvantages
Traditional blood typing anti-serum	Agglutination	RBC antigen	>10	Simplicity Availability	Limited to RBCs Sensitivity
Gel column agglutination	Agglutination	RBC antigen	10	Simplicity Availability	Limited to RBCs Sensitivity
Fluorescent microsphere	Fluorescent labeling	RBC antigen	0.01	Sensitivity	Requires fluorescent microscopy Limited to RBCs
Karyotyping	Microscopy	Chromosome	Limited by nuclei observed	Gross genetic differences	TIme consuming Low throughput
Flow cytometry	Fluorescent labeling	Cell surface antigen	Limited by cells observed 0.05	Dynamic Various applications	Complex Availability May require additional samples
Real-time quantitative PCR	PCR amplification	DNA	0.0001	Simplicity Availability Sensitivity	Requires specific target
Short Tandem Repeat (STR) PCR	PCR amplification	DNA	1-5	Simplicity Availability	Sensitivity
Fluorescent *in situ* hybridization (FISH)	Fluorescent labeling	DNA	Limited by nuclei observed	Cell level analysis in tissue sections	Requires fluorescent microscopy or automated system

1.1 Blood grouping

Many cases of chimerism were initially identified inadvertently by discovery of blood type discrepancies during routine testing. Mixed red cell populations were originally discovered as a consequence of *in utero* anastomoses of chorion blood vessels in cattle [12]. This was first described in humans by Dunsford et al. (1953) in a blood donor who was revealed to have both type A and type O red cell populations, presumably as a result of chimerism from a twin [13]. Common blood bank anti-serum agglutination testing is accomplished by introducing reagent antibodies to the individual's red blood cells and visual agglutination to determine if the cells express the corresponding antigen. Due to the limitations of visual interpretation, this is primarily an effective method for testing the antigen composition of major red cell populations [14]. Additionally, gel column agglutination, another standard blood bank technology, has been reported to detect mixed-field agglutination reactions in an individual with a 10% minor cell population [15]. Combining anti-sera identification methodology with a fluorescent assay assisted the largest study of blood group chimerism prevalence in 415 DZ twin pairs and 57 sets of DZ triplets [1]. In short, this technique utilizes IgG antibodies targeting specific antigens and subsequently introducing fluorescent microspheres coated with anti-human IgG to visualize cells possessing the antigen [1]. Paired with fluorescent microscopy, this technique reported a sensitivity of one positive cell in 10,000 total cells. At this sensitivity, the study reported 8% of twins and 21% of triplets with detectable red blood cells chimerism.

More recently, studies examining blood group chimerism have exploited flow cytometry to automate the process of counting fluorescently labeled cells. In one case study, a pair of monochorionic DZ twins initially reported type AB and type B blood types. Flow cytometry testing at three months of age revealed red cell chimerism of 88% type AB with 12% type B and 99% type B with 1% type AB [16]. In another study, two neonates from a set of triplets produced mixed field reactions in gel card agglutination reactions and flow cytometry was used to establish the presence of both a type A and type O cell populations [17]. Flow cytometry has notable versatility and sensitivity over other blood antigen techniques, including the capacity to measure the prevalence of HLA antigens on white blood cells. Fluorescence activated cell sorting (FACS) is built upon flow cytometry and allows the investigator to separate cells based on fluorescent labeling. This has been used

to enrich microchimeric cells in blood samples using HLA specific monoclonal antibodies with minor cell populations as small as 0.01%, allowing for not only detection of these cells but also further study of this minor cell population [18].

1.2 Karyotyping

Karyotyping is a common cytogenetic technique utilized for examination of gross genetic changes in an individual's chromosomes including aneuploidy and structural changes. This technique examines an individual's chromosomes for genetic disorders which has occasionally led to the discovery of chimerism. Karyotyping can identify major differences between the karyotypes of individual cells including the number and type of sex chromosomes. This approach was previously employed for the identification and diagnosis of chimerism in individuals with a mixture of 46,XX/46,XY cells [19-24]. While this is an effective technique, the approach is limited to the identification of gross chimerism with chromosomal differences such as 46,XX/46,XY and is not able to provide information regarding the potential source of chimerism. Further, this approach does not provide evidence of an individual's alleles making it possible for cases of mosaicism caused by nondisjunction events, as previously reported, to potentially be misdiagnosed as chimerism [25].

1.3 Polymerase chain reaction

One of the most widely accepted methods for identifying chimerism via genomic DNA is by polymerase chain reaction (PCR) [26]. The PCR technique is based on the fundamental principles of DNA replication where the strands of double stranded DNA are denatured, primers are annealed to a complementary sequence, and extension is performed by a thermostable Taq polymerase. This process, repeated over multiple cycles, produces exponential growth in copies of the region of interest which are detected by a variety of different techniques including gel electrophoresis or Sybr Green and TaqMan probes for real-time quantitative PCR. The amplification of signal in PCR is a major reason for the popularity of this technique in microchimerism detection. At a relatively low-cost, researchers can identify a small amount of target genome by amplifying the amount of target available for detection, thereby enhancing the detection signal. It is inherently necessary to examine a target gene that is unique to the chimera cell population or has known discordance in polymorphisms between the minor cell population and host cells. This led to common use of Y-chromosome genes such as *TSPY1* (*DYS14*) and the sex

determining region Y (*SRY*) as targets for identifying microchimerism in women. Several studies have reported obtaining a sensitivity of one male cell per one million female cells with this technique [27-29].

Additional studies have focused on targeting variation and polymorphisms in the genome to identify cells that originate from distinct zygotes and to provide insight into the source of these cells. Researchers are able to exploit known differences between two genomes to identify small populations that have a distinctive characteristic, such as the human leukocyte antigen (HLA) system. As such, the use of HLA-specific PCR assays has been implemented to provide context to the source of microchimeric cells by characterizing the HLA profile of family members to identify the original source [30]. In some studies, it was possible to detect non-inherited HLA sequences representing maternal microchimerism, regardless of the sex of the progeny [31, 32]. While this is a seemingly significant advantage over other techniques, the need to obtain samples from both zygotes to establish unique HLA markers for chimerism analysis, often makes it difficult or not possible to apply. Polymorphic short tandem repeats (STR) are well established in forensic investigations as a PCR target for elucidating the potential sources of genomic material and have also been exploited for chimerism analysis, though with typically lower sensitivity than real-time PCR [33]. Cases of chimerism and mosaicism have been reported as a challenge in forensic cases where STR is used to identify the source of the genomic material [21, 34].

1.4 Fluorescent *in situ* hybridization

The techniques described above have primarily focused on the global identification of the presence of chimerism within a sample, consisting of several cells. To study localization of donor cells on an individual cell level among a background of host cells, researchers may utilize the cytogenetics technique fluorescence *in situ* hybridization (FISH). The FISH assay technique uses labeled nucleotide probes which hybridize to a specific genomic sequence and are subsequently able to be detected by presence of a fluorescent signal. FISH has provided notable advantages to studying the biology of chimerism by allowing researchers to identify the localization of non-host cells within tissue samples [35]. Like PCR, a common technique in tissue samples from women is to use a probe for a Y-chromosome specific sequence which is often paired with an X-chromosome specific probe using dual color probe combinations to identify male cells that feature both signals and female cells that feature two X probe signals.

The sensitivity of this technique is dependent on the number of cells examined and can be measured using automated scanning instruments.

The identification of chimeras may lead to new discoveries and implications of chimeric cells on human health and behavior. Considerable progress has been made in testing hypotheses of microchimerism benefits and consequences because of improved techniques with high sensitivity. However, we have yet to establish the definitive role of this phenomenon, identified among several mammalian species. In this systematic review, we seek to examine the current findings in chimera research regarding health, potential for influencing behavior, points of controversy, and discuss questions for future studies.

2. METHOD

2.1 Search strategy

We performed a systematic review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [36]. The literature search was conducted by searching PubMed and ISI/Web of Science from inception to April 2nd, 2019. The search was conducted to include articles with the terms 'chimera', 'chimerism' and related terms. Excluded terms include 'transplant' and 'chimeric' which primarily contain material related to man-made chimeras and chimeric fusion proteins which are outside of the scope of this review. The process for screening and selecting articles for inclusion is provided in Figure 1. The full systematic review strategy can be found in the supplementary material.

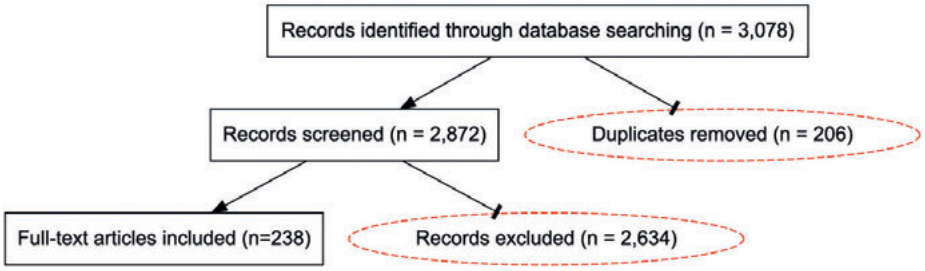


Figure 1: *Flow diagram of article inclusion. At each stage articles included are in solid black rectangles, excluded articles are in dashed red ovals.*

3. ANIMAL STUDIES

Cases of chimerism have been well studied in the animal kingdom and particularly in veterinary science. Several animal taxa have been investigated for the presence of chimerism with a variety of phenotypic outcomes. Among these are cattle, which were the first to be identified as possessing blood cells derived from a co-twin. Monozygotic twinning in cattle is infrequent and similarly, it is not common for full siblings to share blood antigen types across genetically controlled antigens; yet dizygotic twin cattle often share blood type identity [12]. Studies of molecular markers in twin bovine calves has found that dizygotic twinning results in bidirectional cell exchange and subsequent chimerism [37]. One notable consequence of this exchange in cattle is the generation of freemartin heifers, an infertile female calf, which commonly arise in dizygotic opposite-sex twin pregnancies. It has been suggested that freemartinism is the result of anti-Müllerian hormone exchange from the male co-twin due to intrauterine blood transfusion between the twins via anastomoses of the placental vasculature [12, 38, 39]. The occurrence of freemartinism in other mammals with opposite-sex dizygotic twin pregnancy seems to be rare; although further research is warranted as cases of freemartinism with XX/XY chimerism have been described in ovine, caprine, porcine, equine, Cervidae and Camelidae families [40]. A reported case of freemartinism with XX/XY chimerism in Rocky Mountain Bighorn sheep (*Ovis* candensis) demonstrated an association with the development of a masculine appearance and behavior [41]. These observations present an important consequence of chimerism that may have significant contribution to the understanding of biological influence of behavior (see section 4.4).

Primates provide an interesting context to studies in non-human mammals due to their phylogenetic proximity to humans; however, in the context of reproduction particular species have unique characteristics. Studies in rhesus monkeys, which share developmental similarities to humans, demonstrated that male microchimerism in maternal tissues can be measured for several years postpartum [42]. In the study of chimerism, marmosets are of particular interest due to dizygotic twinning being the principal rule of reproduction for the Callithricidae family [43]. Marmoset blastocysts undergo fusion that produces anastomoses of the placental vasculature, yet there are no reports of freemartinism in marmoset females. Studies in marmosets have

discovered opposite-sex cells within the bone marrow and therefore evidence of prolonged chimerism as a result of prenatal blood exchange [43]. Further, DNA fingerprinting of leukocyte rich-tissues in the common marmoset (*Callithrix jacchus*) produces shared profiles between litter mates, whereas leukocyte-poor tissues produce unique profiles supporting the hypothesis of chimerism via intrauterine transfusion [44]. PCR amplification demonstrated that sibling-derived chimerism was present in the majority of marmoset twin sets and the associated chimerism was distributed across various tissue types including gonadal tissue and sperm samples [45]. Subsequent assessment of behavioral characteristics in relation to kin recognition found that marmoset fathers will carry chimeric infants significantly more than nonchimeric infants, while maternal carrying is significantly lower for chimeric infants suggesting that parental care behaviors may be correlated with allele sharing in the offspring [45]. These findings in marmosets further suggest behavioral characteristics as a consequence of chimerism and will require future study in other species to understand if behavioral influences of chimerism are similar across mammals.

Among other families of non-human mammals, there are reports of natural chimerism in mice, dogs, and cats [11, 46, 47]. Multiple studies have concluded that female dogs with prior male birth are able to develop persistent male microchimerism with one suggesting that these male cells could explain male chimerism in daughters of subsequent pregnancies [47, 48]. A case report of a dog with ambiguous genitalia was found to have a leukocyte XX/XY chimerism suggesting the potential for phenotypic abnormalities with canine chimerism [49]. Due to similarities with humans in disease and environment, dogs have been suggested to be useful models for human conditions [48]. Mice are a common model of human disease and have demonstrated pregnancy related chimerism similar to humans [11, 50]. Many studies have investigated the presence of persistent fetal cells in the non-transgenic mother by enhanced green fluorescent protein models for detection [51]. Both maternal and fetal chimerism commonly occur in mice with detection in the tissues of all major organs [11, 52]. Studies of myocardial injury in pregnant mice have provided evidence of fetal cell migration to sites of cardiac injury and capacity for cardiac differentiation [51, 53]. Fetal microchimerism in the lungs and brain of mice feature various cell types and have greater detection at the site of tissue injury from smoking or excitotoxic lesions respectively, suggesting a potential protective effect [54-56]. It has been

proposed that improving our understanding of chimeric cell trafficking across the placenta, blood-brain barrier and migration to sites of tissue injury in mice may provide insights into the properties that also support this behavior in humans [55].

4. HUMAN STUDIES

4.1 Hermaphroditism

Perhaps the most extreme phenotype that is closely associated to human chimerism is true hermaphroditism, a phenomenon of sexual development defined by the simultaneous presentation of both ovarian and testicular tissue. These cases generally present with ambiguous genitalia of variable external severity. There are several reported cases of 46,XX/46,XY karyotypes with true hermaphroditism; however, these only comprise of approximately 10% of total cases of recorded hermaphrodites [25, 57]. Examination of the paternal and maternal genetic contributions have provided a means to establish the origin of these cases; while mostly chimeric, there have been reported cases of mosaicism originating from 47,XXY cells as a result of nondisjunction events [25]. Several hypotheses have been presented to understand the mechanism that could produce 46,XX/46,XY whole-body chimerism. As this is a developmental anomaly these hypotheses have been focused on zygote fusion before sex differentiation; including separately fertilized ova, fertilization of an ovum and second polar body, or a parthenogenetic division of an ovum fertilized by two sperm [22, 23, 58]. Due to the overall rarity and variable nature of this condition it is possible for many cases of hermaphroditism to go undiagnosed or present later in life and therefore the majority of reported cases are young patients with more dramatic phenotypes.

4.2 Autoimmune diseases

Among the studies of microchimerism in human disease, autoimmune diseases have been a primary focus. In the human population approximately 5% are affected by autoimmune diseases of which there is a prominent gender bias with an estimated 78% of affected individuals being females [59-61]. Furthermore, several autoimmune diseases are more predominant in women following the childbearing years and share a similar pathology to chronic graft-versus-host disease (cGVHD), a well-studied condition in transplantation medicine as a result of chimerism [62, 63]. Cumulatively,

these characteristics suggest a role for microchimerism in autoimmune disease etiology. The contributions of microchimerism in autoimmune disease etiology has been largely explored for rheumatic diseases and autoimmune thyroid diseases (AITD).

Rheumatic Autoimmune Diseases

Rheumatic diseases have been well documented with relation to chimerism. Several studies have examined women diagnosed with systemic sclerosis and have produced findings that indicate a greater prevalence and quantification of male microchimerism in peripheral blood and various tissue samples compared to controls [64-69]. A similar study of women with systemic lupus erythematosus (SLE) or rheumatoid arthritis (RA) found greater prevalence of male microchimerism in the peripheral blood compared to controls [70]. HLA associations have been implicated as major risk factors in autoimmune diseases and have therefore been studied extensively with regard to microchimerism. Approximately 80% of patients with RA present with a five-amino-acid motif encoded by *HLA-DRB1* alleles termed the shared epitope [71]. Microchimerism may contribute to RA disease onset as higher levels of minor cell populations with *HLA-DRB1*04* and HLA-*DRB1*01* in addition to the shared epitope have been identified in women with RA [71, 72]. These findings indicate disease-associated alleles acquired via microchimerism may contribute to host disease etiology. The implications of HLA associations extends to systemic sclerosis where detectable fetal microchimeric T lymphocytes are associated with a maternal *HLA-DQA1*0501* allele, suggested to facilitate persistent microchimerism [73]. Together, these data suggest that the molecular relationship between the host and acquired cells aid in facilitating microchimerism and subsequent autoimmune response; however, more studies are needed to elucidate these interactions and associated immunology.

Autoimmune diseases typically result in tissue destruction and similarity to cGVHD, which has led to studies that have investigated microchimerism within host tissues. SLE is characterized by production of antinuclear antibodies that have specificity for nuclear components of the cell resulting in extensive organ complications and similarity to cGVHD [74]. A case study of a woman who died of SLE complications and previously had given birth to two sons presented with detectable male cells in all of her histologically abnormal tissues [74]. In women that present with Sjögren's syndrome, male cells have been identified in labial salivary glands and

bronchoalveolar lavage fluid (36% and 22%, respectively) while none were found in the peripheral blood or in control patients [75]. This localization of non-host cells to regions of inflammation while simultaneously remaining undetectable in the peripheral blood further illustrates the limitations of studies that exclusively examined the peripheral blood for microchimerism [76].

Autoimmune Thyroid Diseases

Research of microchimerism has further explored the conditions that are categorized as AITD, including Hashimoto's thyroiditis and Graves' disease. This class of diseases are known for being suppressed during pregnancy and subsequently exacerbated postpartum, suggesting that changes incurred during pregnancy facilitate the pathogenesis of these diseases. Studies of thyroid tissues have found detectable male microchimerism to occur more frequently in women with AITD when compared to control women with nodular goiters or follicular adenomas, suggesting a potential role in Hashimoto's thyroiditis and Graves' disease etiology [77, 78]. Interestingly, flow cytometry of samples from women with AITD demonstrated that fetal cells in Hashimoto's thyroiditis are primarily $CD8^+$ T cells, whereas in Grave's disease the majority of fetal cells are B cells [60]. It can be hypothesized that T cell mediated cytotoxicity could be directly involved in Hashimoto's thyroiditis, whereas onset of Grave's disease may be dependent on activation via host $CD4^+$ T cells [60]. Additionally, HLA typing of the mothers and offspring found that women with detectable male microchimerism more frequently have alleles HLA-DQA1*0501, DQB1*0201 or DQB1*0301 which are known to be alleles associated with susceptibility for AITD [78].

Autoimmune Diseases of the Liver

Primary biliary cirrhosis (PBC) is a disease of the liver that has autoimmune origin, with clinical and histological similarity to cGVHD. Previous studies of male microchimerism in total female PBC patients (18%) and PBC patients with a male child (70%) found that the prevalence was similar in liver specimens compared to control women with other liver diseases (5% and 72%, respectively) [79, 80]. There are significant differences in male microchimerism prevalence between these studies; it is worth noting that the increased percentage of total female PBC patients with male microchimerism compared to other liver disease patients in the study by Schoniger-Hekele et al. (2002) may be due to a larger proportion of women with male

children in the PBC group that was unaccounted for. Another study examining male microchimerism of presumed fetal origin found similar prevalence in the peripheral blood of PBC (36%) and healthy control (31%) women [81]. Similarly, the potential role of maternal microchimerism in the etiology of PBC has also been explored with similar report of insignificant findings for any association with PBC [82]. Together these studies suggest that the etiology of PBC and other liver diseases are not likely to be influenced by microchimerism.

Juvenile Autoimmune Diseases

Similar to adults, rheumatic diseases in youth are a primary focus of research in maternal microchimerism associated diseases. A group of conditions identified as juvenile idiopathic inflammatory myopathies (JIIM) collectively share similar pathology to cGVHD, suggesting a potentially similar mechanism for disease pathogenesis. Juvenile dermatomyositis (JDM) is the most common JIIM condition and presents with lymphocytic perivascular infiltrates, similar to cGVHD, in the muscle and skin lesions. Individuals with JDM have demonstrated greater prevalence of maternal microchimerism in the peripheral-blood and muscle tissue compared to healthy controls [83, 84]. Likewise, a study of microchimeric HLA-Cw alleles discovered chimerism in 73% of JIIM patients and only 10% in healthy controls [85]. The presence of maternal microchimerism in children with JDM has been associated with the *HLA-DQA1*0501* allele in the mother [83, 86]. These data present a notable similarity to rheumatic autoimmune diseases in adults and further support the hypothesis of an additional route by which HLA genes might contribute to microchimerism and disease susceptibility, as the *HLA-DQA1*0501* allele is strongly associated with both JIIM conditions across multiple ethnicities and the presence of microchimerism [73, 87].

Type 1 diabetes (T1D) presents in children and young adults, leading to the hypothesis of a pathogenic mechanism of autoimmunity caused by maternal microchimerism. The peripheral blood and pancreas tissue in patients with T1D have increased maternal microchimerism compared to controls [31, 88]. Further examination of maternal microchimeric cells in the pancreas has found this to be a common phenomenon among both normal and T1D pancreases; maternal cells of the endocrine, exocrine and vascular endothelial lineages suggest these were derived from a maternal multi/pluripotent progenitor cell [31, 88]. Maternal cells are notably enriched in beta cells in the pancreases of T1D patients however the role of these

cells in immune balance during development remains uncertain [31, 88]. It is not clear what the role of maternal cells in the pancreatic tissues is, as the timing of cell grafting cannot be derived from retrospective analysis. Therefore, these cells may have been present and involved during disease onset or, alternatively, may have migrated to the tissue after in response to active autoimmunity which requires further study.

Contradictory Findings in Autoimmune Diseases

Despite a large body of knowledge that supports a pathogenic hypothesis for microchimerism in autoimmune diseases, there are several studies that have produced contradictory findings. Among these are findings of no difference between cases and controls for the presence of male cells in women with systemic sclerosis, indicating that microchimerism may not be involved with systemic sclerosis pathogenesis [89, 90]. These studies have relied on retrospective case-control studies of disease, examining affected populations for detectable microchimerism, assuming that controls will not later develop systemic sclerosis and that no detectable microchimerism at the time of sampling is indicative of negative lifetime microchimerism. Low sample size may have contributed to insignificant findings of circulating male cells and result in inconsistent conclusions.

Studies among groups of women with thyroid autoimmunity have also been plagued by mixed conclusions. A population study of women in diverse age groups by Bülow Pedersen et al. (n=3,712) found no association between thyroid autoimmunity as indicated by thyroid autoantibodies and previous pregnancy or parity [91]. However, this is inconsistent with findings by Greer et al. who reported in a study of 17,298 prenatal patients at 20 weeks gestation that thyroid peroxidase antibody levels increased with increasing parity [92]. Another report found greater prevalence of male microchimeric cells in healthy control women compared to those with Grave's disease or Hashimoto's thyroiditis (P=0.0004 and P=0.001 respectively) although male cells were identified in the blood vessels and forming follicles in the thyroid of women with AITD [93]. The findings of this study emphasize the need for studies that investigate affected tissue samples for microchimerism, as this will provide insight to the cellular environment of the disease in question. Similarly, some studies of fetal and maternal microchimerism in patients with SLE have reported no significant differences in prevalence compared to healthy controls [94, 95]. These contradictory

findings leave many of the scientific inquiries regarding microchimerism and autoimmune diseases unanswered. However, if these contradictory findings are supported by future studies of affected tissues and sufficient sample sizes, it should be considered to support the null hypothesis that microchimerism is a bystander and does not actively contribute to the onset of autoimmune disease.

4.3 Cancer

Reports of association between parity and reduced cancer susceptibility has led to the study of microchimerism in the course of cancer diseases [96]. Similar to autoimmune diseases, papillary thyroid cancer (PTC) and breast cancer are more prevalent in females; therefore, fetal microchimerism may contribute to the etiology of these diseases. One study of PTC neoplastic thyroid tissue revealed male microchimerism in the tissue of 47.5% of women with previous male pregnancy compared to no detection in female controls with PTC and no previous male pregnancy [97]. Subsequent studies have discovered male cells in neoplastic thyroid tissue, despite lower prevalence of male microchimerism in the peripheral blood of women with PTC compared to healthy controls [27, 98]. Individuals with PTC and peripheral blood male microchimerism have lower prevalence of extra-thyroidal extension or lymph node metastasis and better overall outcome [98]. It has been hypothesized that these findings could be the result of fetal cells protecting the host via a cytotoxic role against preneoplastic cells and as sentinel cells against malignant cells, preventing new tumors and tumor progression respectively [99]. Similar study of breast cancer cases and controls has found male microchimerism in 26% and 56% respectively while a subsequent case-control study detailed that male microchimerism is less prevalent in the unaffected breast tissue of women with breast cancer [96, 100]. Contradictory to the findings in autoimmune diseases, these findings may indicate potential beneficial effects and collectively support the hypothesis that microchimeric cells may be involved in immune surveillance, migrate to damaged or diseased tissue and be recruited for tissue repair. However, further research is necessary to establish the facilitating mechanisms; the discovery of these cells in damaged or diseased tissue may yet constitute null or malevolent effects.

Expanding on these studies, a prospective study of 428 Danish women discovered lower baseline prevalence of male microchimerism in peripheral blood of subjects who later developed breast cancer (40%) compared to controls (70%), whereas women

who later developed colon cancer had an overall greater baseline prevalence of male microchimerism (95%) [101]. Survival rates of individuals from the study following diagnosis of either cancer type were improved in individuals with detectable male microchimerism [102]. The resulting hypothesis insinuates that microchimerism may contribute to each cancer differently while immunological and repair roles of chimeric cells may contribute to improved survival among women regardless of cancer type. These seemingly contradictory findings were summarized in a hypothesis by Geck (2013) that this is an observation of evolution in action where the positive and negative effects regarding the presence of chimerism via fetomaternal exchange are being weighed in an evolutionary experiment [103]. Additional research is needed to specifically examine the mechanism of action for fetal cells to have a protective or pathogenic role in women with cancer.

A variety of other cancers including uterine, melanoma, glioblastoma and meningioma have found measurable microchimerism in tumor samples via FISH or PCR techniques [104-106]. Male microchimerism, although common in endometrial tissues, was found less frequently in the tissues of type 1 endometrial cancer compared to controls [105]. Lower male microchimerism in women with uterine cancer pointed towards better prognoses for factors including histological grade, tumor subtype and stage [105]. Melanoma tissues in both humans and mouse models demonstrated that fetal cells selectively migrate to melanoma tissues based on increased prevalence of chimeric cells compared to nevi (birthmark) sections [106]. Among brain tumors, glioblastoma cases were found to have greater prevalence of male microchimerism compared to women with meningioma [104]. Furthermore, women with glioblastoma and microchimerism had longer average survival time suggesting that the greater prevalence of microchimerism among these patients may have a positive effect on disease outcome [104]. The presence of microchimerism in brain tissues introduces several questions including the influence of chimeric cells on brain function and cognition that requires further investigation.

4.4 Behavior

Studies of chimerism have largely investigated the role of chimerism in somatic diseases, however associations between physical biology and behavioral traits reasonably suggest a hypothesis for chimerism in individual and social behavior. The presence of chimeric cells in autoimmunity presents the potential for changes

in hormone levels and associated behavior due to subsequent disease states. Particularly, diseases of the thyroid have had significant reports of changes in mood and behavior. Current studies of human microchimerism have explored the important potential of involvement in the course of autoimmune thyroid disease (AITD), however known changes in the endocrine system due to such conditions have been previously established. Hypothyroidism found in Hashimoto's thyroiditis has been known to be associated with deficits in cognitive abilities as well as with depression, while hyperthyroidism in Graves' disease is related to depression in addition to anxiety and irritability [107]. The relationship between the endocrine system and behavior is complex with bidirectional relationships and studies of a role of chimerism in disorders of the endocrine system are necessary [108]. Knowledge of an association between immune dysregulation and the etiology of psychiatric disorders has increased interest in the role of an altered immune system in the pathogenesis and treatment of depression and other major psychiatric disorders [109, 110]. We hypothesize that the currently detailed associations between microchimerism and autoimmune disease (Section 4.2) suggest this pathway may be independent of the endocrine pathway.

Sports and exercise behavior, especially in athletic competition, are largely segregated into male and female domains due to the physiological advantages of males in strength and endurance. Following puberty, males exhibit 15-fold greater circulating testosterone than females which leads to increased interaction with androgen receptors leading to biological enhancement of muscle, bone and hemoglobin [111]. To maintain balanced competition in athletics at the elite level, female competition is protected by rules of eligibility. Recently, sex differentiation has come under significant scrutiny in athletic competition based on the masculine appearance and elevated testosterone of two female athletes in advance of the 2016 Summer Olympics [112]. Disorders of sex development, including phenotypically female individuals with 46,XY karyotype and androgen insensitivity complicate the established rules that define eligibility for female competition and individuals who present with XX/XY chimerism provide additional challenges for defining males and females based on genetic testing.

Testosterone is involved in social behavior, with bidirectional relationships and Y-chromosomal chimerism potentially having a role in multiple behaviors that are influenced by testosterone: including behaviors associated with survival,

reproduction, and dominance [113]. Developmental studies of psychosocial and psychosexual outcomes have found associations between androgens and generally male behavior in females [114]. It should be considered that male chimerism may be involved in producing changes in sex hormone production and function. It is equally possible that localized chimerism resulting in testosterone production may be sufficient for producing observable traits in females, but further studies are required to understand the scale of chimerism necessary to produce a sufficient change in phenotype. Current studies of disorders of sexual development, including hermaphroditism, have found that women born with atypical genitalia were more comparable to reference men than reference women for psychosocial behaviors, indicating more masculine personality traits. The relationship of 46,XX/46,XY chimera true hermaphrodites and resulting hormone profiles requires further study [115]. Another example involves an individual with a 46,XX/46,XY karyotype who was brought up female, with an ovotestis and prominent phallus in addition to a functional urethral opening and rudimentary vaginal orifice. The person had undergone corrective surgery removing the ovotestis and had a typical hormone profile and feminine gender identity [19]. These studies suggest that the presence of allogeneic cells of different sex may influence the endocrine function of the gonads, modifying hormone signal profile and subsequent behaviors. Despite known associations of microchimerism in thyroid diseases and T1D there is a need for future studies to describe the behavioral implications.

As suggested in marmosets, the presence of chimerism can shape social behaviors and manipulate involvement or interaction with offspring [45]. Maternal microchimerism may indeed be a biological process involved in imprinting oneself on their offspring and, perhaps unintentionally, leading to greater investment and evolutionary implications on survival. Further, increasing maternal exposure to fetal cells during the course of pregnancy may ultimately be involved in actively encouraging processes that support postnatal care of the child, from maintaining maternal health to increased lactation [116]. Expanding on previous hypotheses presented by Boddy et al. (2015) of microchimerism in an evolutionary framework, maternal and fetal microchimerism may have been selected for during mammalian evolution to have distinct roles, including psychological manipulation, which encourage maternal interaction with their child to facilitate survival. Future

longitudinal studies may seek to investigate the role of chimerism in behavior in addition to psychological health.

4.5 Female-Specific Diseases

Microchimerism is increasingly inspiring research on other conditions that exhibit a female propensity or clinical similarity to cGVHD. One such group of conditions are pregnancy related complications. There is an association between experiencing fetal loss and subsequently having detectable microchimerism [117]. It is therefore important to recognize that parity, the number of pregnancies carried to viable gestational age, is potentially less preferable for microchimerism studies than gravidity which accounts for all pregnancies. It should be noted that research by Gammill et al. established that fetal microchimerism concentration and prevalence does not increase with parity suggesting the possibility of dynamic graft-graft interaction [118].

Preeclampsia presents a complex and dynamic relationship with microchimerism as the condition appears to be associated with a specific chimerism source. Fetal microchimerism is seemingly pathogenic and is significantly more prevalent in women with preeclampsia, whereas a protective role is suggested for maternal microchimerism which has been found in lower frequency among women with preeclampsia as well as recurrent miscarriage [119-121]. First time preeclampsia onset is more common among women with a change in paternity between pregnancies; conversely, a reduction in preeclampsia prevalence is demonstrated for women with preeclampsia in the first pregnancy [122]. Exposure to non-shared paternal alleles is hypothesized to be implicated in preeclampsia etiology, potentially exacerbated by fetal microchimerism. Therefore, a change in paternity in women with previous preeclampsia may eliminate exposure to an offending antigen. It is possible that these immunological interactions are exacerbated by fetal microchimerism and further research is warranted to investigate if HLA sharing between maternal and paternal genomes may contribute to the overall risk for preeclampsia.

4.6 Other Human Studies

Male microchimerism is common in the female liver and present in fetal, juvenile, and adult liver tissues suggesting a possible role of both maternal and fetal microchimerism in liver diseases [123]. In addition to studies in PBC (see Section

4.2), biliary atresia is a neonatal liver disease with maternal chimerism has been observed in the liver tissue of affected infants. Study of these cells found that they include $CD8^+$, $CD45^+$ and cytokeratin positive and are therefore may originate as stem cells capable of differentiation into progenitor lymphocytes, effector lymphocytes, or biliary epithelial cells [124]. Additionally, study of the bile duct epithelium and hepatocytes of female biliary atresia patients has demonstrated the presence of antimaternal HLA class I antibodies [125]. These findings suggest an immunological role of chimerism in the pathogenesis of biliary atresia; yet distinguishing GvHD or host-vs.-graft disease requires additional study of the mechanisms of disease pathogenesis [124, 126].

Similar to studies in mice (see Section 3), microchimerism of heart tissue had also been seen in humans. The study by Bayes-Genis et al. (2005) investigated the cardiac tissue of two women with male offspring, identifying the presence of male cells via real-time PCR and FISH [127]. The male cells which were identified via FISH also demonstrated cardiomyocyte phenotype and protein expression. While this report did not specifically examine a role of microchimerism in cardiac conditions it presents important findings regarding the integration of presumably fetal cells into maternal tissues. This and other studies have demonstrated the plasticity of chimerism and capacity to differentiate and integrate in a variety of tissue types. It will be important for future studies to investigate this plasticity further and understand the tissue-specific implications of chimerism.

4.7 Prevalence

Perhaps the most elusive information in the study of human chimerism is its overall prevalence. Limitations of studying chimerism on a population scale include several of the limitations presented by the current techniques. Limitations of sensitivity and inability to obtain a global sample of chimerism throughout the individual in all tissues make definitive diagnosis of chimerism complicated. Karyotyping, FISH and PCR approaches based on Y-chromosome detection as an indicator of male microchimerism, limit identification to chimerism originating from a male donor. Remarkably, studies of women without sons have identified detectable male microchimerism. One study that explored male microchimerism in a population of nulliparous 10-15 year old Danish girls found 13.6% to have male microchimerism in their peripheral blood, supporting the findings of 13% prevalence previously

found in healthy nullgravid women [128, 129]. We do not know if male chimerism originated from an older brother with fraternal DNA passed through the mother, male miscarriages, vanishing twins, transfusion history or sexual intercourse [29, 128-132]. Research by Kamper-Jørgensen et al. in a comprehensive study of numerous reproductive traits, health and lifestyle did not establish a model-based prediction of male microchimerism and concluded there is little known about the relationship between exposure to allogeneic cells and maintaining persistent chimerism; although, histocompatibility with transgenerational HLA relationships has been suggested and should be further investigated [133].

Chimerism has also presented in dizygotic twins. The single largest study of chimerism in twins examined peripheral blood-chimerism in 472 individuals and observed chimerism in 8% and 21% of twin and triplet pairs, respectively [1]. Another twin study found that both male and female opposite-sex twins had an increased frequency of thyroid autoantibodies compared to monozygotic twins, supporting the hypothesis that twins may have increased prevalence and risks from microchimerism [134]. Observed cases of monochorionic dizygotic twinning have facilitated the hypothesis that chimerism between twins can be the result of twin-to-twin transfusion between dizygotic twins *in utero* [135-140]. The majority of documented cases of monochorionic dizygotic twin pregnancies have been the result of assisted reproductive technology and a review of monochorionic dizygotic twins found similar risk for perinatal morbidity and mortality in addition to pregnancy loss before 24 weeks when compared to other monochorionic twin pregnancies [141]. It has been suggested that intrauterine blood exchange may be the cause of genital abnormalities observed in 15.4% of monochorionic dizygotic twin cases. In a recent study of Mayer-Rokitansky-Küster-Hauser syndrome, a phenotypically analogous condition to freemartinism in animals, we unexpectedly found that male microchimerism in the peripheral blood was less frequent in adult patients with this condition compared to healthy controls [29]. An important consideration of the findings in monochorionic dizygotic twin pregnancies is that many dizygotic twin pregnancies have been assumed to be dichorionic and therefore scientists may be underestimating the overall prevalence of this phenomenon. More detailed assessment of chronicity in dizygotic twin pregnancies in future studies will be necessary to provide more comprehensive insight into associations with health complications.

5. FUTURE RESEARCH

Despite improvements in molecular techniques, achieving a sensitivity of one in one million genome equivalents, studies continue to be limited when attempting to identify the presence or absence of single chimeric cells on a scale as grand as a human organism [29]. It is possible that several studies have an inadequate sample size to properly identify an appropriate number of individuals with microchimerism due to the limited sensitivity of modern molecular and cytological techniques, as well as insufficient knowledge of the minimum level of chimerism required to produce pathology if these pathogenic hypotheses are correct. A better understanding of the individual interactions of donor cells with the host environment is needed to establish the role of single allogeneic cells in pathology. As chimerism may be transient and therefore become undetectable following pregnancy, it cannot be excluded that chimeric cells contribute to pathology after they are no longer detected. Current approaches for assessing the implications of chimerism are dependent on the sensitivity of modern techniques, such that only measurable microchimerism can be directly assessed for association with disease etiology. The modern techniques described in current literature are unable to provide a noninvasive approach to comprehensively assess global chimerism status in humans. However, given these limitations it must be considered that alternative approaches may be needed to compare non-shared genomic traits between a mother and offspring which may provide a better understanding of underlying mechanisms of disease pathogenesis than direct measurement of microchimerism. Moreover, we hypothesize that study of non-shared (paternal) alleles between the mother and offspring, using readily available genotyping platforms, may be useful to identify non-inherited genes that have a previously undefined role in fetal or maternal disease.

The majority of studies have obtained samples for analysis of chimerism status after disease onset at which point chimerism level may have decreased to undetectable levels. As a consequence of such retrospective case-control study designs, investigators cannot firmly establish a causal role for transient microchimerism in human health. Further, preferential selection or elimination of allogeneic cells may be distorting the associations made in current studies without examining longitudinal chimerism burden. It will be critical that future studies investigate the role of both transient and persistent chimerism as a factor in both diseases listed in this review

and those to be later established. To address this missing component in the current body of knowledge, prospective cohort studies examining chimerism status in women and offspring from the time of fetomaternal exposure to the time of disease onset are needed. These studies will provide opportunity to understand the biology that facilitates persistent chimerism and association with human health. However, such prospective study designs may be difficult and expensive to implement. Twin studies have provided significant insight into various omics studies, controlling for confounders that complicate other research designs [142]. The disease-discordant monozygotic (MZ) twin design is of particular interest as it controls for age, sex and genetic contributors, allowing for investigation of environmental contribution [142-144]. The discordant MZ twin design may be able to help establish both patterns of allele sharing and associated risk of persistent microchimerism in disease onset.

There may be different forms of each disease that have a unique etiology, such that one form includes a mechanism of pathogenesis activated by chimeric cells and another does not. A study by Rak et al. (2009) compared the presence of male microchimerism in blood samples from females with two different clinical presentations of systemic sclerosis to come to the conclusion that male microchimerism in whole blood occurs more frequently in limited cutaneous systemic sclerosis than diffuse cutaneous systemic sclerosis; however, the opposite was found in peripheral blood mononuclear cells despite no notable difference from controls [65]. Subsequent study of the HLA compatibility between the maternal and fetal cells for HLA-DRB1 found having an HLA-DRB1 compatible child was significantly associated with limited cutaneous systemic sclerosis, but not diffuse cutaneous systemic sclerosis [65]. These findings suggest that different clinical presentations of systemic sclerosis may be the result of different mechanisms of onset which may also translate to other conditions associated with chimerism. Despite evidence indicating association between HLA-DQA1*0501 and microchimerism in autoimmune diseases (see section 4.2), others did not find association between HLA-DQA1*0501 and the presence of microchimerism in patients with systemic sclerosis or JIIM; proposing that discrepancies may be due to methodology, population, study size or confounding association of HLA-DQA1*0501 and autoimmune diseases [145]. To control for several of these confounders, future studies should consider utilizing a discordant twin design, as described above; this may be able to provide further clarity to these ambiguous conclusions.

Currently chimerism has been most widely explored in women with XX/XY male chimerism due to the simplicity and sensitivity achieved using PCR for targeting the Y-chromosome, leaving chimerism in males understudied. Use of HLA typing via PCR and flow cytometry has demonstrated excellent sensitivity and capacity for study of chimerism, not limited by sex chromosomes. Similarly, blood group chimerism has been shown to be effective for study of chimerism that is not reliant on sex chromosomes [146]. Published results have also used FISH to identify and quantify the sex chromosomes of cells in tissue sections providing an opportunity to study maternal microchimerism in male subjects [31, 84, 130]. Researchers must address current challenges in studying mixed chimerism samples by developing techniques with adequate sensitivity and versatility to provide confidence in defining cell sources. A relatively new technology that has shown promise for application in the field of microchimerism is droplet-digital PCR (ddPCR), which one group demonstrated an assay limit of detection and limit of quantification of 0.008% and 0.023% respectively [147]. Future studies will require comprehensive assessment of subject samples within pedigrees to define the source of microchimerism within the proband. Approaches may rely on panels of informative polymorphisms for PCR or targeted sequencing to identify similarities with potential cell sources. Improvements in the sensitivity, specificity and costs of these techniques may provide the tools necessary to better understand chimerism in the general population. Further, larger studies that examine different types of chimerism among various tissue sources will be necessary to elucidate the overall prevalence of human chimerism and better understand the biological relationships facilitating health and disease. Obviously, all techniques and study designs face questions about the appropriate tissue to use for analysis and the comparability of findings across different tissues.

Significant progress has been made in advancing the current understanding of chimerism and the implications hypothesized for human health and wellness. However, two significant questions remain which warrant future study. First, what are the mechanisms by which chimerism produces disease and tissue specific protection or pathogenic consequences? These studies will most likely include experiments of chimerism in animal models for longitudinal study of chimerism in disease progression or tissue repair. To truly understand the role of chimerism, it will be necessary to study chimerism at the tissue level instead of exclusively via noninvasive approaches, such as blood samples. The primary focus of most research

to date has been to establish association with disease. While these studies have produced exciting findings, they have not defined the mechanisms that drive disease pathogenesis. As chimerism has been hypothesized to have beneficial, pathogenic or null effects on health and disease, hypotheses of a role in disease etiology warrants further validation of the proposed biological mechanisms. Second, what is the proportion of the population exposed to allogeneic cells via microchimerism? It will be exceedingly important to establish the prevalence of lifetime chimerism in the human population to understand the positive and negative effects of chimerism in human biology and possibly psychology. These studies will require large population studies to examine the epidemiology of chimerism and systematic collection of samples at the time of chimerism incidence, for example, during pregnancy. Furthermore, to facilitate this aim it will be necessary to develop high throughput approaches to identifying microchimerism in mixed chimerism samples. Prospective studies are desperately needed in the study of chimerism and will subsequently illustrate biological and behavioral changes in subjects longitudinally. By addressing these questions future research will detail the diverse implications and mechanisms underlying chimerism in human biology. This knowledge will facilitate new insight into complex diseases and behavior, potentially inspiring novel approaches to proactive screening, diagnosis and treatment including novel strategies and therapeutics. Similarly, as chimerism is the direct result of medical transplantation and transfusion medicine, it is of increasing importance that we have a comprehensive understanding of the implications of natural chimerism to better prevent long-term consequences, including graft rejection. This research will be able to provide additional context to our understanding of alloimmunization, immune tolerance, the need for immunosuppressants, guide clinical decisions and inform subsequent diagnostic screening. For example, understanding the mechanisms that produce protective and pathogenic consequences of chimerism will provide invaluable information to promote better outcomes in patients receiving allogeneic transplants.

6. CONCLUSIONS

Chimerism in humans has been documented for over 60 years; however, we have only recently begun to elucidate the sources, prevalence, and relationship

of chimerism in human behavior and disease. The continuous improvements in sensitivity and specificity of techniques for chimerism detection has provided researchers with the means of detecting minute levels of allogeneic cells and explore hypotheses relating chimerism to various human conditions. Also, with increases in assisted reproductive technology more dizygotic twin pregnancies are seen; creating a potential increase in chimerism prevalence and stressing another need for understanding the implications of chimerism. The study of chimerism in human behavior and health appears to have a significant amount of complexity that we have only just begun to explain. Current findings suggest that chimerism may have a pathogenic role in autoimmune diseases through mechanisms similar to chronic graft-versus-host disease (cGVHD). Similarities in the pathophysiology of autoimmune diseases and cGVHD observed in transplantation patients provides further evidence to support a pathogenic consequence for sequestering allogeneic cells. Understanding the potential roles of allogeneic cells in the health of a host organism has direct implications in defining the long-term influences on the health of recipients of allogeneic transplants and transfusions. In contrast, among cancer patients, chimerism has primarily been found to be less prevalent in the peripheral blood compared to controls. In addition, chimeric cells appear to migrate to tumor tissues and result in better overall outcomes suggesting a protective role. However, there is immense need for understanding the underlying mechanisms that may be driving contribution to disease etiology. As such, we cannot exclude that allogeneic cells sequestered in disease tissues may also have no direct role in disease prevention or pathology. Studies of chimerism in PBC has shown no significant association, yet biliary atresia has demonstrated a cGVHD role in pathogenesis. Due to the variety of findings that support pathogenic, protective or neutral roles for various human diseases, it is also possible for chimerism to have different roles and mechanisms depending on the associated host conditions. Future studies will be necessary to explain the intricacies that underlie these various conditions and describe the prevalence of microchimerism among human populations to establish additional associated phenotypes.

Elucidating the prevalence of chimerism, both in males and females, and associations of chimerism with health and behavior may facilitate better care and quality of life in individuals with disorders characterized by significant sex differences in prevalence

or unknown etiology. Ultimately, these studies will provide information to improve our understanding of human health and behavior.

ACKNOWLEDGEMENTS:

The authors would like to acknowledge the KNAW Academy Professor Award (PAH/6635) to D.I.B.

AUTHOR CONTRIBUTION STATEMENT:

Systematic review design, literature search, article screening, and drafting of manuscript performed by B.N.J. Supervision of review design and manuscript draft performed by D.I.B., E.A.E. and G.E.D. All authors provided critical feedback, read and approved the final manuscript prior to submission.

CONFLICT OF INTEREST:

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT:

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1. van Dijk, B.A., D.I. Boomsma, and A.J. de Man, *Blood group chimerism in human multiple births is not rare.* Am J Med Genet, 1996. **61**(3): p. 264-8.
2. Race, R.R. and R. Sanger, *Blood groups in man.* 6th ed. 1975, Philadelphia: Oxford : Blackwell Scientific Publications.
3. Eckardt, S., K.J. McLaughlin, and H. Willenbring, *Mouse chimeras as a system to investigate development, cell and tissue function, disease mechanisms and organ regeneration.* Cell Cycle, 2011. **10**(13): p. 2091-9.
4. Bianchi, D.W., et al., *Isolation of fetal DNA from nucleated erythrocytes in maternal blood.* Proc Natl Acad Sci U S A, 1990. **87**(9): p. 3279-83.
5. Bianchi, D.W., et al., *Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum.* Proc Natl Acad Sci U S A, 1996. **93**(2): p. 705-8.
6. Evans, P.C., et al., *Long-term fetal microchimerism in peripheral blood mononuclear cell subsets in healthy women and women with scleroderma.* Blood, 1999. **93**(6): p. 2033-7.
7. Mueller, U.W., et al., *Isolation of fetal trophoblast cells from peripheral blood of pregnant women.* Lancet, 1990. **336**(8709): p. 197-200.
8. Ando, T. and T.F. Davies, *Self-recognition and the role of fetal microchimerism.* Best Pract Res Clin Endocrinol Metab, 2004. **18**(2): p. 197-211.
9. O'Donoghue, K., et al., *Identification of fetal mesenchymal stem cells in maternal blood: implications for non-invasive prenatal diagnosis.* Mol Hum Reprod, 2003. **9**(8): p. 497-502.
10. Davies, A.J., *Immigration control in the vertebrate body with special reference to chimerism.* Chimerism, 2012. **3**(1): p. 1-8.
11. Khosrotehrani, K., et al., *Natural history of fetal cell microchimerism during and following murine pregnancy.* J Reprod Immunol, 2005. **66**(1): p. 1-12.
12. Owen, R.D., *Immunogenetic Consequences of Vascular Anastomoses between Bovine Twins.* Science, 1945. **102**(2651): p. 400-1.
13. Dunsford, I., et al., *A human blood-group chimera.* Br Med J, 1953. **2**(4827): p. 81.
14. Mujahid, A. and F.L. Dickert, *Blood Group Typing: From Classical Strategies to the Application of Synthetic Antibodies Generated by Molecular Imprinting.* Sensors (Basel), 2015. **16**(1).
15. Hong, X., et al., *A dispermic chimera was identified in a healthy man with mixed field agglutination reaction in ABO blood grouping and mosaic 46, XY/46, XX karyotype.* Transfus Apher Sci, 2013. **48**(2): p. 223-8.
16. Aoki, R., et al., *Blood chimerism in monochorionic twins conceived by induced ovulation: case report.* Hum Reprod, 2006. **21**(3): p. 735-7.
17. Chung, Y.N., et al., *The first case of congenital blood chimerism in two of the triplets in Korea.* J Clin Lab Anal, 2018. **32**(8): p. e22580.
18. Eikmans, M. and F.H. Claas, *HLA-targeted cell sorting of microchimeric cells opens the way to phenotypical and functional characterization.* Chimerism, 2011. **2**(4): p. 114-6.
19. Farag, T.I., et al., *Unilateral true hermaphrodite with 46,XX/46,XY dispermic chimerism.* J Med Genet, 1987. **24**(12): p. 784-6.

20. Jang, J.H., et al., *Blood chimerism in a dizygotic dichorionic pregnancy.* Korean J Lab Med, 2010. **30**(5): p. 521-4.

21. Chen, C.P., et al., *Prenatal diagnosis, sonographic findings and molecular genetic analysis of a 46,XX/46,XY true hermaphrodite chimera.* Prenat Diagn, 2005. **25**(6): p. 502-6.

22. Repas-Humpe, L.M., et al., *A dispermic chimerism in a 2-year-old Caucasian boy.* Ann Hematol, 1999. **78**(9): p. 431-4.

23. Shin, S.Y., et al., *Identification of the mechanism underlying a human chimera by SNP array analysis.* Am J Med Genet A, 2012. **158a**(9): p. 2119-23.

24. Green, A.J., et al., *Chimaerism shown by cytogenetics and DNA polymorphism analysis.* J Med Genet, 1994. **31**(10): p. 816-7.

25. Niu, D.M., et al., *Mosaic or chimera? Revisiting an old hypothesis about the cause of the 46,XX /46,XY hermaphrodite.* J Pediatr, 2002. **140**(6): p. 732-5.

26. Mullis, K.B., *The unusual origin of the polymerase chain reaction.* Sci Am, 1990. **262**(4): p. 56-61, 64-5.

27. Cirello, V., et al., *Fetal cell microchimerism in papillary thyroid cancer: studies in peripheral blood and tissues.* Int J Cancer, 2010. **126**(12): p. 2874-8.

28. Gilmore, G.L., et al., *Fetal-maternal microchimerism in normal parous females and parous female cancer patients.* Exp Hematol, 2008. **36**(9): p. 1073-7.

29. Peters, H.E., et al., *Low prevalence of male microchimerism in women with Mayer-Rokitansky-Kuster-Hauser syndrome.* Hum Reprod, 2019.

30. Lambert, N.C., *Microchimerism and rheumatic diseases.* Joint Bone Spine, 2012. **79**(6): p. 563-4.

31. Nelson, J.L., et al., *Maternal microchimerism in peripheral blood in type 1 diabetes and pancreatic islet beta cell microchimerism.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1637-42.

32. Lambert, N.C., et al., *Quantification of maternal microchimerism by HLA-specific real-time polymerase chain reaction: studies of healthy women and women with scleroderma.* Arthritis Rheum, 2004. **50**(3): p. 906-14.

33. Kristt, D., et al., *Assessing quantitative chimerism longitudinally: technical considerations, clinical applications and routine feasibility.* Bone Marrow Transplant, 2007. **39**(5): p. 255-68.

34. Castella, V., M. Lesta Mdel, and P. Mangin, *One person with two DNA profiles: a(nother) case of mosaicism or chimerism.* Int J Legal Med, 2009. **123**(5): p. 427-30.

35. Johnson, K.L., D.K. Zhen, and D.W. Bianchi, *The use of fluorescence in situ hybridization (FISH) on paraffin-embedded tissue sections for the study of microchimerism.* Biotechniques, 2000. **29**(6): p. 1220-4.

36. Moher, D., et al., *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.* Syst Rev, 2015. **4**: p. 1.

37. Ron, M., et al., *Chimaerism detection in bovine twins, triplets and quadruplets using sex chromosome-linked markers.* Anim Genet, 2011. **42**(2): p. 208-11.

38. Lillie, F.R., *The Theory of the Free-Martin.* Science, 1916. **43**(1113): p. 611-3.

39. Vigier, B., et al., *Origin of anti-Mullerian hormone in bovine freemartin fetuses.* J Reprod Fertil, 1984. **70**(2): p. 473-9.

40. Padula, A.M., *The freemartin syndrome: an update.* Anim Reprod Sci, 2005. **87**(1-2): p. 93-109.

41. Kenny, D.E., et al., *Freemartinism in a captive herd of Rocky Mountain bighorn sheep (Ovis canadensis).* J Wildl Dis, 1992. **28**(3): p. 494-8.

42. Jimenez, D.F., et al., *Fetal CD34+ cells in the maternal circulation and long-term microchimerism in rhesus monkeys (Macaca mulatta).* Transplantation, 2005. **79**(2): p. 142-6.

43. Benirschke, K., J.M. Anderson, and L.E. Brownhill, *Marrow Chimerism in Marmosets.* Science, 1962. **138**(3539): p. 513-5.

44. Signer, E.N., G. Anzenberger, and A.J. Jeffreys, *Chimaeric and constitutive DNA fingerprints in the common marmoset (Callithrix jacchus).* Primates, 2000. **41**(1): p. 49-61.

45. Ross, C.N., J.A. French, and G. Orti, *Germ-line chimerism and paternal care in marmosets (Callithrix kuhlii).* Proc Natl Acad Sci U S A, 2007. **104**(15): p. 6278-82.

46. Lyons, L.A., *Genetic testing in domestic cats.* Mol Cell Probes, 2012. **26**(6): p. 224-30.

47. Axiak-Bechtel, S.M., et al., *Y-chromosome DNA is present in the blood of female dogs suggesting the presence of fetal microchimerism.* PLoS One, 2013. **8**(7): p. e68114.

48. Kumar, S.R., et al., *The health effects of fetal microchimerism can be modeled in companion dogs.* Chimerism, 2013. **4**(4): p. 139-41.

49. Szczerbal, I., et al., *A case of leucocyte chimerism (78,XX/78,XY) in a dog with a disorder of sexual development.* Reprod Domest Anim, 2014. **49**(3): p. e31-4.

50. Perlman, R.L., *Mouse models of human disease: An evolutionary perspective.* Evol Med Public Health, 2016. **2016**(1): p. 170-6.

51. Kara, R.J., et al., *A mouse model for fetal maternal stem cell transfer during ischemic cardiac injury.* Clin Transl Sci, 2012. **5**(4): p. 321-8.

52. Su, E.C., et al., *Murine maternal cell microchimerism: analysis using real-time PCR and in vivo imaging.* Biol Reprod, 2008. **78**(5): p. 883-7.

53. Kara, R.J., et al., *Fetal cells traffic to injured maternal myocardium and undergo cardiac differentiation.* Circ Res, 2012. **110**(1): p. 82-93.

54. Pritchard, S., et al., *Comprehensive analysis of genes expressed by rare microchimeric fetal cells in the maternal mouse lung.* Biol Reprod, 2012. **87**(2): p. 42.

55. Tan, X.W., et al., *Fetal microchimerism in the maternal mouse brain: a novel population of fetal progenitor or stem cells able to cross the blood-brain barrier?* Stem Cells, 2005. **23**(10): p. 1443-52.

56. Vogelgesang, A., et al., *Cigarette smoke exposure during pregnancy alters fetomaternal cell trafficking leading to retention of microchimeric cells in the maternal lung.* PLoS One, 2014. **9**(5): p. e88285.

57. Hadjiathanasiou, C.G., et al., *True hermaphroditism: genetic variants and clinical management.* J Pediatr, 1994. **125**(5 Pt 1): p. 738-44.

58. Ramsay, M., et al., *Chimerism in black southern African patients with true hermaphroditism 46,XX/47XY,+21 and 46,XX/46,XY.* Ann N Y Acad Sci, 2009. **1151**: p. 68-76.

59. Wang, L., F.S. Wang, and M.E. Gershwin, *Human autoimmune diseases: a comprehensive update.* J Intern Med, 2015. **278**(4): p. 369-95.

60. Lepez, T., et al., *Fetal microchimeric cells in blood of women with an autoimmune thyroid disease.* PLoS One, 2011. **6**(12): p. e29646.

61. Jacobson, D.L., et al., *Epidemiology and estimated population burden of selected autoimmune diseases in the United States.* Clin Immunol Immunopathol, 1997. **84**(3): p. 223-43.

62. Lambert, N.C., et al., *From the simple detection of microchimerism in patients with autoimmune diseases to its implication in pathogenesis.* Ann N Y Acad Sci, 2001. **945**: p. 164-71.

63. Furst, D.E., et al., *A syndrome resembling progressive systemic sclerosis after bone marrow transplantation.* Arthritis & Rheumatism, 1979. **22**(8): p. 904-910.

64. Burastero, S.E., et al., *Cellular microchimerism as a lifelong physiologic status in parous women: an immunologic basis for its amplification in patients with systemic sclerosis.* Arthritis Rheum, 2003. **48**(4): p. 1109-16.

65. Rak, J.M., et al., *Male microchimerism and HLA compatibility in French women with sclerodema: a different profile in limited and diffuse subset.* Rheumatology (Oxford), 2009. **48**(4): p. 363-6.

66. Lambert, N.C., et al., *Male microchimerism in healthy women and women with scleroderma: cells or circulating DNA? A quantitative answer.* Blood, 2002. **100**(8): p. 2845-51.

67. Johnson, K.L., et al., *Fetal cell microchimerism in tissue from multiple sites in women with systemic sclerosis.* Arthritis Rheum, 2001. **44**(8): p. 1848-54.

68. Ohtsuka, T., et al., *Quantitative analysis of microchimerism in systemic sclerosis skin tissue.* Arch Dermatol Res, 2001. **293**(8): p. 387-91.

69. Sawaya, H.H., S.A. Jimenez, and C.M. Artlett, *Quantification of fetal microchimeric cells in clinically affected and unaffected skin of patients with systemic sclerosis.* Rheumatology (Oxford), 2004. **43**(8): p. 965-8.

70. Kekow, M., et al., *Long-term persistence and effects of fetal microchimerisms on disease onset and status in a cohort of women with rheumatoid arthritis and systemic lupus erythematosus.* BMC Musculoskelet Disord, 2013. **14**: p. 325.

71. Rak, J.M., et al., *Transfer of the shared epitope through microchimerism in women with rheumatoid arthritis.* Arthritis Rheum, 2009. **60**(1): p. 73-80.

72. Yan, Z., et al., *Acquisition of the rheumatoid arthritis HLA shared epitope through microchimerism.* Arthritis Rheum, 2011. **63**(3): p. 640-4.

73. Lambert, N.C., et al., *Cutting edge: persistent fetal microchimerism in T lymphocytes is associated with HLA-DQA1*0501: implications in autoimmunity.* J Immunol, 2000. **164**(11): p. 5545-8.

74. Johnson, K.L., et al., *Microchimerism in a female patient with systemic lupus erythematosus.* Arthritis Rheum, 2001. **44**(9): p. 2107-11.

75. Kuroki, M., et al., *Detection of maternal-fetal microchimerism in the inflammatory lesions of patients with Sjogren's syndrome.* Ann Rheum Dis, 2002. **61**(12): p. 1041-6.

76. Toda, I., et al., *Lack of evidence for an increased microchimerism in the circulation of patients with Sjogren's syndrome.* Ann Rheum Dis, 2001. **60**(3): p. 248-53.

77. Klintschar, M., et al., *Evidence of fetal microchimerism in Hashimoto's thyroiditis.* J Clin Endocrinol Metab, 2001. **86**(6): p. 2494-8.

78. Renne, C., et al., *Thyroid fetal male microchimerisms in mothers with thyroid disorders: presence of Y-chromosomal immunofluorescence in thyroid-infiltrating lymphocytes is more prevalent in Hashimoto's thyroiditis and Graves' disease than in follicular adenomas.* J Clin Endocrinol Metab, 2004. **89**(11): p. 5810-4.

79. Tanaka, A., et al., *Fetal microchimerism alone does not contribute to the induction of primary biliary cirrhosis.* Hepatology, 1999. **30**(4): p. 833-8.

80. Schoniger-Hekele, M., et al., *Lack of evidence for involvement of fetal microchimerism in pathogenesis of primary biliary cirrhosis.* Dig Dis Sci, 2002. **47**(9): p. 1909-14.

81. Invernizzi, P., et al., *Blood fetal microchimerism in primary biliary cirrhosis.* Clin Exp Immunol, 2000. **122**(3): p. 418-22.

82. Nomura, K., et al., *Lack of evidence for leukocyte maternal microchimerism in primary biliary cirrhosis.* World J Gastroenterol, 2004. **10**(16): p. 2415-6.

83. Reed, A.M., et al., *Chimerism in children with juvenile dermatomyositis.* Lancet, 2000. **356**(9248): p. 2156-7.

84. Ye, Y., et al., *Maternal microchimerism in muscle biopsies from children with juvenile dermatomyositis.* Rheumatology (Oxford), 2012. **51**(6): p. 987-91.

85. Artlett, C.M., F.W. Miller, and L.G. Rider, *Persistent maternally derived peripheral microchimerism is associated with the juvenile idiopathic inflammatory myopathies.* Rheumatology (Oxford), 2001. **40**(11): p. 1279-84.

86. Reed, A.M., *Microchimerism in children with rheumatic disorders: what does it mean?* Curr Rheumatol Rep, 2003. **5**(6): p. 458-62.

87. Reed, A.M. and J.D. Stirling, *Association of the HLA-DQA1*0501 allele in multiple racial groups with juvenile dermatomyositis.* Hum Immunol, 1995. **44**(3): p. 131-5.

88. Ye, J., M. Vives-Pi, and K.M. Gillespie, *Maternal microchimerism: increased in the insulin positive compartment of type 1 diabetes pancreas but not in infiltrating immune cells or replicating islet cells.* PLoS One, 2014. **9**(1): p. e86985.

89. Selva-O'Callaghan, A., et al., *Lack of evidence of foetal microchimerism in female Spanish patients with systemic sclerosis.* Lupus, 2003. **12**(1): p. 15-20.

90. Lambert, N.C., et al., *Male microchimerism in women with systemic sclerosis and healthy women who have never given birth to a son.* Ann Rheum Dis, 2005. **64**(6): p. 845-8.

91. Bulow Pedersen, I., et al., *Lack of association between thyroid autoantibodies and parity in a population study argues against microchimerism as a trigger of thyroid autoimmunity.* Eur J Endocrinol, 2006. **154**(1): p. 39-45.

92. Greer, L.G., et al., *Antithyroid antibodies and parity: further evidence for microchimerism in autoimmune thyroid disease.* Am J Obstet Gynecol, 2011. **205**(5): p. 471.e1-4.

93. Cirello, V., et al., *Fetal cell microchimerism: a protective role in autoimmune thyroid diseases.* Eur J Endocrinol, 2015. **173**(1): p. 111-8.

94. Mosca, M., et al., *Correlations of Y chromosome microchimerism with disease activity in patients with SLE: analysis of preliminary data.* Ann Rheum Dis, 2003. **62**(7): p. 651-4.

95. Kanold, A.M., et al., *A research study of the association between maternal microchimerism and systemic lupus erythematosus in adults: a comparison between patients and healthy controls based on single-nucleotide polymorphism using quantitative real-time PCR.* PLoS One, 2013. **8**(9): p. e74534.

96. Gadi, V.K., et al., *Case-control study of fetal microchimerism and breast cancer.* PLoS One, 2008. **3**(3): p. e1706.

97. Cirello, V., et al., *Fetal cell microchimerism in papillary thyroid cancer: a possible role in tumor damage and tissue repair.* Cancer Res, 2008. **68**(20): p. 8482-8.

98. Cirello, V., et al., *Fetal cell microchimerism in papillary thyroid cancer: A role in the outcome of the disease.* Int J Cancer, 2015. **137**(12): p. 2989-93.

99. Cirello, V. and L. Fugazzola, *Positive effect of fetal cell microchimerism on tumor presentation and outcome in papillary thyroid cancer.* Chimerism, 2014. **5**(3-4): p. 106-8.

100. Gadi, V.K., *Fetal microchimerism in breast from women with and without breast cancer.* Breast Cancer Res Treat, 2010. **121**(1): p. 241-4.

101. Kamper-Jorgensen, M., et al., *Opposite effects of microchimerism on breast and colon cancer.* Eur J Cancer, 2012. **48**(14): p. 2227-35.

102. Kamper-Jorgensen, M., *Microchimerism and survival after breast and colon cancer diagnosis.* Chimerism, 2012. **3**(3): p. 72-3.

103. Geck, P., *Symptotic detection of chimerism: Y does it matter?* Chimerism, 2013. **4**(4): p. 144-6.

104. Broestl, L., J.B. Rubin, and S. Dahiya, *Fetal microchimerism in human brain tumors.* Brain Pathol, 2018. **28**(4): p. 484-494.

105. Hromadnikova, I., et al., *The occurrence of fetal microchimeric cells in endometrial tissues is a very common phenomenon in benign uterine disorders, and the lower prevalence of fetal microchimerism is associated with better uterine cancer prognoses.* DNA Cell Biol, 2014. **33**(1): p. 40-8.

106. Nguyen Huu, S., et al., *Fetal microchimeric cells participate in tumour angiogenesis in melanomas occurring during pregnancy.* Am J Pathol, 2009. **174**(2): p. 630-7.

107. Grigorova, M. and B.B. Sherwin, *Thyroid hormones and cognitive functioning in healthy, euthyroid women: a correlational study.* Horm Behav, 2012. **61**(4): p. 617-22.

108. Garland, T., Jr., M. Zhao, and W. Saltzman, *Hormones and the Evolution of Complex Traits: Insights from Artificial Selection on Behavior.* Integr Comp Biol, 2016. **56**(2): p. 207-24.

109. Arteaga-Henriquez, G., et al., *Low-Grade Inflammation as a Predictor of Antidepressant and Anti-Inflammatory Therapy Response in MDD Patients: A Systematic Review of the Literature in Combination With an Analysis of Experimental Data Collected in the EU-MOODINFLAME Consortium.* Front Psychiatry, 2019. **10**: p. 458.

110. Gibney, S.M. and H.A. Drexhage, *Evidence for a dysregulated immune system in the etiology of psychiatric disorders.* J Neuroimmune Pharmacol, 2013. **8**(4): p. 900-20.

111. Handelsman, D.J., A.L. Hirschberg, and S. Bermon, *Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance.* Endocr Rev, 2018. **39**(5): p. 803-829.

112. Genel, M., J.L. Simpson, and A. de la Chapelle, *The Olympic Games and Athletic Sex Assignment.* JAMA, 2016. **316**(13): p. 1359-1360.

113. Geniole, S.N. and J.M. Carre, *Human social neuroendocrinology: Review of the rapid effects of testosterone.* Horm Behav, 2018. **104**: p. 192-205.

114. Sandberg, D.E., M. Gardner, and P.T. Cohen-Kettenis, *Psychological aspects of the treatment of patients with disorders of sex development.* Semin Reprod Med, 2012. **30**(5): p. 443-52.

115. de Neve-Enthoven, N.G., et al., *Psychosocial well-being in Dutch adults with disorders of sex development.* J Psychosom Res, 2016. **83**: p. 57-64.

116. Boddy, A.M., et al., *Fetal microchimerism and maternal health: a review and evolutionary analysis of cooperation and conflict beyond the womb.* Bioessays, 2015. **37**(10): p. 1106-18.

117. Khosrotehrani, K., et al., *The influence of fetal loss on the presence of fetal cell microchimerism: a systematic review.* Arthritis Rheum, 2003. **48**(11): p. 3237-41.

118. Gammill, H.S., et al., *Effect of parity on fetal and maternal microchimerism: interaction of grafts within a host?* Blood, 2010. **116**(15): p. 2706-12.

119. Gammill, H.S., et al., *Pregnancy, microchimerism, and the maternal grandmother.* PLoS One, 2011. **6**(8): p. e24101.

120. Gammill, H.S., et al., *Microchimerism in recurrent miscarriage.* Cell Mol Immunol, 2014. **11**(6): p. 589-94.

121. Gammill, H.S., et al., *Cellular fetal microchimerism in preeclampsia.* Hypertension, 2013. **62**(6): p. 1062-7.

122. Li, D.K. and S. Wi, *Changing paternity and the risk of preeclampsia/eclampsia in the subsequent pregnancy.* Am J Epidemiol, 2000. **151**(1): p. 57-62.

123. Guettier, C., et al., *Male cell microchimerism in normal and diseased female livers from fetal life to adulthood.* Hepatology, 2005. **42**(1): p. 35-43.

124. Muraji, T., et al., *Maternal microchimerism in underlying pathogenesis of biliary atresia: quantification and phenotypes of maternal cells in the liver.* Pediatrics, 2008. **121**(3): p. 517-21.

125. Kobayashi, H., et al., *Maternal microchimerism in biliary atresia.* J Pediatr Surg, 2007. **42**(6): p. 987-91; discussion 991.

126. Muraji, T., *Maternal microchimerism in biliary atresia: are maternal cells effector cells, targets, or just bystanders?* Chimerism, 2014. **5**(1): p. 1-5.

127. Bayes-Genis, A., et al., *Identification of male cardiomyocytes of extracardiac origin in the hearts of women with male progeny: male fetal cell microchimerism of the heart.* J Heart Lung Transplant, 2005. **24**(12): p. 2179-83.

128. Yan, Z., et al., *Male microchimerism in women without sons: quantitative assessment and correlation with pregnancy history.* Am J Med, 2005. **118**(8): p. 899-906.

129. Muller, A.C., et al., *Microchimerism of male origin in a cohort of Danish girls.* Chimerism, 2015. **6**(4): p. 65-71.

130. de Bellefon, L.M., et al., *Cells from a vanished twin as a source of microchimerism 40 years later.* Chimerism, 2010. **1**(2): p. 56-60.

131. Dierselhuis, M.P., et al., *Transmaternal cell flow leads to antigen-experienced cord blood.* Blood, 2012. **120**(3): p. 505-10.

132. Utter, G.H., et al., *Transfusion-associated microchimerism.* Vox Sang, 2007. **93**(3): p. 188-95.

133. Kamper-Jorgensen, M., et al., *Predictors of male microchimerism.* Chimerism, 2012. **3**(3): p. 1-8.

134. Brix, T.H., et al., *Aggregation of thyroid autoantibodies in twins from opposite-sex pairs suggests that microchimerism may play a role in the early stages of thyroid autoimmunity.* J Clin Endocrinol Metab, 2009. **94**(11): p. 4439-43.

135. Ekelund, C.K., et al., *Dizygotic monochorionic twin pregnancy conceived following intracytoplasmic sperm injection treatment and complicated by twin-twin transfusion syndrome and blood chimerism.* Ultrasound Obstet Gynecol, 2008. **32**(6): p. 832-4.

136. Le Bras, A.M., et al., *Confined blood chimerism in a monochorionic dizygotic sex discordant twin pregnancy conceived after induced ovulation.* Birth Defects Research Part a-Clinical and Molecular Teratology, 2016. **106**(4): p. 298-303.

137. Rodriguez-Buritica, D., et al., *Sex-discordant monochorionic twins with blood and tissue chimerism.* Am J Med Genet A, 2015. **167a**(4): p. 872-7.

138. Shalev, S.A., et al., *Evidence for blood chimerism in dizygotic spontaneous twin pregnancy discordant for Down syndrome.* Prenat Diagn, 2006. **26**(9): p. 782-4.

139. Smeets, D., et al., *Monochorionic dizygous twins presenting with blood chimerism and discordant sex.* Twin Res Hum Genet, 2013. **16**(4): p. 799-801.

140. Walker, S.P., S. Meagher, and S.M. White, *Confined blood chimerism in monochorionic dizygous (MCDZ) twins.* Prenat Diagn, 2007. **27**(4): p. 369-72.

141. Peters, H.E., et al., *Unusual Twinning Resulting in Chimerism: A Systematic Review on Monochorionic Dizygotic Twins.* Twin Res Hum Genet, 2017. **20**(2): p. 161-168.

142. van Dongen, J., et al., *The continuing value of twin studies in the omics era.* Nat Rev Genet, 2012. **13**(9): p. 640-53.

143. Finnicum, C.T., et al., *Metataxonomic Analysis of Individuals at BMI Extremes and Monozygotic Twins Discordant for BMI.* Twin Res Hum Genet, 2018. **21**(3): p. 203-213.

144. Li, W., et al., *On the power of epigenome-wide association studies using a disease-discordant twin design.* Bioinformatics, 2018. **34**(23): p. 4073-4078.

145. Artlett, C.M., et al., *HLA-DQA1 is not an apparent risk factor for microchimerism in patients with various autoimmune diseases and in healthy individuals.* Arthritis Rheum, 2003. **48**(9): p. 2567-72.

146. vanDijk, B.A., D.I. Boomsma, and A.J.M. deMan, *Blood group chimerism in human multiple births is not rare.* American Journal of Medical Genetics, 1996. **61**(3): p. 264-268.

147. Kliman, D., et al., *Ultra-Sensitive Droplet Digital PCR for the Assessment of Microchimerism in Cellular Therapies.* Biol Blood Marrow Transplant, 2018. **24**(5): p. 1069-1078.

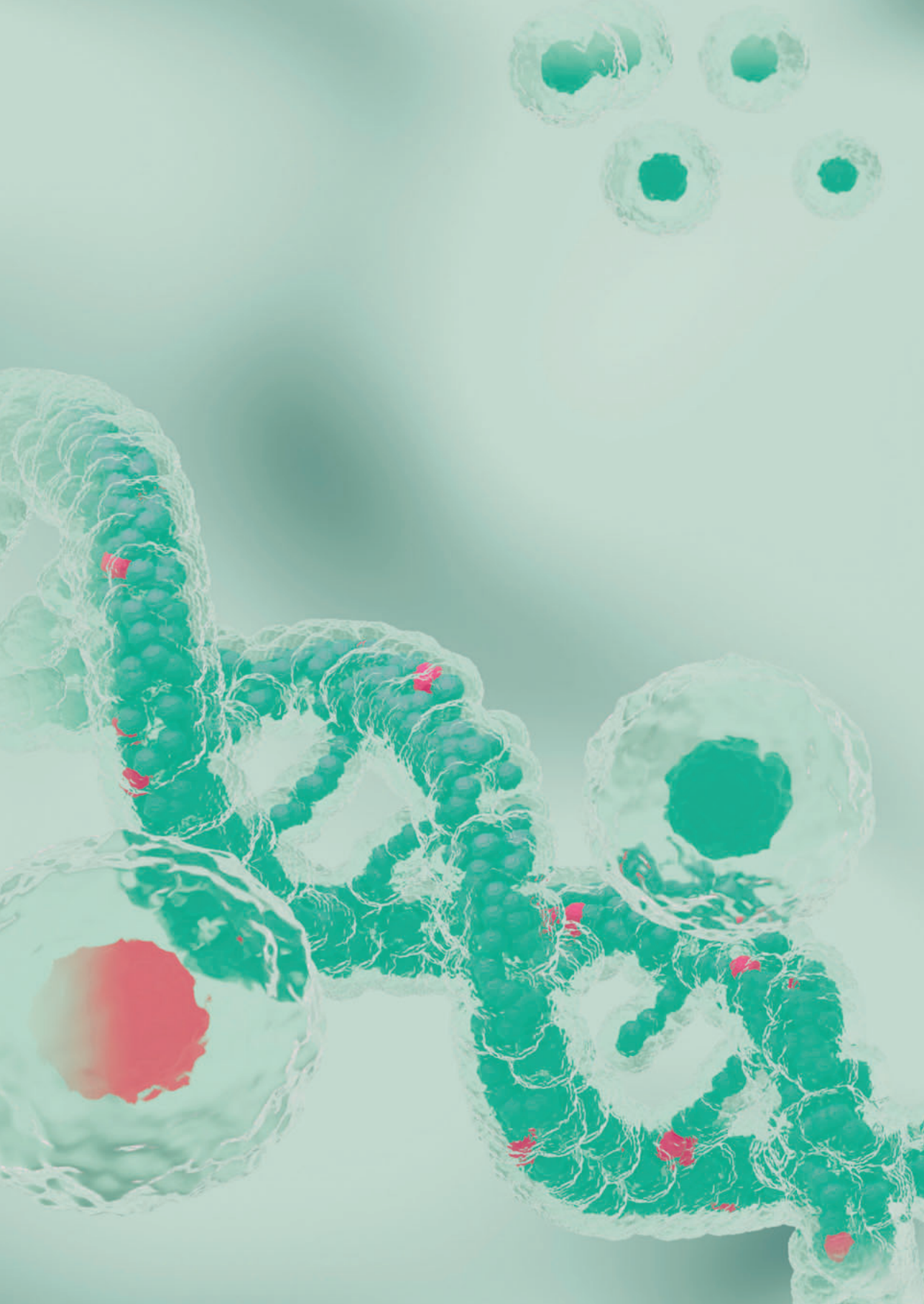
148. Ciolino, A.L., M.E. Tang, and R. Bryant, *Statistical treatment of fluorescence in situ hybridization validation data to generate normal reference ranges using Excel functions.* J Mol Diagn, 2009. **11**(4): p. 330-3.

149. Sharpe, C., et al., *Mixed field reactions in ABO and Rh typing chimerism likely resulting from twin haematopoiesis.* Blood Transfusion, 2014. **12**(4): p. 608-610.

150. Summers, T., Jr., et al., *The value of automated gel column agglutination technology in the identification of true inherited D blood types in massively transfused patients.* Transfusion, 2009. **49**(8): p. 1672-7.

151. Thiele, K., et al., *Comparative sensitivity analyses of quantitative polymerase chain reaction and flow cytometry in detecting cellular microchimerism in murine tissues.* J Immunol Methods, 2014. **406**: p. 74-82.

152. Greendyke, R.M., J.L. Wormer, and J.C. Banzhaf, *Quality assurance in the blood bank. Studies of technologist performance.* Am J Clin Pathol, 1979. **71**(3): p. 287-90.



3

LOW PREVALENCE OF MALE MICROCHIMERISM IN WOMEN WITH MAYER-ROKITANSKY-KÜSTER-HAUSER SYNDROME

Published as:

Peters, H. E.[†], Johnson, B. N.[†], Ehli, E. A., Micha, D., Verhoeven, M. O., Davies, G. E., Dekker, J., Overbeek, van den Berg, M.H., M., van Dulmen-den Broeder, E., van Leeuwen, F.E., Mijatovic, V., Boomsma, D. I., & Lambalk, C. B. (2019).

Low prevalence of male microchimerism in women with Mayer-Rokitansky-Kuster-Hauser syndrome. *Human Reproduction*, 34(6), 1117-1125.

https://doi.org/10.1093/humrep/dez044

[†]The authors consider that the first two authors should be regarded as joint first authors.

ABSTRACT

Study question: Is there an increased prevalence of male microchimerism in women with Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome, as evidence of fetal exposure of blood and anti-Müllerian hormone (AMH) from a (vanished) male co-twin resulting in regression of the Müllerian duct derivatives?

Summary answer: Predominant absence of male microchimerism in adult women with MRKH syndrome does not support our hypothesis that intrauterine blood exchange with a (vanished) male co-twin is the pathophysiological mechanism.

What is known already: The etiology of MRKH is unclear. Research on the phenotype analogous condition in cattle (freemartinism) has yielded the hypothesis that Müllerian duct development is inhibited by exposure to AMH *in utero*. In cattle, the male co-twin has been identified as the source for AMH which is transferred via placental blood exchange. In human twins a similar exchange of cellular material has been documented by detection of chimerism but it is unknown whether this has clinical consequences.

Study design, size, duration: An observational case–control study was performed to compare the presence of male microchimerism in women with MRKH syndrome and control women. Through recruitment via the Dutch patients' association of women with MRKH (comprising 300 members who were informed by email or regular mail), we enrolled 96 patients between January 2017 and July 2017. The control group consisted of 100 women who reported never having been pregnant.

Participants/materials, setting, methods: After written informed consent, peripheral blood samples were obtained by venipuncture, and genomic DNA was extracted. Male microchimerism was detected by Y-chromosome–specific real-time quantitative PCR, with use of DYS14 marker. Possible other sources for microchimerism, for example older brothers, were evaluated using questionnaire data.

Main results and the role of chance: The final analysis included 194 women: 95 women with MRKH syndrome with a mean age of 40.9 years and 99 control women with a mean age of 30.2 years. In total, 54 women (56.8%) were identified as having typical MRKH syndrome, and 41 women (43.2%) were identified as having atypical MRKH syndrome (when extra-genital malformations were present). The prevalence of male microchimerism was significantly higher in the control group than in the MRKH group (17.2% versus 5.3%, P = 0.009). After correcting for age, women in the control group were 5.8 times more likely to have male microchimerism (odds ratio 5.84 (CI 1.59–21.47), P = 0.008). The mean concentration of male microchimerism in the positive samples was 56.0 male genome equivalent per 1 000 000 cells. The prevalence of male microchimerism was similar in women with typical MRKH syndrome and atypical MRKH syndrome (5.6%

versus 4.9%, P = 0.884). There were no differences between women with or without microchimerism in occurrence of alternative sources of XY cells, such as older brothers, previous blood transfusion, or history of sexual intercourse.

Limitations, reason for caution: We are not able to draw definitive conclusions regarding the occurrence of AMH exchange during embryologic development in women with MRKH syndrome. Our subject population includes all adult women and therefore is reliant on long-term prevalence of microchimerism. Moreover, we have only tested blood, and, theoretically, the cells may have grafted anywhere in the body during development. It must also be considered that the exchange of AMH may occur without the transfusion of XY cells and therefore cannot be discovered by chimerism detection.

Wider implications of the findings: This is the first study to test the theory that freemartinism causes the MRKH syndrome in humans. The study aimed to test the presence of male microchimerism in women with MRKH syndrome as a reflection of early fetal exposure to blood and AMH from a male (vanished) co-twin. We found that male microchimerism was only present in 5.3% of the women with MRKH syndrome, a significantly lower percentage than in the control group (17.2%). Our results do not provide evidence for an increased male microchimerism in adult women with MRKH as a product of intrauterine blood exchange. However, the significant difference in favor of the control group is of interest to the ongoing discussion on microchimeric cell transfer and the possible sources of XY cells.

Study funding/competing interest(s): None.

Trial registration number: Dutch trial register, NTR5961

Key words: Mayer-Rokitansky-Küster-Hauser syndrome, Müllerian aplasia, microchimerism, etiology, freemartinism

INTRODUCTION

Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome is characterized by congenital aplasia of the uterus and the upper part of the vagina. It affects around 1 in 5000 females [1]. The diagnosis is usually made in adolescence after the presentation of primary amenorrhea. Further examination reveals vaginal agenesis with absence of the uterus, normal secondary sex characteristics and normal female 46,XX karyotype [2]. It may also be associated with renal and/or skeletal malformations, classified as atypical MRKH syndrome [3, 4].

The etiology of the MRKH syndrome is unclear [5]. Embryological evidence supports the hypothesis that this syndrome occurs as a result of failure of Müllerian duct development. In the normal male embryo the Sertoli cells of the testes produce anti-Müllerian hormone (AMH), resulting in regression of the Müllerian duct. This inhibitory action of AMH on the Müllerian duct starts during the fifth week of pregnancy and is progressive during the critical time window of uterine development [6]. Genetic activation of AMH or its receptor has been implicated as a cause of MRKH syndrome but without any supporting results [7, 8].

In cattle, a similar phenotypical syndrome to MRKH exists: a so-called freemartin. In this infertile female calf, the absence of Müllerian structures occurs due to intrauterine AMH exposure, originating from a male co-twin [9]. Vascular connections in the placenta transport AMH from the male to the female calf [10]. As a result of this intrauterine blood exchange via the placenta, the freemartins are ‘chimeras’; in addition to the normal XX cells, they have an extra XY cell line, originating from their co-twin. The term chimerism means that two genetically different cell lines are present in one individual, originating from more than one zygote.

In humans, intrauterine cell trafficking between twins can also result in chimerism [11, 12]. It is unknown whether this has clinical consequences. Bogdanova et al. (2010) reported a possible case of ‘human freemartinism’ in a female twin with aplasia of the uterus. In this female, blood exchange via the placenta with her brother had resulted in male chimerism (the presence of XY cells) [13]. In addition, a vanishing twin can leave its traces in the form of microchimerism, in which a second cell population is present at variable concentrations in the surviving fetus [14].

It should also be mentioned that fetal-maternal exchange in a normal pregnancy is a well-known source for chimeric cells [15]. Fetal microchimerism refers to the phenomenon of fetal cells entering the maternal circulation during pregnancy. As the exchange between fetus and mother is bidirectional, the presence of maternal cells in the circulation of their children is called maternal chimerism. Chimerism can also occur following transplantation or blood transfusion [16].

There have been no studies to date which examined the presence of male chimerism in women with the MRKH syndrome. The main purpose of the research reported here was to study whether the etiology associated with freemartinism could be the cause of the MRKH syndrome in humans. We hypothesized that a male – possibly vanished – co twin was present during the early embryological development of women with MRKH syndrome, resulting in exchange of AMH via placental vascular connections. Therefore, we investigated the presence of male microchimerism (as a result of cell trafficking from male twin to female twin) in MRKH patients and compared these results with control women.

PATIENTS AND METHODS

Study subjects

This observational case-control study compared the presence of male microchimerism in women diagnosed with the MRKH syndrome and control women. Through recruitment via the Dutch patients' association of women with MRKH ('Stichting MRK-vrouwen') we enrolled 96 women diagnosed with the MRKH syndrome (in total, the 300 members were informed by email or regular mail and the participation rate was 32%). All participants were 18 years or older and provided written informed consent prior to enrollment. Between January 2017 and July 2017 blood samples were obtained by venipuncture, by one researcher (H.E.P.). The blood samples were collected in EDTA vacutainer® tubes. All participants completed a questionnaire, comprising questions about demographic information, medical history, MRKH diagnosis, and family history.

The control group comprised of women who volunteered to participate in an earlier study at our hospital. This study evaluated reproductive functioning in female childhood cancer survivors (DCOG LATER-VEVO study, NL15106.029.06) [17, 18]. For

this study women who did not have a history of cancer in the past were included as control subjects. Phenotypic data was collected between 2008 and 2014 by questionnaires, and biological material was collected via blood sampling. Out of the 390 women who provided biological material for the study, we selected 100 women who reported to have never been pregnant for the control group.

MRKH diagnosis

The MRKH diagnosis of the participants was confirmed by contacting the general practitioner or gynecologist, to retrieve the detailed information about the diagnosis based on physical examination, imaging or laparoscopy results. The patients were identified as having typical MRKH syndrome (also referred to as type 1) in case of no known other malformations and as atypical (type 2) MRKH syndrome when renal and/or skeletal malformations were present (Oppelt, et al., 2006).

Preparation of samples

Genomic DNA was extracted from peripheral venous blood according to standard procedures. The extracted DNA was stored at -80°C until further processing. The DNA concentration ranged from 2.48 to 261 ng/µL (mean 77.47 ng/µL). The A260/A280 ratio ranged from 1.33 – 2.46 (mean 1.85).

Analysis of male microchimerism

Real-time quantitative polymerase chain reaction (qPCR) was performed using a QuantStudio™ 7 Flex system (Applied Biosystems, Waltham, MA, USA). The DNA sequence utilized for detection of male genome targets a Y-chromosome specific region, *DYS14*. Amplification primers and probe sequences for *DYS14* have been described previously [19]. Also the rational for using this region as a Y-chromosome target is detailed [19]. As a measure of the total input quantity of genomic mass per reaction, qPCR of the *β-globin* gene was measured in parallel. Primers and probes for *β-globin* are also previously described [20]. All genome equivalent (GEq) conversions utilized a conversion factor of 6.6 pg DNA per cell [21]. Male GEq was reported per 1 million cells and calculated using the following equation:

$$C = \frac{Q_{DYS14}}{6.6pg} \times \frac{6.6pg \times 10^6}{Q_{\beta-\text{globin}}}$$

where C = concentration of male GEq per 1 million cells; Q_{DYS14} = mean mass of *DYS14* (ng) determined by qPCR; and $Q_{\beta\text{-Globin}}$ = mean mass of *β-globin* (ng) determined by qPCR.

Amplification measurement was accomplished using a dual labeled minor groove binder (MGB) probe with a 5'-bound fluorescent dye reporter [6-carboxyfluorescein (FAM)] and a 3'-bound non-fluorescent quencher. While both are bound to the probe the fluorescent signature of the reporter is quenched. Taq DNA polymerase 5'-3' exonuclease activity during the extension phase of PCR causes cleavage of the probe, releasing the FAM reporter [22, 23]. Once separated from the quencher the reporter will emit its fluorescence [19, 23]. This fluorescent emission is detected by the Quantstudio™ 7 Flex system and analyzed by Quantstudio™ Real-Time PCR software (v 1.1) (Applied Biosystems, Waltham, MA, USA).

Calibration standards for the *DYS14* assay were produced diluting known male DNA in female DNA to a total concentration of 66ng/mL, with subsequent 10-fold dilutions ranging from 66ng to 0.0066ng male genome (mean r^2 = 0.994, n = 11). Calibration standards for the β-globin assay were produced by ten-fold serial dilution of DNA, ranging from 500ng – 0.5ng (mean r^2 = 0.997, n = 11). All standards were measured in triplicate on each reaction plate simultaneously with subject reactions. Also included on each reaction plate were positive and negative control subjects for male genome diluted in female DNA and a no template control (NTC) for each assay. The negative controls and NTC tested consistently negative during all experiments.

Performance of the assay was measured for reproducibility and linearity by using female samples spiked with male DNA. A series of 10-fold dilutions of male DNA in female DNA demonstrated target specificity of the *DYS14* assay. The measured male GEq of our validation controls demonstrated excellent linear regression with a correlation coefficient of 0.998. Reproducibility was measured using three samples with mean male GEqs per million of 1594.64 ± 137.41 (CV = 8.62), 10.50 ± 2.31 (CV=22.02), and 6.72 ± 1.82 (CV=27.14).

The qPCR reactions were set up in a 10uL reaction on 384-well reaction plates. The target genome input of blood-extracted DNA was 66 ng per reaction, thus achieving 10 000 genomes per reaction. Each reaction also included 5.0 mL TaqMan® Fast Advanced Master Mix (Applied Biosystems, Waltham, MA, USA), 0.5 mL 20X custom assay (Applied Biosystems, Waltham, MA, USA) including a dual labeled probe as well as forward and reverse primers. Each sample simultaneously tested 12 aliquots for *DYS14* and in duplicate for *β-globin* on the same reaction plate. The amplification

thermal cycling initiated with 50.0°C for 2 min then denaturation at 95.0°C for 2 min then 46 cycles of 95.0°C for 15 s and 56.0°C for 1 min.

All qPCR reactions and analysis were performed by a scientist who was blinded to MRKH status. Results were reported as male GEq per 1 million cells, calculated by using the mean measure of *DYS14* and *β-globin* for each subject. Male microchimerism was defined as a quantifiable measure of *DYS14* in any of 12 aliquots tested per sample.

Precautions against contamination

We utilized strict precautions while handling these samples to prevent contamination. Standard precautions for PCR reaction set up were followed [24]. Separate pre-amplification and post-amplification areas were strictly adhered to. All handling of samples was done inside a class II biosafety cabinet which pulls contaminated air from the work surface and exterior environment through a high efficiency particulate air (HEPA) filter to sterilize before returning to the work surface. All pipetting was carried out using aerosol-resistant filtered pipette tips. Each reaction plate included multiple NTC wells for each assay. After each use and preparation of each plate, the biosafety cabinet was thoroughly cleaned using a system of 10% bleach, cleaned with deionized water, followed by DNAZap™ (Invitrogen, Waltham, MA, USA) as per manufacturer's directions, two cycles of deionized water cleaning and 1 h of an ultraviolet lamp.

Statistical analysis

For sample size calculation, we assumed 25% microchimerism in the control group [according to results in previous studies [25]], and we assumed that in at least 50% of the adult MRKH women male microchimerism would still be detectable in the blood. With a significance level (alpha) of 5% and a power of 95%, we calculated a sample size of 91 women in both groups. Anticipating a 5% margin of error, in total 192 women needed to be included.

Not normally distributed variables were compared using the Mann-Whitney U test. For categorical variables the chi-squared (χ^2) or Fisher's exact test was used as appropriate. We used logistic regression to examine the association between microchimerism and MRKH syndrome by calculating the odds ratio. Age (years) was included as covariate as it may affect the presence of microchimerism. $P < 0.05$ was considered statistically significant. The statistical analysis was performed using SPSS 22.0 (SPSS, Chicago, IL, USA).

Ethics

The study protocol has been approved by the Institutional Review Board of the VU Medical Center, Amsterdam, the Netherlands, on 5 January 2017 (METC VUmc 2016.374). The trial was registered in the Dutch National Trial Registry (trial registration number NTR5961).

RESULTS

For this study, 96 women were included in the MRKH group and 100 control women were selected who reported to have never been pregnant. After collection of all available information, we excluded one woman in the patient group because, in retrospect, she did not have the MRKH syndrome. In the control group, one subject was excluded due to the limited amount of extracted DNA available for the microchimerism analysis. The final analysis included 194 women.

Patient characteristics

All women in the patient group reported having been diagnosed with the MRKH syndrome, with a mean time since diagnosis of 25.5 years (range 2 - 64 years). The mean age at the time of diagnosis was 16.5 years (range 6-26 years). In five women (5.3%), confirmation of the MRKH diagnosis by medical records could not be achieved. 54 women (56.8%) were identified as having typical MRKH syndrome and 41 women (43.2%) were identified as having atypical MRKH syndrome. For the women with atypical MRKH syndrome 28 out of 41 (53.7%) had a renal malformation, 18 (24.4%) had a skeletal malformation, and 9 (22.0%) had combined malformations (Table I). Of the participants in the patient group, three (3.2%) reported a positive family history for MRKH syndrome: two sisters were both participants in this study, and one woman reported an affected cousin (daughter of sister of father of participant). Two women with MRKH syndrome were part of dizygotic twin pairs, one with a twin brother and one with a twin sister, the latter twin sister being unaffected for the MRKH syndrome. All pregnancies resulting in the birth of a woman with MRKH syndrome were conceived naturally. The mean age of the mother at the time of birth of the women with MRKH syndrome was 28.7 years (range 18-42). All women in the control group reported having reached menarche and had never been pregnant. No women in the control group were part of a twin pair.

Table I - *Reported malformations in 41 women with atypical Mayer-Rokitansky-Küster-Hauser syndrome*

Renal malformations (28/41)*	(n)
Unilateral renal agenesis	15
Pelvic kidney	3
Hypoplastic kidney	3
Horse shoe kidney	2
Abnormal position	2
Malrotation	1
Cirrhotic kidney	1
Duplex kidney	1
Skeletal (18/41)*	
Scoliosis	13
Scoliosis and fusion of vertebrae	1
Cervical ribs	1
Arm agenesis	1
KFS, SD and scoliosis	1
Spina bifida, KFS, SD and scoliosis	1
Other malformations (4/41)*	
Hearing loss	2
Scaphoid hypoplasia and hearing loss	1
Clubfoot	1

** Combined malformations in nine women. KFS Klippel-Feil syndrome (cervical vertebral fusion); SD, Sprengel's deformity (high scapula).*

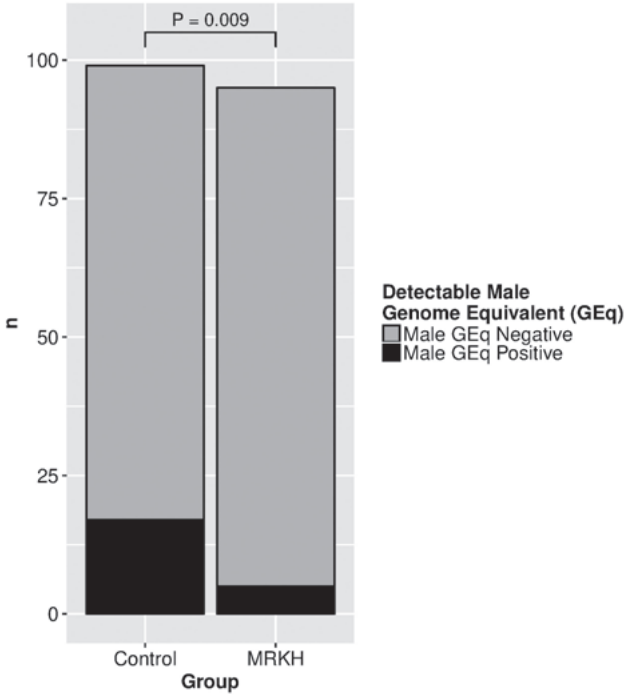


Figure 1: *Number of subjects with male microchimerism in blood in control women and women with Mayer– Rokitansky–Küster–Hauser syndrome. MRKH, Mayer-Rokitansky-Küster-Hauser; statistical test used: chi-squared test.*

Male microchimerism

In total, 194 women were included in the analyses concerning male microchimerism (Table II). In the total group, 22 out of 194 (11.3%) women tested positive for male microchimerism. The prevalence of male microchimerism was significantly higher in the control group than in the MRKH group (17.2% versus 5.3%, P = 0.009) (Fig. 1). After correcting for age, women in the control group were 5.8 times more likely to have male microchimerism (OR 5.84 (CI 1.59 – 21.47), P = 0.008). The mean concentration of male microchimerism in the positive samples was 56.0 male GEq per 1 000 000 cells (Fig. 2). The mean concentration of male microchimerism was significantly higher in the control group than in the MRKH group (P = 0.007). The prevalence of male microchimerism was similar in women with typical MRKH syndrome and atypical MRKH syndrome (5.6% versus 4.9%, P > 0.999).

To account for a possible older brother as a source for chimerism [26], we made a subgroup analysis. From the total group, 134 women (69.1%) had no older brother (Table II). Also in this subgroup, the prevalence and concentration of male microchimerism were higher in the controls than in the MRKH group.

Table II. *Male microchimerism in study subjects.*

Total group (n=194)	**MRKH (n= 95)**	**Control (n= 99)**	**P-value**
Age (years)	40.9 (15.5)	30.2 (6.6)	-
BMI (kg/m^2)	24.3 (5.3)	22.8 (3.5)	-
Presence of male microchimerism	5 (5.3%)	17 (17.2%)	0.009 [a]
	OR[c]: 0.27 (CI 0.10 – 0.76)		
	OR[d]: 0.17 (CI 0.05 – 0.63)		
Concentration male microchimerism (GEq*)			
- Positive samples (n=22)	2.3 (1.5)	71.8 (127.0)	0.055 [b]
- Total group	0.1 (0.6)	12.3 (58.1)	0.007 [b]
Selection of study subjects with no older brother (n = 134)	**MRKH (n=63)**	**control (n=71)**	**P-value**
Age (years)	38.1 (12.9)	30.5 (6.5)	
Presence of male microchimerism	3 (4.8%)	10 (14.1%)	0.069 [a]
	OR[c]: 0.31 (CI 0.08 – 1.16)		
	OR[d]: 0.18 (CI 0.03 – 0.90)		
Concentration male microchimerism (GEq*)			
- Positive samples (n=13)	1.3 (0.3)	27.3 (67.1)	0.018 [b]
- Total group	0.1 (0.3)	3.9 (25.9)	0.053 [b]

Data are mean (SD) or n (%). MRKH: Mayer-Rokitansky-Küster-Hauser syndrome.
**Male genome equivalents per 1,000,000 female cells [a] Chi square test [b] Mann-Whitney U test [c] Unadjusted odds ratio (95% confidence interval), logistic regression [d]Odds ratio (95% confidence interval), adjusted for age*

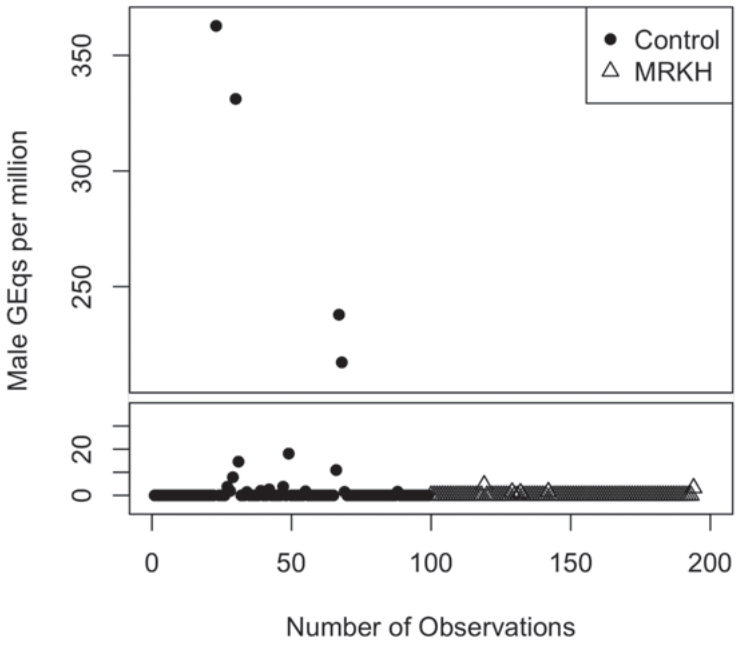


Figure 2: *Calculated male genome equivalents (GEq) per 1 million cells in blood, for control women and women with MRKH syndrome.*

Table III presents possible sources for male microchimerism. No significant difference was found for the prevalence of male microchimerism in the group with an older brother as opposed to the group without (P = 0.263). In the MRKH group, five women reported having received a blood transfusion: these women all tested negative for male microchimerism. The prevalence of male microchimerism was similar when comparing the group with and without blood transfusion (P > 0.999). In the control group, eight women reported never having had sexual intercourse; of these women, one tested positive for male microchimerism. The prevalence of male microchimerism was similar when comparing the group with and without sexual intercourse (P > 0.999).

Table III. *Identifying additional sources for male microchimerism*

	Male microchimerism positive (n=22)	Male microchimerism negative (n=172)	P-value
Age, years	35.3 (11.6)[+]	35.4 (13.1)	-
Study group			
- MRKH	5 (22.7%)	90 (52.3%)	0.009[a]
- Control	17 (77.3%)	82 (47.7%)	
Older brother			
- yes	9 (40.9%)	50 (29.1%)	0.263[a]
- no	13 (59.1%)	121 (70.3%)[#]	
Blood transfusion *			
- yes	0	5 (5.6%)	>0.999[b]
- no	5 (100%)	85 (94.4%)	
Sexual intercourse **			
- yes	16 (94.2%)	75 (91.5%)	>0.999[b]
- no	1 (5.9%)	7 (8.5%)	

Data are mean (SD) or n (%).
**Data missing for control women, this question was only asked in questionnaire for MRKH women; ** Data missing for MRKH women, this question was only asked in questionnaire for control women; [+]Mean (SD) age in women with MRKH syndrome who tested positive for microchimerism = 49.9 (10.9) years, age in control women = 31.0 (7.9) years; [#]One MRKH woman with a twin brother tested negative for male microchimerism. Statistical analyses: [a]Chi square test; [b]Fisher's exact test*

DISCUSSION

MRKH is a condition that has been well characterized yet the etiology has remained elusive [27]). Research of the phenotype analogous condition in cattle, freemartinism, has yielded the hypothesis that uterine development is inhibited by exposure to AMH *in utero* [9]. In cattle, the male co-twin has been identified as the source for AMH, which is transferred via placental blood exchange. In humans, a similar exchange of cellular material has been documented between twins by detection of blood chimerism [11, 12]. We sought to explain fetal exposure to blood and AMH from a male (vanished) co-twin in women with MRKH by the measure of male blood microchimerism.

The use of a male genome target has been previously implemented in other studies of chimerism including fetal microchimerism [19]. However, to our knowledge, there are no studies to date that have explored the presence of male microchimerism in women with MRKH. Our approach utilized a qPCR technique that was able to demonstrate exceptional target specificity and sensitivity for male genome. In this study of 194 women, 5 of 95 women (5.3%) with MRKH demonstrated male

microchimerism while the control group demonstrated male microchimerism in 17 of 99 (17.2%). This difference shows a significantly lower frequency in detection of male microchimerism for women with MRKH (P=0.009). Similarly, when subjects with older brothers are removed the percentage of male microchimerism in controls and cases are 14.1% and 4.8%, respectively. Our results do not support increased male microchimerism in adult women with MRKH as a consequence of intrauterine blood exchange.

Although these results contradict our underlying hypothesis of inadvertent AMH exposure in early pregnancy as the pathophysiological mechanism, it is too early to definitively conclude that occurrence of AMH exchange during early embryologic development in MRKH was not involved. There are several limitations to our approach that must be considered. First, this method relies on the detection of male genome in these subjects and is therefore contingent on the occurrence of cell grafting as a consequence of this exchange. There are several factors that are essential to effective grafting of cellular material in a host to evade immune targeting and thrive [28]. It is possible for intrauterine cellular exchange to occur passively between twins and not result in a persistent chimerism. Similarly, our subject population includes all adult women and therefore is reliant on long term prevalence of the male (XY) cell population to be detected at this stage of life. These XY cells may be eliminated or never graft in the host body, thereby shortening the timespan where detection is achievable. This technique exclusively examines the blood-derived genome, resulting in a limited scope of microchimerism detection and is not a global representation of chimerism throughout the individual. Owing to the proposed transfusion mechanism, it is most plausible that a male cell population may be localized in the blood; however, the cells may have grafted anywhere in the body during development [29]. It must also be considered that the exchange of AMH may occur without the transfusion of XY cells and therefore cannot be discovered by chimerism detection.

Our finding of male microchimerism in a larger percentage of women (17.2%) in the control group without a history of pregnancy is in line with findings from the literature. Prior studies demonstrated male microchimerism being present in 13.6% of adolescent girls [30] and 13.3% of healthy null gravid women [25]. Potential sources for the male cells are considered to be transfusion [31], older brothers (or discontinued male pregnancies from their mother) [26], unrecognized male miscarriages [25],

vanishing twins [14], or possibly sexual intercourse without pregnancy [30]. The significant difference in our study, in favour of the control group, suggests that a substantial proportion of the microchimerism could be explained by unrecognized pregnancies or the harbouring of microchimeric cells after sexual intercourse. Moreover, the higher concentration in the control group could reflect a larger transfusion of microchimeric cells by these sources. In our study, there were no differences in occurrence of older brothers, previous blood transfusion or history of sexual intercourse between women with or without microchimerism, both in participants and controls. We are unable to rule out the potential for microchimerism resulting from an unrecognized pregnancy or a vanishing twin.

The findings in women with MRKH illustrate that this population has a decreased prevalence of male microchimerism relative to others. This significant difference demonstrates that women with MRKH may serve as a suitable control group in chimerism research, in part due to the certainty that they have no history of pregnancy. Still, three women with MRKH and no older brother had detectable microchimerism which may have been obtained through several mechanisms including a vanishing twin or miscarriage in their mother.

Considering the several potential sources of XY cells, our results from the control population demonstrate that there may be a larger prevalence of microchimerism in the general population which requires further investigation. In general, there is progressive interest in the clinical consequences of microchimerism [32, 33]. It has been associated with various autoimmune disorders [34], but possible beneficial consequences have also been described. Microchimeric cells could replace injured cells in diseased tissues [32]. Recently it has been suggested that naturally acquired microchimerism has 'protective effects in promoting success of future pregnancies' [35].

In summary, we hypothesized that women with MRKH syndrome had an unrecognized male co-twin that was lost early in gestation (a vanishing twin) and have investigated possible fetal exposure to male blood by measurement of male microchimerism. We are led to the conclusion that women with MRKH syndrome are not evident (micro)chimeras. However, we are not able to draw definitive conclusions regarding the occurrence of AMH exchange. Further research should involve the presence of MRKH syndrome in a large cohort of girls with a twin brother and in girls born after a pregnancy complicated by vanishing twin syndrome. Additional

research on twin chimerism is needed to study the true prevalence of intrauterine blood exchange between dizygotic twins.

AUTHORS' ROLES

Study design, study organization, data collection, data analysis and manuscript draft was performed by H.P, supervised by C.L.; Analysis of male microchimerism and manuscript draft was performed by B.J., supervised by E.E and G.D.; D.M. organized DNA isolation and data analysis; A.O. M.B., E.D.B, F.L. collected data of control group; Supervision of study design and manuscript draft were performed by J.D., V.M., M.V., D.B.; All authors had read the manuscript before submission.

ACKNOWLEDGEMENTS

We thank all participating women and the patients' association of women with MRKH ('Stichting MRK-vrouwen'). We thank T.J. Korsen (research nurse) and R.J. van Schooten (Genome Diagnostics VUmc) for their contributions in study organization.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

1. Herlin, M., et al., *Prevalence and patient characteristics of Mayer-Rokitansky-Kuster-Hauser syndrome: a nationwide registry-based study.* Hum Reprod, 2016. **31**(10): p. 2384-90.
2. Choussein, S., et al., *Mullerian dysgenesis: a critical review of the literature.* Arch Gynecol Obstet, 2017. **295**(6): p. 1369-1381.
3. Rall, K., et al., *Typical and Atypical Associated Findings in a Group of 346 Patients with Mayer-Rokitansky-Kuester-Hauser Syndrome.* J Pediatr Adolesc Gynecol, 2015. **28**(5): p. 362-8.
4. Oppelt, P., et al., *Clinical aspects of Mayer-Rokitansky-Kuester-Hauser syndrome: recommendations for clinical diagnosis and staging.* Hum Reprod, 2006. **21**(3): p. 792-7.
5. Patnaik, S.S., et al., *Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome: a historical perspective.* Gene, 2015. **555**(1): p. 33-40.
6. Taguchi, O., et al., *Timing and Irreversibility of Mullerian Duct Inhibition in the Embryonic Reproductive-Tract of the Human Male.* Developmental Biology, 1984. **106**(2): p. 394-398.
7. Resendes, B.L., et al., *Role for anti-Mullerian hormone in congenital absence of the uterus and vagina.* Am J Med Genet, 2001. **98**(2): p. 129-36.
8. Oppelt, P., et al., *DNA sequence variations of the entire anti-Mullerian hormone (AMH) gene promoter and AMH protein expression in patients with the Mayer-Rokitanski-Kuster-Hauser syndrome.* Hum Reprod, 2005. **20**(1): p. 149-57.
9. Vigier, B., et al., *Origin of anti-Mullerian hormone in bovine freemartin fetuses.* J Reprod Fertil, 1984. **70**(2): p. 473-9.
10. Lillie, F.R., *The Theory of the Free-Martin.* Science, 1916. **43**(1113): p. 611-3.
11. van Dijk, B.A., D.I. Boomsma, and A.J. de Man, *Blood group chimerism in human multiple births is not rare.* Am J Med Genet, 1996. **61**(3): p. 264-8.
12. Peters, H.E., et al., *Unusual Twinning Resulting in Chimerism: A Systematic Review on Monochorionic Dizygotic Twins.* Twin Res Hum Genet, 2017. **20**(2): p. 161-168.
13. Bogdanova, N., et al., *Blood chimerism in a girl with Down syndrome and possible freemartin effect leading to aplasia of the Mullerian derivatives.* Hum Reprod, 2010. **25**(5): p. 1339-43.
14. de Bellefon, L.M., et al., *Cells from a vanished twin as a source of microchimerism 40 years later.* Chimerism, 2010. **1**(2): p. 56-60.
15. Lo, Y.M., et al., *Two-way cell traffic between mother and fetus: biologic and clinical implications.* Blood, 1996. **88**(11): p. 4390-5.
16. Lee, T.H., et al., *Survival of donor leukocyte subpopulations in immunocompetent transfusion recipients: frequent long-term microchimerism in severe trauma patients.* Blood, 1999. **93**(9): p. 3127-39.
17. van den Berg, M.H., et al., *Long-term effects of childhood cancer treatment on hormonal and ultrasound markers of ovarian reserve.* Hum Reprod, 2018.
18. Overbeek, A., et al., *A nationwide study on reproductive function, ovarian reserve, and risk of premature menopause in female survivors of childhood cancer: design and methodological challenges.* BMC Cancer, 2012. **12**: p. 363.

19. Lambert, N.C., et al., *Male microchimerism in healthy women and women with scleroderma: cells or circulating DNA? A quantitative answer.* Blood, 2002. **100**(8): p. 2845-51.

20. Lo, Y.M., et al., *Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis.* Am J Hum Genet, 1998. **62**(4): p. 768-75.

21. Saiki, R.K., et al., *Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase.* Science, 1988. **239**(4839): p. 487-91.

22. Holland, P.M., et al., *Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase.* Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7276-80.

23. Heid, C.A., et al., *Real time quantitative PCR.* Genome Res, 1996. **6**(10): p. 986-94.

24. Kwok, S. and R. Higuchi, *Avoiding false positives with PCR.* Nature, 1989. **339**(6221): p. 237-8.

25. Yan, Z., et al., *Male microchimerism in women without sons: quantitative assessment and correlation with pregnancy history.* Am J Med, 2005. **118**(8): p. 899-906.

26. Dierselhuis, M.P., et al., *Transmaternal cell flow leads to antigen-experienced cord blood.* Blood, 2012. **120**(3): p. 505-10.

27. Fontana, L., et al., *Genetics of Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome.* Clin Genet, 2017. **91**(2): p. 233-246.

28. Ichinohe, T., *Long-term feto-maternal microchimerism revisited: Microchimerism and tolerance in hematopoietic stem cell transplantation.* Chimerism, 2010. **1**(1): p. 39-43.

29. Korbling, M., et al., *Hepatocytes and epithelial cells of donor origin in recipients of peripheral-blood stem cells.* N Engl J Med, 2002. **346**(10): p. 738-46.

30. Muller, A.C., et al., *Microchimerism of male origin in a cohort of Danish girls.* Chimerism, 2015. **6**(4): p. 65-71.

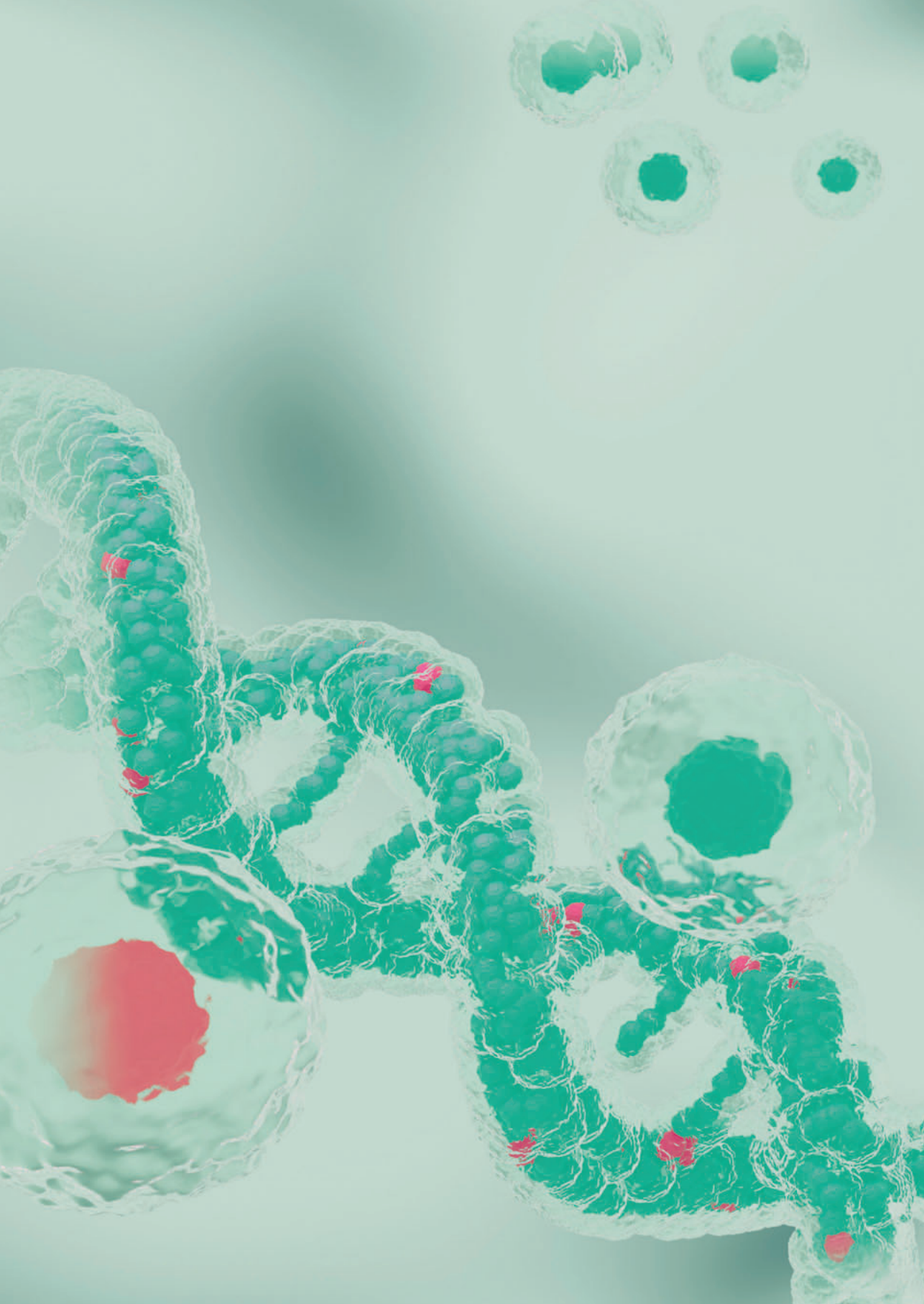
31. Utter, G.H., et al., *Transfusion-associated microchimerism.* Vox Sang, 2007. **93**(3): p. 188-95.

32. Kinder, J.M., et al., *Immunological implications of pregnancy-induced microchimerism.* Nat Rev Immunol, 2017. **17**(8): p. 483-494.

33. Gammill, H.S. and W.E. Harrington, *Microchimerism: Defining and redefining the prepregnancy context - A review.* Placenta, 2017. **60**: p. 130-133.

34. Nelson, J.L., *The otherness of self: microchimerism in health and disease.* Trends Immunol, 2012. **33**(8): p. 421-7.

35. Kinder, J.M., et al., *Cross-Generational Reproductive Fitness Enforced by Microchimeric Maternal Cells.* Cell, 2015. **162**(3): p. 505-15.



4

MALE MICROCHIMERISM IN FEMALES: A QUANTITATIVE STUDY OF TWIN PEDIGREES TO INVESTIGATE MECHANISMS

Published as:

Johnson, B. N.†, Peters, H. E.†, Lambalk, C. B., Dolan, C. V., Willemsen, G., Ligthart, L., Mijatovic, V., Hottenga, J. J., Ehli, E. A., & Boomsma, D. I. (2021).

Male microchimerism in females: a quantitative study of twin pedigrees to investigate mechanisms. Hum Reprod, 36(9), 2529-2537.

https://doi.org/10.1093/humrep/deab170

†The authors consider that the first two authors should be regarded as joint first authors.

ABSTRACT

Study question: Does having a male co-twin, older brothers, or sons lead to an increased probability of persistent male microchimerism in females members of twin pedigrees?

Summary answer: The presence of a male co-twin did not increase risk of male microchimerism and the prevalence of male microchimerism was not explained by having male offspring or by having an older brother.

What is known already: Microchimerism describes the presence of cells within an organism that originate from another zygote and is commonly described as resulting from pregnancy in placental mammals. It is associated with diseases with a female predilection including autoimmune diseases and pregnancy-related complications. However, microchimerism also occurs in nulliparous women; signifying gaps in the understanding of risk factors contributing to persistent microchimerism and the origin of the minor cell population.

Study design, size, duration: This cross-sectional study comprised of 446 adult female participants of the Netherlands Twin Register (NTR).

Participants/materials, setting, methods: Participants included in the study were female monozygotic (MZ) twins, female dizygotic same-sex twins and females of dizygotic opposite-sex twin pairs, along with the mothers and sisters of these twins. Peripheral blood samples collected from adult female participants underwent DNA extraction and were biobanked prior to the study. To detect the presence of male-origin microchimerism, DNA samples were tested for the relative quantity of male specific Y chromosome gene *DYS14* compared to the common *β-globin* gene using a highly sensitive quantitative PCR assay.

Main results and the role of chance: We observed a large number of women (26.9%) having detectable male microchimerism in their peripheral blood samples. The presence of a male co-twin did not increase risk of male microchimerism (odds ratio (OR) = 1.23: SE 0.40, P = 0.61) and the prevalence of male microchimerism was not explained by having male offspring (OR 0.90: SE 0.19, P = 0.63) or by having an older brother (OR = 1.46: SE 0.32, P = 0.09). The resemblance (correlation) for the presence of microchimerism was similar (P = 0.66) in monozygotic pairs (0.27; SE 0.37) and in first-degree relatives (0.091; SE 0.092). However, age had a positive relationship with the presence of male microchimerism (P = 0.02).

Limitations, reasons for caution: After stratifying for variables of interest, some participant groups resulted in a low numbers of subjects. We investigated microchimerism in peripheral blood due to the proposed mechanism of cell acquisition via transplacental blood exchange; however, this does not represent global chimerism in the individual and microchimerism may localize to numerous other tissues.

Wider implications of the findings: Immune regulation during pregnancy is known to mitigate allosensitization and support tolerance to non-inherited antigens found on donor cells. While unable to identify a specific source that promotes microchimerism prevalence within pedigrees, this study points to the underlying complexities of natural microchimerism in the general population. These findings support previous studies which have identified the presence of male microchimerism among women with no history of pregnancy, suggesting alternative sources of microchimerism. The association of detectable male microchimerism with age is suggestive of additional factors including time, molecular characteristics and environment playing a critical role in the prevalence of persistent microchimerism. The present study necessitates investigation into the molecular underpinnings of natural chimerism to provide insight into women's health, transplant medicine and immunology.

Study funding/competing interest(s): This work is funded by Royal Netherlands Academy of Science Professor Award (PAH/6635) to DIB; Genotype / phenotype database for behavior genetic and genetic epidemiological studies (ZonMw 911-09-032); Biobanking and Biomolecular Research Infrastructure, (BBMRI–NL, 184.021.007; 184.033.111); Netherlands Twin Registry Repository (NWO-Groot 480-15-001/674); The Rutgers University Cell and DNA Repository cooperative agreement (NIMH U24 MH068457-06); Grand Opportunity grants Integration of genomics and transcriptomics in normal twins and major depression (NIMH 1RC2MH089951-01); Developmental trajectories of psychopathology (NIMH 1RC2 MH089995); and European Science Council - Genetics of Mental Illness (ERC 230374).

CBL declares a competing interest as editor-in-chief of *Human Reproduction* and his department receives unrestricted research grants from Ferring, Merck and Guerbet. All remaining authors have no conflict-of-interest to declare in regards to this work.

Trial registration number: N/A

Keywords: Microchimerism, Chimerism, Twins, Female, Monozygotic Twins, Dizygotic Twins, Pedigree

INTRODUCTION

Microchimerism is defined as the presence of a small number of cells within an organism that originate from another genetically distinct zygote. There is an interest in microchimerism and its consequences for health, with a focus on bidirectional feto-maternal exchange of blood cells during pregnancy [1]. Such intrauterine blood exchange can result in long-term persistent microchimerism, suggested to be the result of grafting and proliferation of 'transfused' stem cells in the tissues of the recipient. Early exposure to foreign cells may teach the immune system to develop a tolerance for these cells, supporting long-term persistence of microchimerism [2, 3]. These examples of immune tolerance have exacerbated complications for the traditional self versus non-self criterion of immunology [4]. Exposure to foreign cells could contribute to human health and disease (for a review see Johnson et al., 2020 [1]), as a contributing factor in the pathogenesis of autoimmune diseases [5]. Interestingly, chimerism has also been associated with paternal care in marmosets, suggesting a role in behavior [6].

The presence of male cells in maternal circulation is indicative of microchimerism and has been well documented in women following pregnancy with a male fetus [7-9]. The method of delivery, presence of placental complications and hypertensive disorders all can influence the amount of cell trafficking and therefore the exposure to microchimerism [10, 11]. Foetal loss in early pregnancy may also result in microchimerism [12]. However, several studies have reported male microchimerism in women without any history of male pregnancy [13-15], leading to a search for natural sources which may result in male microchimerism, such as having older brothers, having a male co-twin, sexual intercourse or unrecognized pregnancies [16]. It has even been proposed that every human is born as a microchimera, with a yet undefined source of donor cells [17].

Intrauterine exchange of cells occurs between twins through foeto-maternal transfer or direct transfer between the twin fetuses via placental transfusion. For many years, twin chimerism was considered to be an exception in humans, with the 30-40 cases reported in literature discovered by coincidence, most commonly by blood typing discrepencies [18]. In 1996, a systematic search by Van Dijk et al. showed an 8% prevalence of blood group chimerism in a study of blood samples from 552 dizygotic twin pairs and 24 triplet sets for multiple red cell blood group antigens

[19]. This finding placed the concept of twin chimerism in humans in a new light. A recent study of 35 dizygotic twin pairs detected no ABO or D blood group chimerism through serologic assays and additional short tandem repeat (STR) assays also did not detect any blood chimerism [20]. Improvements in detecting microchimerism by advanced molecular techniques now allows quantitative real-time polymerase chain reaction (qPCR) to quantify chimerism prevalence in DNA samples [15].

The presence of non-inherited maternal antigens via maternal microchimerism in the offspring may be involved in the development of immune tolerance to both the maternal antigens and future overlapping fetal antigen from their own offspring [21]. With multiple associations between microchimerism and risk for disease [1], we designed a study to look into the etiology of microchimerism. DNA samples from twin pedigrees provide a unique opportunity to investigate multiple mechanisms for microchimerism, both within and across generations, providing an overview of presence of male microchimerism in women. Here we document the patterns and transmission of microchimerism in multi-generation families. We present rates of male microchimerism as quantified by qPCR in mono- and dizygotic female twins from same- and opposite-sex twin pairs, their singleton (non-twin) sisters and their mothers. The women come from a general population sample and are characterized for the presence of older brothers and male offspring. By studying male microchimerism in female twins and their relatives, we can investigate the effects of having a male co-twin, the prevalence of microchimerism in twins and their non-twin sisters, generation differences, and the degree of shared microchimerism among family members.

MATERIALS AND METHODS

Participants

The participants in this study are enrolled in the Netherlands Twin Register (NTR) Biobank [22, 23]. We included females and identified families with female monozygotic (MZ), female dizygotic same-sex (DZss) or dizygotic opposite-sex (DOS) twins with blood derived DNA samples. Twins were included if a DNA sample was also available from their mother and from at least one singleton sister. Due to the amount of genomic material needed, an additional inclusion criterion was the quantity of genomic material of the sample available in the biobank. The failure to meet this

criterion resulted in some incomplete pedigrees. The study included 446 women from 152 families: 62 females from DOS twin pairs, 80 females from MZ twin pairs, 68 females from DZss twin pairs, 106 mothers and 130 non-twin sisters. The median age in the total study population was 34 years (range 18-83), with the mothers of twins having a median age of 59 and their offspring of 32 years. Information on age, the presence of older brothers and the presence of a son was retrieved from the NTR database.

Quantitative real-time PCR

Blood derived DNA samples were tested for the presence of male microchimerism via a qPCR approach for measuring male microchimerism. This approach targets the Y-chromosome specific gene *DYS14* as previously described [15]. In brief, each sample was tested in parallel for *DYS14* in 12 replicates as a measure of male genome mass and for *β-globin* in duplicate to obtain a measure of total genome mass. The PCR cycling and fluorescent measurement was completed on a Quantstudio™ 7 Flex instrument and analyzed by Quantstudio™ Real-Time PCR software (v1.1) (Applied Biosystems, Waltham, MA, USA). Standards were produced using known male and female extracted DNA samples. The *β-globin* standard included 10-fold dilutions of extracted DNA to produce standards from 500 ng to 0.5 ng (Figure 1A). The standards for *DYS14* were produced by simulating sample conditions by diluting known male extracted DNA into female extracted DNA. The resulting standards maintained a constant 66 ng/mL with ten-fold dilutions of male genome from 66 ng/mL to 0.0066 ng/mL (Figure 1A). The lower detection limit of the qPCR assay implied a minimum threshold of detection for male microchimerism of one genome equivalent per one million.

A number of precautions were implemented to prevent potential contamination of the samples and qPCR reactions [15]. All samples from the NTR biobank and reaction preparation were handled in a Class II biosafety cabinet that was rigorously cleaned by chemical and ultraviolet decontamination. Additionally, all pipetting utilized previously established guidelines for PCR reaction preparation [24]. Each reaction plate contained various quality control measures. Six no template control (NTC) wells were included as a true negative for the *DYS14* and *β-globin* assays. Negative control was produced using a known female extracted DNA sample, whereas the positive control was produced by spiking known male extracted DNA sample into female

sample. The controls tested as expected for all experiments, such that negative control tested negative for *DYS14* (Figure 1B) while the NTC tested negative for both assays consistently. The positive control demonstrated consistent detection of both *β-globin* and *DYS14* across reaction plates (Figure 1C).

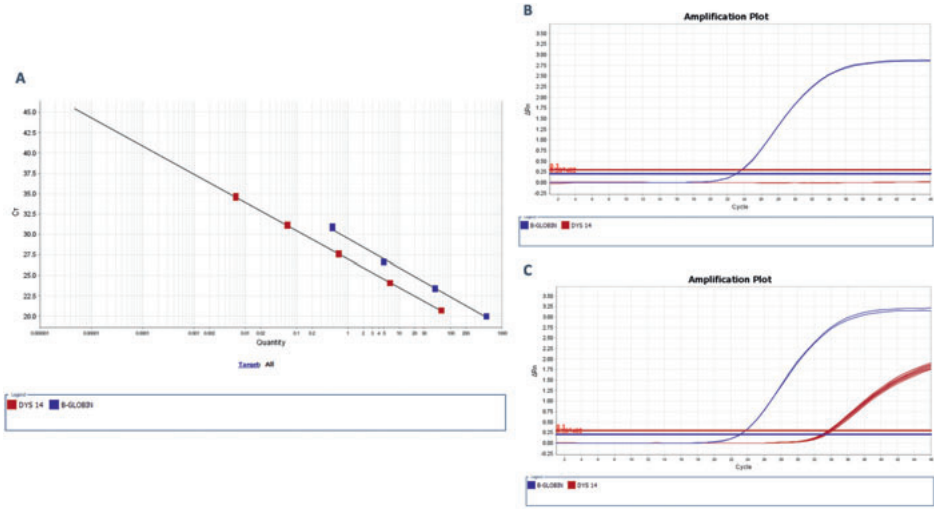


Figure 1: *qPCR data for the analysis of male genome equivalents by measure of DYS14 and β-globin targets. A) A plot of standard curves generated using known human genomic samples; B) An amplification plot for a negative control sample with measure of β-globin and no detection of DYS14. C) An amplification plot for a positive control sample with detectable measurement of both DYS14 and β-globin targets.*

Statistical Analyses

Participant group male microchimerism status presented as categorical data were compared by Chi-squared (c^2) tests. Associations of male microchimerism status with age and the presence of sons or older brothers were assessed by generalized estimating equations (GEE) and logistic regression with a correction for family clustering. Concordance for male microchimerism between pairs of relatives was summarized in 2 x 2 contingency tables. Quantitative data for male microchimerism burden were evaluated by a Kruskal-Wallis test. Data on presence of brothers (n = 18) or presence of son (n = 19) were missing in some families; for analyses concerning these variables, these families were excluded. Statistical analyses were performed using SPSS 26.0 (SPSS, Chicago, IL, USA) and R programming language. $P < 0.05$ was considered statistically significant.

To quantify familial resemblance for presence of male microchimerism, we estimated the tetrachoric correlations for 2 types of genetic relations: for MZ twin pairs, whose genetic relatedness is ~100% as they derive from the same fertilized egg, and for all others, who are first-degree relatives. All first-degree relatives share either exactly 50% of their segregating genes (mother and daughter), or 50% on average (DZ twin and sister pairs). Tetrachoric correlations represent the relation between variables on the underlying continuous liability scale. Tetrachoric correlations and their standard errors (SE) between family members [25] were estimated and tested for significance in OpenMx [26].

Ethics

The Netherlands Twin Register Biobank study was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam, an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance FWA00017598; IRB/ institute codes, NTR 03-180) and informed consent was obtained from all participants [23].

RESULTS

Prevalence

Male microchimerism was detected in 120 of the 446 participating women (26.9%). Prevalence was 16.3% in MZ females and did not differ from the prevalence in DZss twins (P = 0.27). In the group of females from DZss twin pairs 23.5% tested positive for male microchimerism, compared to 27.4% of females from DOS twin pairs (P = 0.61); OR 1.23 (SE 0.40). Of all 130 singleton sisters included, 33 tested positive for male microchimerism (25.4%). There were 106 mothers of twins, of whom 41 tested positive for male microchimerism (38.7%). Figure 2 summarizes these results for all members of each participant group.

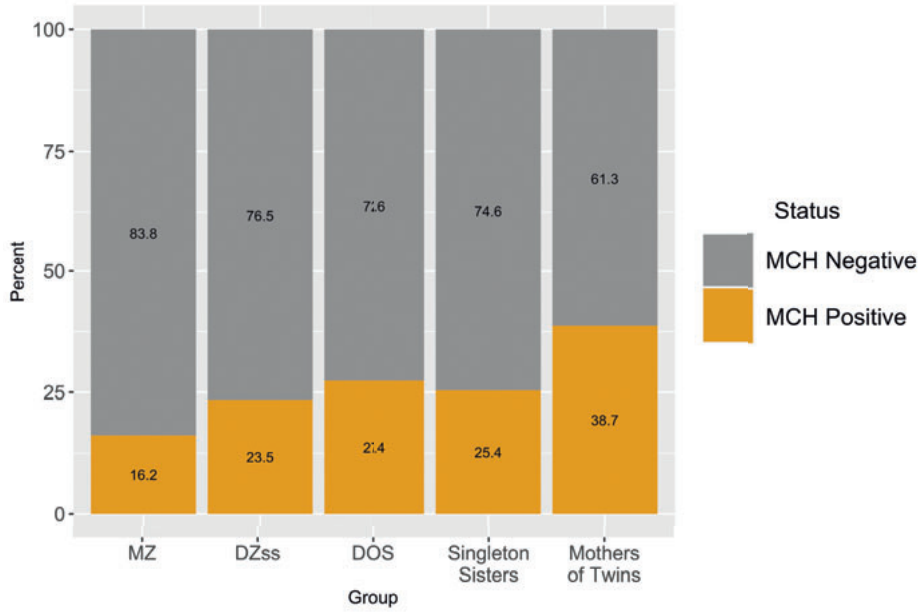


Figure 2: *Prevalence of male microchimerism (MCH) in women, stratified by pedigree position: MZ; monozygotic twins, DZss; dizygotic same-sex twins, DOS; dizygotic opposite-sex twins.*

Associations with microchimerism

The prevalence of detectable male microchimerism was compared among females who have an older brother, male offspring or both as shown in Figure 3. The prevalence of male microchimerism tended to be greater in females with an older brother (31.4%) compared to those without (24.0%); OR 1.46 (SE 0.32), P = 0.09. Females with and without male offspring had a similar prevalence of male microchimerism (26.0% and 28.0%, respectively); OR 0.90 (SE 0.19), P = 0.63. Further assessment of an additive effect of having both an older brother and a male offspring revealed no significant difference among all participants (P = 0.16) or in the mother of twins group (P = 0.50).

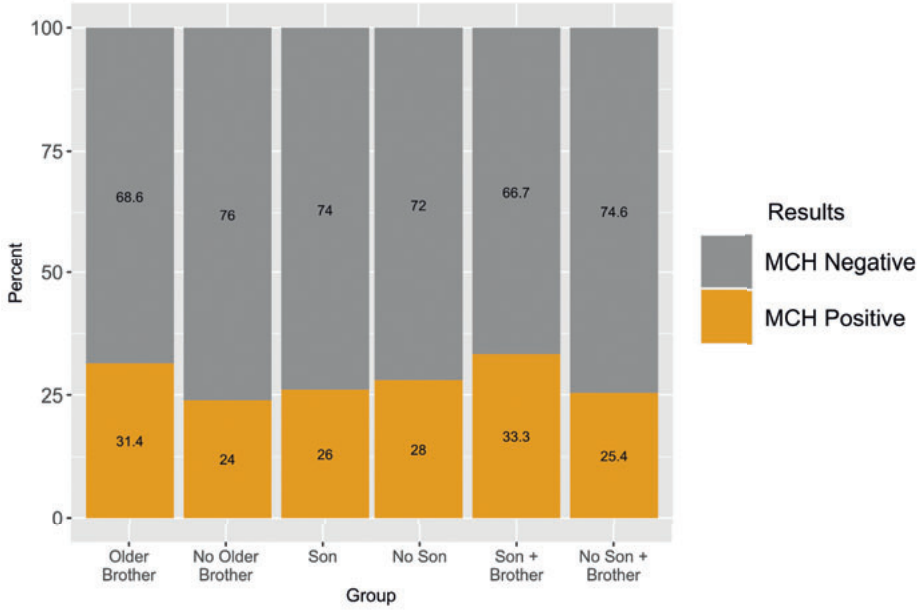


Figure 3: *Prevalence of male microchimerism (MCH) in females with and without an older brother, a son, or both an older brother and a son.*

The mothers of twins group, which showed the largest proportion of detectable male microchimerism, had similar prevalence with and without having male offspring (38.9% and 38.5%); OR 1.018 (SE 0.34), P = 0.96. There was a positive relationship between age at the time of biobanking and presence of microchimerism (P = 0.02; Nagelkerke R^2 = 0.017).

Figure 4 summarizes conconcordances for DZss and MZ twin, twin-sister, twin-mother and sister-mother pairs for male microchimerism. In the complete MZ twin pairs and DZss twin pairs, twin-twin comparison reveals that 75% and 61% were concordant for microchimerim status (P = 0.35). Male microchimerism was more prevalent in mothers (P = 0.003 for mother-twin and P = 0.06 for mother-non twin offspring comparison). There was no difference between the twins and their sisters for male microchimerism (P = 0.71).

Figure 4: *Male microchimerism concordance within families. Data are presented for pairs of relatives including: DZss (dizygotic same-sex) twins, MZ (monozygotic) twins, and mother with singleton sister and twin, and twin-sister.*

Familial resemblance

The offspring (data from twins and their sisters combined) of mothers with male microchimerism presented with male microchimerism somewhat more frequently than female offspring of male microchimerism negative mothers (26.8% and 17.6%, respectively; P = 0.08). We investigated the family resemblance by calculating tetrachoric correlations for microchimerism status separately in MZ twin pairs and in all first degree relatives. The correlations were estimated at 0.27 (SE 0.37) for MZ twin pairs and 0.091 (SE 0.092) for all first-degree relative pairs. The two correlations can equated to be the same (P = 0.66) and did not differ from zero (P = 0.25).

Microchimerism concentration

After selection of the positive samples (i.e. 120 subjects with >1 male GEq per 1 000 000 cells), the median concentration of male microchimerism is similar among twins, their sisters and their mothers (P = 0.28) and is summarized in Figure 5. Outlier samples with elevated levels of detectable male genome (>80 male GEq) were investigated compared to the other participants with detectable male microchimerism revealing

this group to have no male offspring or other distinguishing features and a lower age at time of collection (P = 0.034) (Figure 6).

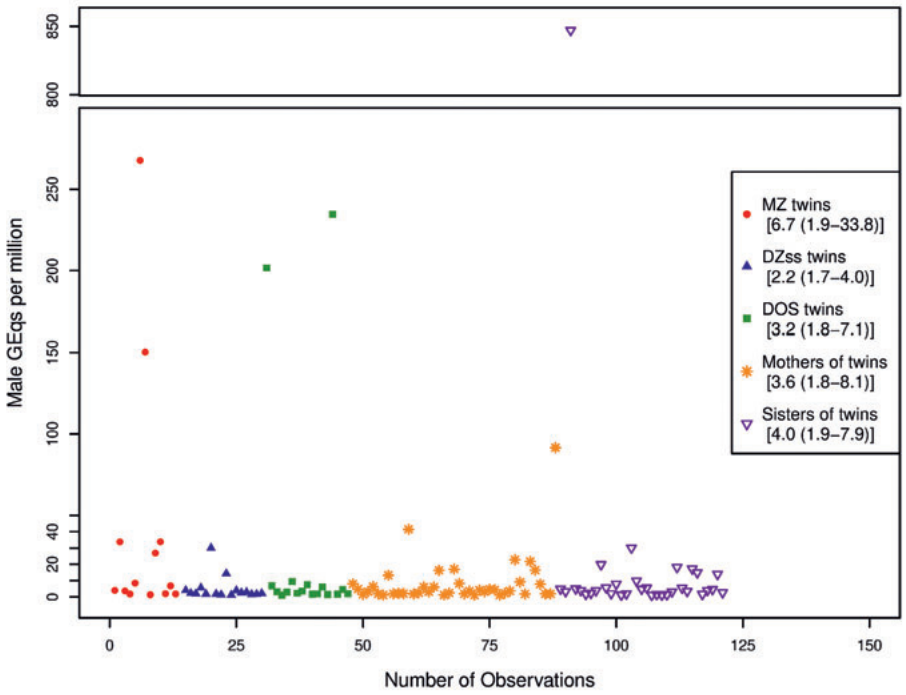


Figure 5: *Male genome equivalents (GEq) per one million cells by each participant with detectable male GEq. The data are stratified by participant group. Descriptive statistics presented as Median (IQR). MZ; monozygotic twins, DZss; dizygotic same-sex twins, DOS; dizygotic opposite-sex twins. Male GEqs; male genome equivalents per million.*

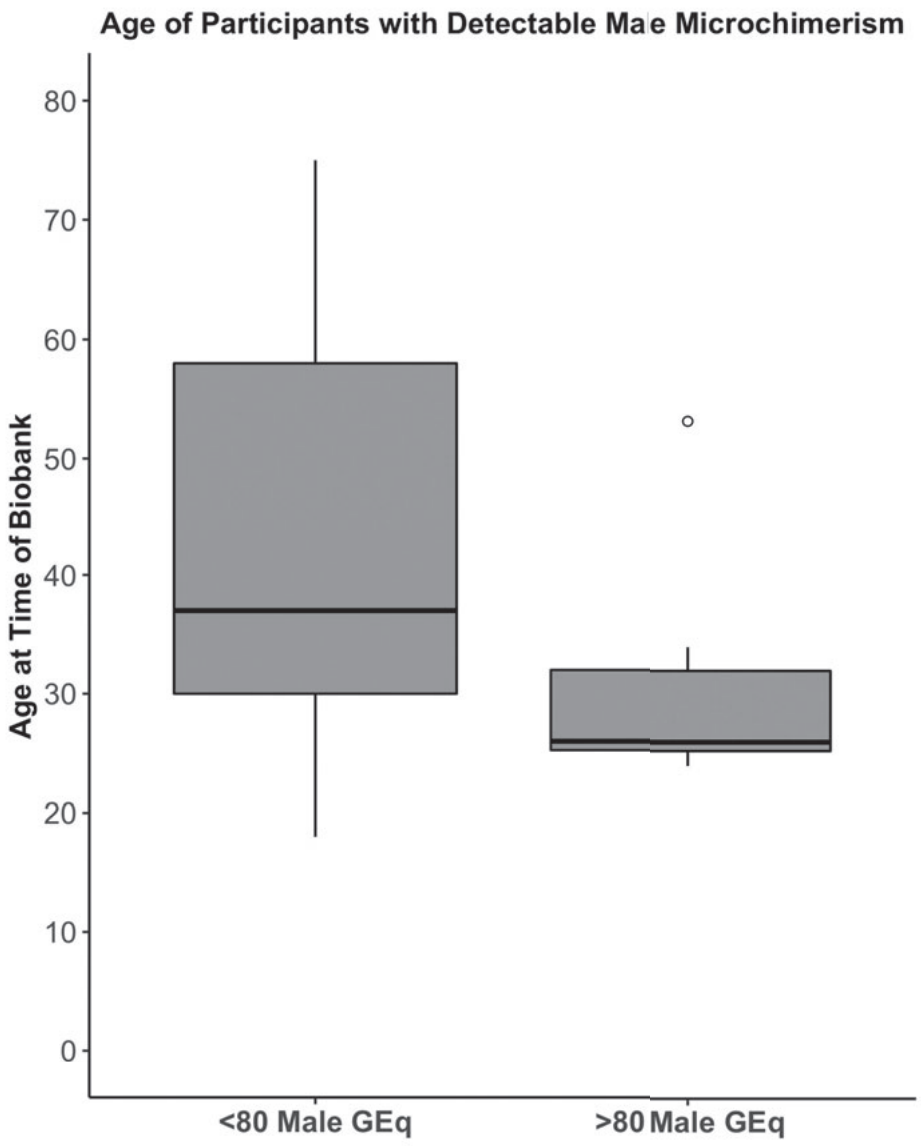


Figure 6: *Age distribution of individuals with detectable male microchimerism at a concentration of <80 male genome equivalents per million (GEq) (n = 114) or outliers with >80 male GEq (n = 6).*

DISCUSSION

We investigated the presence of male chimerism in women and its etiology in members of twin pedigrees. We found that about 27% of the adult women had detectable male microchimerism. The highest prevalence of male microchimerism was detected in the older participants, i.e. mothers of twins (38.7%) and we observed a positive relation between age and presence of microchimerism. This observation indicates that microchimeric cells may persist and stay detectable long-term after they are acquired, although increased exposure to unexplored variables due to age could also be involved in chimerism risk and persistence.

Previous research in twins looked at the presence of blood chimerism via red blood cell antigens and discovered an 8% chimerism prevalence in twins [19]. We expanded upon the previous work by investigating the source of male microchimerism and the familial relationships in two generation twin pedigrees. We found that females from a DOS twin pair do not present with male microchimerism more frequently than females from a DZss twin pair, despite the adjacent in utero presence of the male co-twin. This refutes our hypothesis that a male co-twin promotes persistent male microchimerism. Our data also showed that rates between DZ twins are not different from those among MZ twins and singleton siblings.

Male microchimerism could arise in women having an older brother, which has been suggested in earlier studies [13, 16]. This would support trans-maternal cell flow as an explanation and is supported by a higher prevalence of male microchimerism in female offspring of mothers with male microchimerism. We observed only a tendency of a higher prevalence of male microchimerism in women with an older brother. We also saw no evidence that male microchimerism status was related to the presence of a son. Further, among participants with an exceptionally high level of male microchimerism (>80 male GEq), none had a son at the time of sample collection. Thus, our study inidicates that the origin of the microchimeric cells may not necessarily be a close family member. One source is repeated sequential fetal-maternal exchanges across generations [27]. Possible further sources are unreported or unrecognized interrupted pregnancies [12], breastfeeding, placental structure, pregnancy complications including preeclampsia [10, 12, 28, 29] and it has also been suggested that sexual intercourse may play a role [13, 30]. We did not have these data available for a sufficient number of participants for comprehensive analyses of these altenatives.

There are several limitations that must be considered when interpreting our findings. First, despite obtaining samples from a large register, after stratifying for various analyses we had a relatively low number of subjects within some participant groups. Furthermore, the data do not illustrate the source of the minor cell population beyond that it is of male origin. Due to the suggested mechanism of chimerism acquisition via blood exchange we exclusively investigated blood microchimerism which may not represent global chimerism. Several studies have identified the presence of allogeneic cells in various tissues throughout the body; some of which include the skin [31], breast [32-35], pancreas [36] and brain [37, 38] among others. It is probable that cells obtained via blood exchange may localize to any variety of tissues in the host where they may possibly persist, the so-called 'adult stem cell plasticity phenomenon' [39, 40]. These findings help demonstrate that the temporality and localization of microchimerism in the blood and other tissues may be a factor of exposure, time and environmental influences such as tissue injury which may partially explain the present findings in adult participants.

As with many other studies of microchimerism, the use of a Y-chromosome target has been proven effective at identifying low levels of chimerism due to the target specificity. However this technique is limited to the study of females. While other targets, including RBC antigens and HLA typing have been developed, these have a lower sensitivity or increased complexity that often require prior knowledge of the donor's genotype [1]. As this work adds to the growing body of knowledge on human chimerism, continued advancements in molecular technologies will provide new opportunities to expand upon our findings presented here. We found that microchimerism occurs frequently, in over one quarter of women. This necessitates research into its etiology and significance.

Others have previously described that exposure to non-inherited maternal antigens improves success of future pregnancies as well as mitigating allosensitization [21]. As many changes occur during pregnancy, the complication of immune tolerance is balanced by interaction with non-inherited maternal antigens and is likely to provide insight into other areas of immune sensitization and allotransplantation [27]. Already, outcomes in cord blood transplantation have shown improvements when mismatched HLA alleles between donor and recipient include a non-inherited maternal antigen [41] and a beneficial effect against leukemia relapse when graft and donor share inherited paternal HLA antigens, suggesting an effect of maternal

microchimerism of T memory cells [42, 43]. Our findings provide further support for the remarkable complexity of microchimerism and the challenge it may pose the immune system both during and following pregnancy. Based on our findings of elevated microchimerism concentrations in younger participants (Figure 6) and a positive relationship between prevalence and age, it is possible that long term conservation of microchimerism is best maintained with low levels of allogeneic cells. These individual observations of persistent microchimerism may be due to successful evasion of overt sensitization of the immune system. Maintaining balance between microchimerism concentration and immune detection may be essential to long term microchimerism persistence.

Further, there is a growing body of research that has suggested microchimerism to be associated with the pathogenesis of disease. Recently, a study established that women without a protective HLA allele are more likely to develop rheumatoid arthritis when they have microchimerism carrying an HLA protective allele [44]. Similarly, maternal microchimerism has been identified to be associated with type 1 diabetes and contribute to islet b cells in the offspring [36]. Our findings provide further support for the hypothesis that microchimerism sensitizes the immune system. The inverse relationship between quantitative male microchimerism outliers and age (Figure 6) may illustrate the pressures of age and future pregnancy on maintenance of microchimerism, resulting in lower quantitative levels later in life. Possible mechanisms producing these findings could include immune system responses clearing allogeneic cells or an individual's own cells outcompeting the donor cells over multiple generations of cellular replication. Such studies and findings continue to amplify the need for researching microchimerism prevalence within the general population.

Improvements in future research will require expanding upon current knowledge of human chimerism with additional sample types to further explain chimerism tissue localization and subsequent health implications. Such work will likely necessitate larger consortia driven studies to achieve sample numbers to achieve significance in understanding the underlying complexity of human chimerism. Expanding on the findings presented here requires larger studies of microchimerism in mixed sex populations to evaluate alternative sources of microchimerism and contribution in health and disease. Further, due to the seemingly variable nature of microchimerism

over time, longitudinal studies of chimerism presence and concentration are warranted to further understand this phenomenon in human biology.

AUTHORS' ROLES

Study design, data collection, data analysis and interpretation, and drafting of manuscript were performed by B.N.J. and H.E.P., supervised by E.A.E. and C.B.L. respectively; supervision and contribution to study design by C.B.L., V.M., D.I.B. and E.A.E.; G.W. was responsible for the supervision of the NTR biobank project; assessment of participants and biobank sample selection by G.W. and L.L.; C.V.D. and J.J.H. contributed to data analysis and interpretation; founding of the Netherlands Twin Register (NTR), oversight of NTR participant recruitment and concomitant sample and phenotype collection by D.I.B.; all authors contributed to critical evaluation and approval of final manuscript.

FUNDING

This work is funded by Royal Netherlands Academy of Science Professor Award (PAH/6635) to DIB; Genotype / phenotype database for behavior genetic and genetic epidemiological studies (ZonMw 911-09-032); Biobanking and Biomolecular Research Infrastructure, (BBMRI–NL, 184.021.007; 184.033.111); Netherlands Twin Registry Repository (NWO-Groot 480-15-001/674); The Rutgers University Cell and DNA Repository cooperative agreement (NIMH U24 MH068457-06); Grand Opportunity grants Integration of genomics and transcriptomics in normal twins and major depression (NIMH 1RC2MH089951-01); Developmental trajectories of psychopathology (NIMH 1RC2 MH089995); and European Science Council - Genetics of Mental Illness (ERC 230374).

CONFLICT-OF-INTEREST

CBL declares a competing interest as editor-in-chief of *Human Reproduction* and his department receives unrestricted research grants from Ferring, Merck and Guerbet. All remaining authors have no conflict-of-interest to declare in regards to this work.

DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

1. Johnson, B.N., et al., *Chimerism in health and potential implications on behavior: A systematic review.* Am J Med Genet A, 2020.
2. Bianchi, D.W., et al., *Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum.* Proc Natl Acad Sci U S A, 1996. **93**(2): p. 705-8.
3. Gammill, H.S. and W.E. Harrington, *Microchimerism: Defining and redefining the prepregnancy context - A review.* Placenta, 2017. **60**: p. 130-133.
4. Pradeu, T. and E.D. Carosella, *On the definition of a criterion of immunogenicity.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17858-61.
5. Nelson, J.L., *The otherness of self: microchimerism in health and disease.* Trends Immunol, 2012. **33**(8): p. 421-7.
6. Ross, C.N., J.A. French, and G. Orti, *Germ-line chimerism and paternal care in marmosets (Callithrix kuhlii).* Proc Natl Acad Sci U S A, 2007. **104**(15): p. 6278-82.
7. Lo, Y.M., et al., *Two-way cell traffic between mother and fetus: biologic and clinical implications.* Blood, 1996. **88**(11): p. 4390-5.
8. Mosca, M., et al., *Correlations of Y chromosome microchimerism with disease activity in patients with SLE: analysis of preliminary data.* Ann Rheum Dis, 2003. **62**(7): p. 651-4.
9. Murata, H., H. Nakauchi, and T. Sumida, *Microchimerism in Japanese women patients with systemic sclerosis.* Lancet, 1999. **354**(9174): p. 220.
10. Gammill, H.S., et al., *Cellular fetal microchimerism in preeclampsia.* Hypertension, 2013. **62**(6): p. 1062-7.
11. Shree, R., et al., *Fetal microchimerism by mode of delivery: a prospective cohort study.* BJOG, 2019. **126**(1): p. 24-31.
12. Bianchi, D.W., et al., *Significant fetal-maternal hemorrhage after termination of pregnancy: implications for development of fetal cell microchimerism.* Am J Obstet Gynecol, 2001. **184**(4): p. 703-6.
13. Muller, A.C., et al., *Microchimerism of male origin in a cohort of Danish girls.* Chimerism, 2015. **6**(4): p. 65-71.
14. Yan, Z., et al., *Male microchimerism in women without sons: quantitative assessment and correlation with pregnancy history.* Am J Med, 2005. **118**(8): p. 899-906.
15. Peters, H.E., et al., *Low prevalence of male microchimerism in women with Mayer-Rokitansky-Kuster-Hauser syndrome.* Hum Reprod, 2019. **34**(6): p. 1117-1125.
16. Dierselhuis, M.P., et al., *Transmaternal cell flow leads to antigen-experienced cord blood.* Blood, 2012. **120**(3): p. 505-10.
17. Dierselhuis, M.P. and E. Goulmy, *We are all born as microchimera.* Chimerism, 2013. **4**(1): p. 18-9.
18. Tippett, P., *BLOOD-GROUP CHIMERAS - A REVIEW.* Vox Sanguinis, 1983. **44**(6): p. 333-359.
19. van Dijk, B.A., D.I. Boomsma, and A.J. de Man, *Blood group chimerism in human multiple births is not rare.* Am J Med Genet, 1996. **61**(3): p. 264-8.
20. Tavares, L., et al., *Blood chimerism in twins.* Immunohematology, 2018. **34**(4): p. 151-157.

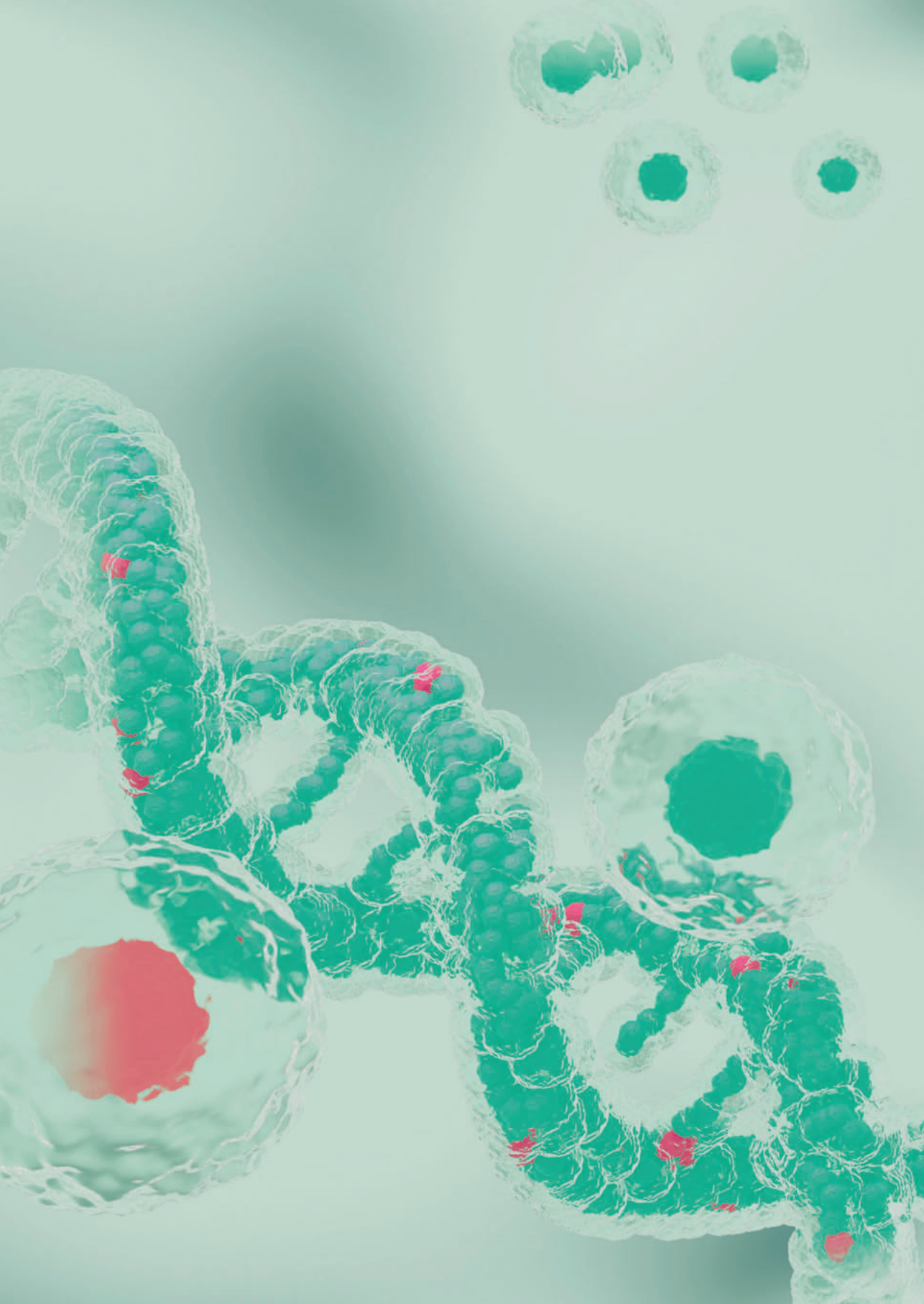
21. Kinder, J.M., et al., *Cross-Generational Reproductive Fitness Enforced by Microchimeric Maternal Cells.* Cell, 2015. **162**(3): p. 505-15.
22. Ligthart, L., et al., *The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs.* Twin Res Hum Genet, 2019: p. 1-14.
23. Willemsen, G., et al., *The Netherlands Twin Register biobank: a resource for genetic epidemiological studies.* Twin Res Hum Genet, 2010. **13**(3): p. 231-45.
24. Kwok, S. and R. Higuchi, *Avoiding false positives with PCR.* Nature, 1989. **339**(6221): p. 237-8.
25. Falconer, D. and T. Mackay, *Introduction to quantitative genetics. Essex*. 1996, UK: Pearson Educational Limited.
26. Neale, M.C., et al., *OpenMx 2.0: Extended Structural Equation and Statistical Modeling.* Psychometrika, 2016. **81**(2): p. 535-49.
27. Kinder, J.M., et al., *Immunological implications of pregnancy-induced microchimerism.* Nat Rev Immunol, 2017. **17**(8): p. 483-494.
28. Peters, H.E., et al., *Unusual Twinning Resulting in Chimerism: A Systematic Review on Monochorionic Dizygotic Twins.* Twin Res Hum Genet, 2017. **20**(2): p. 161-168.
29. Hassiotou, F. and D.T. Geddes, *Immune cell-mediated protection of the mammary gland and the infant during breastfeeding.* Adv Nutr, 2015. **6**(3): p. 267-75.
30. Yan, Z., et al., *Mate microchimerism in women without sons: Quantitative assessment and correlation with pregnancy history.* American Journal of Medicine, 2005. **118**(8): p. 899-906.
31. Mahmood, U. and K. O'Donoghue, *Microchimeric fetal cells play a role in maternal wound healing after pregnancy.* Chimerism, 2014. **5**(2): p. 40-52.
32. Dhimolea, E., et al., *High male chimerism in the female breast shows quantitative links with cancer.* Int J Cancer, 2013. **133**(4): p. 835-42.
33. Gadi, V.K., *Fetal microchimerism in breast from women with and without breast cancer.* Breast Cancer Res Treat, 2010. **121**(1): p. 241-4.
34. Kamper-Jorgensen, M., et al., *Opposite effects of microchimerism on breast and colon cancer.* Eur J Cancer, 2012. **48**(14): p. 2227-35.
35. Nemescu, D., et al., *Heterogeneous Distribution of Fetal Microchimerism in Local Breast Cancer Environment.* PLoS One, 2016. **11**(1): p. e0147675.
36. Nelson, J.L., et al., *Maternal microchimerism in peripheral blood in type 1 diabetes and pancreatic islet beta cell microchimerism.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1637-42.
37. Chan, W.F., et al., *Male microchimerism in the human female brain.* PLoS One, 2012. **7**(9): p. e45592.
38. Broestl, L., J.B. Rubin, and S. Dahiya, *Fetal microchimerism in human brain tumors.* Brain Pathol, 2018. **28**(4): p. 484-494.
39. Koopmans, M., et al., *Chimerism in kidneys, livers and hearts of normal women: implications for transplantation studies.* Am J Transplant, 2005. **5**(6): p. 1495-502.
40. Hong, Y.C., et al., *Hair follicle: a reliable source of recipient origin after allogeneic hematopoietic stem cell transplantation.* Bone Marrow Transplant, 2007. **40**(9): p. 871-4.

41. van Rood, J.J., et al., *Reexposure of cord blood to noninherited maternal HLA antigens improves transplant outcome in hematological malignancies.* Proc Natl Acad Sci U S A, 2009. **106**(47): p. 19952-7.

42. Burlingham, W.J. and J.L. Nelson, *Microchimerism in cord blood: mother as anticancer drug.* Proc Natl Acad Sci U S A, 2012. **109**(7): p. 2190-1.

43. van Rood, J.J., A. Scaradavou, and C.E. Stevens, *Indirect evidence that maternal microchimerism in cord blood mediates a graft-versus-leukemia effect in cord blood transplantation.* Proc Natl Acad Sci U S A, 2012. **109**(7): p. 2509-14.

44. Kanaan, S.B., et al., *Immunogenicity of a rheumatoid arthritis protective sequence when acquired through microchimerism.* Proc Natl Acad Sci U S A, 2019.



5

MORE THAN RESULTS: THE CLINICAL AND RESEARCH RELATIONSHIP IN THE EVOLVING DETECTION AND SURVEILLANCE OF SARS-COV-2

Based on:

Johnson, B.N., Cooper, L., Beck, J.J., Finnicum, C.T., Davis, C.M., Van Asselt, A.J., Kallsen, N., Silvernail, C., Viet, S., Long, K., Huizenga, P., Vaselaar, E., Pfeifle, M., Nelson, L., Streier, Z., Katz, C., Leonard, K.B., Matthiesen, N., Buschette, N., Weisser, L., Jahnke, M., Bohlen, K., Young, S.L., Sulaiman, R.A., Prouse, B.R., Basel, D.A., Black, M., Hansen, R., Petersen, J.L., Willemsen, G., Boomsma, D.I., Bleile, M.J., Ehli, E.A. (2022)

More than results: the clinical and research relationship in the evolving detection and surveillance of SARS-CoV-2. Submitted for publication.

ABSTRACT

Introduction: During the Coronavirus Disease 2019 (COVID-19) pandemic, real-time reverse transcription polymerase chain reaction (RT-PCR) became an essential tool for laboratories to provide high-sensitivity qualitative diagnostic testing for patients and real-time data to public health officials. Here we explore the potential predictive value of quantitative data from RT-PCR cycle threshold (Ct) values in epidemiological measures, symptom presentation, and variant transition.

Methods: Data from 74479 patients referred to the Avera Institute for Human Genetics (AIHG) for COVID-19 testing in 2020 were matched by calendar week to epidemiological data reported by the South Dakota Department of Health to examine correlation with state hospitalizations and deaths. Symptom data for 101 unique patients were reduced by exploratory factor analysis and investigated for association with patient age and sample metrics. We also explore changes in Ct values during variant transition detected by genomic surveillance sequencing within the AIHG testing population during 2021.

Results: Measures from AIHG diagnostic testing strongly explain variance in the South Dakota state positivity percentage (R^2 = 0.758), a two-week delay in hospitalizations (R^2 = 0.856), and a four-week delay in deaths (R^2 = 0.854). Patient symptoms produced three groups by factor analysis and demonstrated variation in trends associated with age, Ct value, and time from collection. Additionally, findings among SARS-CoV-2 variants indicate that variant transition may influence community quantitative measures.

Conclusions: Together, these findings suggest that measures of Ct value in RT-PCR diagnostic assays combined with routine screening have valuable applications in monitoring the dynamics of SARS-CoV-2 within communities.

1. INTRODUCTION

In December of 2019 the attention of the world turned to China as reports emerged involving cases of acute respiratory distress of unknown origin at health facilities in Wuhan, Hubei Province. The condition which has come to be known as Coronavirus Disease 2019 (COVID-19), was discovered to be caused by a novel betacoronavirus of the *Coronaviridae* family, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [1-3]. The sequence of the virus was published in January 2020 with 96.2% genetic similarity to a known SARS-related coronavirus found in bats, the natural reservoir host [4, 5]. In March of 2020 the World Health Organization (WHO) declared COVID-19 a pandemic, which led to the initiation of a series of federal and local responses to prepare health care systems in the United States for testing, tracing, and treating patients with COVID-19 [6]. Early in the pandemic, testing in the United States (U.S.) was only conducted at the U.S. Centers for Disease Control and Prevention (CDC) and public health laboratories [2, 7]. Due to the recognized potential impact of the virus on public health, the U.S. department of Health and Human Services enabled the Food and Drug Administration (FDA) to issue Emergency Use Authorization (EUA) for medical products for the diagnosis and treatment of COVID-19 [8, 9]. The EUA issued on March 31, 2020 included *in vitro* diagnostic laboratory developed tests (LDT) for the detection and diagnosis of COVID-19, such as targeted real-time reverse transcription polymerase chain reaction (RT-PCR) for detecting the genomic material of SARS-CoV-2 in human upper respiratory samples [10, 11].

Based in Sioux Falls, South Dakota, the Avera Health system is a vertically integrated health care delivery network providing health services at hospitals and clinics to a population of nearly one million people in South Dakota and the surrounding region including the states of Minnesota, Iowa, Nebraska, and North Dakota. As a part of an integrated health care system, the Avera Institute for Human Genetics (AIHG) incorporates both clinical and research experience and was thus uniquely positioned in the spring of 2020 to rapidly develop, validate, and operationalize an RT-PCR assay for the detection of SARS-CoV-2 and diagnosis of COVID-19. The research staff started the process of development of a diagnostic assay, leveraging their experience and technical proficiency with low-concentration detection of genetic material via quantitative PCR as obtained in studies of telomere repeat mass [12] and quantification of male microchimerism [13, 14]. Simultaneously, the clinical staff at

AIHG began preparing the laboratory space for workflow optimization and evaluating the regulatory requirements to develop a comprehensive validation plan. Together, the research and clinical teams successfully developed, tested, and validated a SARS-CoV-2 diagnostic assay to address the needs of the community. Further, as SARS-CoV-2 variants emerged, the genomic sequencing capabilities at AIHG supported genomic surveillance through the available samples, technical proficiency, and on-site sequencing instrumentation (see Figure 1 for a timeline).

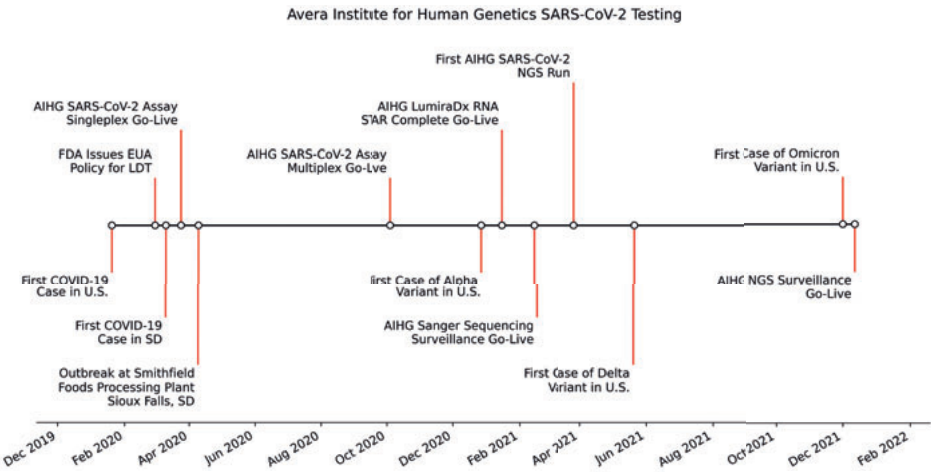


Figure 1: *Key events for SARS-CoV-2 test development and surveillance programs at the Avera Institute for Human Genetics (AIHG) are presented alongside major COVID-19 related events in the state of South Dakota (SD) and United States (U.S.). Additional acronyms: Food and Drug Administration (FDA), Emergency Use Authorization (EUA), Laboratory Developed Test (LDT), Next Generation Sequencing (NGS).*

Here we present studies to examine how population-level molecular genetic testing data may be informative for evaluating dynamics of emerging pathogens in a local population and the associated demand for healthcare services. We first report the findings of a retrospective study investigating how changes in weekly variance in diagnostic test positivity and future hospitalizations and deaths for the local population of South Dakota (SD) and surrounding area (catchment area of approximately one million persons) are reflected in real-time RT-PCR molecular diagnostic assays. We further examine variation in clinical presentation among patients tested in the year 2020 (23 March to 31 December) by clustering symptoms and exploring how these are explained by age, timing of symptom documentation,

and quantitative RT-PCR data. We investigate if SARS-CoV-2 variants present different viral loads in the local patient population as measured by surveillance sequencing and associated quantitative RT-PCR diagnostic assay data generated at AIHG during times of SARS-CoV-2 variant transition in the spring and winter of 2021.

Box 1. Terms Frequently Used

Reverse Transcription Polymerase Chain Reaction (RT-PCR)
: Molecular genetic technique for transcribing RNA to DNA using a reverse transcriptase enzyme, followed by PCR amplification of nucleic acids achieved by DNA replication induced by a series of temperature changes, called thermocycling.

Primer
: A short single-stranded nucleic acid complementary to 3' end of DNA target. Used as a starting point for polymerase to begin synthesis of complementary strand during polymerase chain reaction.

Probe
: A short single-stranded nucleic acid complementary to a specific target sequence labelled with a fluorescent reporter molecule for measuring DNA amplification.

Singleplex PCR
: A type of PCR assay used to detect and measure a single target in each reaction.

Multiplex PCR
: A type of PCR assay used to detect and measure multiple different targets in each reaction, accomplished through use of different fluorescent reporter molecules with different emission spectra for each target.

Cycle Threshold (Ct)
: A measure of real-time PCR indicating the cycle where the level of fluorescence crosses the defined threshold to indicate a positive result for the assay target. This value is inversely related to the amount of starting material.

Sanger Sequencing
: Also called chain termination method, this sequencing is accomplished by adding labeled di-deoxyribonucleotides to PCR which stops DNA replication. The fragments of various lengths are separated by electrophoresis to read the order of nucleotides by fragment size, shortest to longest.

Next-Generation Sequencing (NGS)
: Also called massively parallel sequencing, this sequencing method is accomplished by preparation of a DNA library of small fragments from the sample genome that are all simultaneously sequenced on a NGS instrument. Bioinformatics are used to map the fragments to the reference genome and piece together the original sample genome.

2. METHODS

Sample collection and storage

Samples for testing were delivered to AIHG from facilities across the Avera Health system throughout the course of the pandemic. Upper respiratory samples (e.g. nasopharyngeal and nasal) were collected in Universal Transport Media (UTM), Viral Transport Media (VTM), or phosphate buffered saline. Specimens were transported to the AIHG in an appropriate media at ambient temperature or at 2°C to 8°C for up to 72 hours after sample collections. Samples with delayed shipping time were stored at -70°C and shipped on dry ice. Upon sample receipt at the AIHG, samples were stored at 2°C to 8°C for up to 72 hours after sample collection or at -70 °C or colder for longer term storage.

Upper respiratory samples collected and confirmed acceptable for testing were processed by laboratory personnel in a Class II biological safety cabinet. Samples were extracted utilizing the QIAGEN RNeasy® Mini Kit (QIAGEN), RNAdvance Viral XP kit (Beckman Coulter), or the QiaSymphony® DSP Virus/Pathogen Midi Kit according to the manufacturer's instructions. Automated extractions were validated for use on the QIAsymphony (QIAGEN) and Biomek i7 (Beckman Coulter). Following RNA extraction samples were either tested immediately or stored at -70 °C or colder until samples could be processed for genetic testing.

Diagnostic testing for COVID-19

Real-time reverse transcription polymerase chain reaction (RT-PCR) diagnostic testing for COVID-19

In March of 2020 AIHG produced a laboratory developed test (LDT) for use in the diagnosis of patients with COVID-19. The Avera Institute for Human Genetics SARS-CoV-2 Assay is a real-time RT-PCR assay which includes three separate primer/TaqMan probe sets; two sets targeting SARS-CoV-2 specific regions of the nucleocapsid (N) gene, labeled N1 and N2, in addition to a human RNase P (RP) gene target for sample and extraction quality control. The sequences of the assay primers and probes (Supplemental Table 1) were selected to match the Centers for Disease Control and Prevention (CDC) 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel [15]. The Avera Institute for Human Genetics SARS-CoV-2 Assay was developed as a singleplex assay validated for use on the following Applied Biosystems™ PCR

instruments: 7500 Fast, ViiA™ 7, and QuantStudio™ 7 Flex. Each 96-well PCR reaction plate could include a maximum of 32 tests (32 tests x 3 reactions = 96 wells) or 30 patient samples plus a positive and negative reaction control. Submission through the FDA EUA process required a comprehensive comparison study of the assay. In coordination with the South Dakota Public Health Laboratory, AIHG received samples previously tested using an FDA EUA diagnostic test to conduct the comparison study. Submission of the validation to the FDA was completed at the end of March 2020, followed by official FDA authorization of the test under the EUA (EUA200410) (https://www.fda.gov/media/138331/download). The singleplex Avera Institute for Human Genetics SARS-CoV-2 Assay was implemented into AIHG clinical operations for diagnostic testing and was used in testing over 35,000 samples from the end of March 2020 to September 2020.

In response to a surge in testing demand, the assay was multiplexed in October 2020 to consolidate space on the 96-well PCR reaction plates to allow for up 94 patients plus a positive and negative control to be run on each plate. This was accomplished by including distinct reporter dyes containing unique emission spectra (see Table 2) [16]. As part of the assay validation, the multiplex Avera Institute for Human Genetics SARS-CoV-2 Assay was tested in a systematic panel including all assays in singleplex, duplex, and triplex to assess for any inhibition or unexpected amplification in the primer/TaqMan probe pool. Subsequent validation and comparison studies were successful, and the multiplex Avera Institute for Human Genetics SARS-CoV-2 Assay was submitted as an amendment to the FDA EUA. The multiplex assay was validated for use on the following Applied Biosystems™ PCR instruments: ViiA™ 7, QuantStudio™ 7 Flex, and QuantStudio™ 12K Flex. The decrease in plate space required by each sample resulted in 313% (94/30 x 100% = 313%, excluding controls) increase in sample throughput per 96-well PCR reaction plate compared to the singleplex assay. AIHG completed over 42,000 diagnostic tests using the multiplex Avera Institute for Human Genetics SARS-CoV-2 Assay from October 2020 to January 2021.

Qualitative result interpretation of the assay was conducted according to the information provided in Supplemental Table 2. In brief, the RNase P target was utilized as a measure of sample and extraction quality control which is required for interpretation of a negative PCR result. Alternatively, the assay requires the detection of both nucleocapsid targets (N1 and N2) for confirmation of a SARS-CoV-2 positive sample. As both nucleocapsid (N) gene targets must be detected for a positive test

result, the mean of the N1 and N2 test results is used for analyses in this study. Included on each test plate were two controls: a positive control and non-template control (NTC) which is required to achieve the expected results (Supplemental Table 3) before confirmation of patient sample results. An additional extraction control was tested for every extraction batch to evaluate the quality of the extraction procedure. The cycle threshold (Ct) value of the test is a continuous variable defined as the cycle of the RT-PCR reaction where signal intensity crosses the defined threshold of detection, which is inversely correlated with the amount of starting material.

Patient selection for study of diagnostic testing in 2020

At Avera Health, SARS-CoV-2 diagnostic testing was completed for inpatients, outpatients and as a part of standard screening (e.g. pre-surgical). When clinic and hospital laboratories were challenged with testing supply shortages and the immense community demand for testing services, samples were directed to AIHG as a reference laboratory for SARS-CoV-2 testing. Patients were referred for diagnostic testing via a physician based on standard of care clinical testing criteria, including known close contact, exposure, symptoms, or mandatory screening.

To test variation in clinical presentation and state epidemiology explained by Ct values and positivity of SARS-CoV-2 molecular diagnostic assays, we included patients who had an Avera Institute for Human Genetics SARS-CoV-2 Assay RT-PCR diagnostic test completed between 23 March 2020 and 31 December 2020 on either the original singleplex Avera Institute for Human Genetics SARS-CoV-2 assay (N = 35524) or its successor the multiplex assay (N = 38955). During this time the Avera Institute for Human Genetics SARS-CoV-2 Assay was exclusively utilized for diagnostic testing at AIHG. In addition, data from the year 2020 preceded prominent availability of FDA EUA vaccines and SARS-CoV-2 variants that could confound findings. Samples that returned an inconclusive result that was not resolved on follow-up tests were excluded.

Symptom information was obtained from the Avera Health system electronic medical record (EMR) for all patients who had a test completed using the Avera Institute for Human Genetics SARS-CoV-2 RT-PCR assay in the year 2020. Symptom entries (N = 54883 entries; 9171 unique patients) were matched to testing data and included all data within a three-week interval of sample collection (>7 days before and >14 days after sample collection), leaving 1703 individual entries for 101 unique patients. Symptoms documented as text were manually interpreted for presentation

of primary symptom categories including ageusia (loss of taste), anosmia (loss of smell), body aches, chills, cough, diarrhea, fatigue, fever, headache, nausea/vomiting, shortness of breath, and sinus congestion. The resulting binary classification (presence = 1 or absence = 0) of each primary symptom category was used for all comparisons. To examine differences in presentation at different timepoints relative to sample collection, the data were binned in 7-day windows relative to the date of sample collection for each patient. This binning of categorical symptom data resulted in one seven-day window (0-7 days) before sample collection and two seven-day windows (0-7 days and 7-14 days) after the day of sample collection (N = 101 unique patients; 125 weekly symptom entries) (Table 1).

Table 1: *Patient sex and age for entries from electronic medical record (EMR) included in the analysis of symptoms.*

Characteristic	Total (unique patients)	0-7 Days Prior to Sample Collection	0-7 Days Post-Sample Collection	7-14 Days Post-Sample Collection
Total	101	21	52	52
Male	52	10	29	27
Female	49	11	23	25
Age Range (Mean +/- SD)	20-98 (61.8 +/- 17.5)	20-98 (58.6 +/- 20.6)	26-96 (63.8 +/- 18.1)	23-85 (63.1 +/- 14.1)

Sequencing for SARS-CoV-2 Genomic Surveillance

Sanger Sequencing

At the end of 2020 information from the United Kingdom indicated the emergence of a SARS-CoV-2 variant, Pango lineage B.1.1.7 (WHO nomenclature: Alpha), that was more transmissible and potentially more virulent than the original strain [17]. Soon after, news of other emerging variants began to follow, including B.1.351 (Beta), P.1 (Gamma), and B.1.427/B.1.429 (Epsilon). [18, 19] In response, AIHG developed a Sanger sequencing test to selectively sequence the receptor binding domain (RBD) region of the SARS-CoV-2 S gene, a region known to feature key mutations of early variants of concern (VOC) and variants of interest (VOI). Beginning February 2021, each week a selection of positive samples was processed for surveillance sequencing of the RBD region to monitor changes in the prevalence of notable mutations associated with SARS-CoV-2 VOC and VOI. The N501Y mutation (point substitution A23063T) of the

spike protein was initially indicative of the first VOC Alpha and in combination with other key mutations such as E484K could indicate VOCs such as Beta or P.1 Gamma.

The RBD region of interest is selectively amplified via a nested PCR reaction (Supplemental Figure 1) using two primer sets specific to the target region of the SARS-CoV-2 genome (Supplemental Table 4). The samples first underwent PCR for reverse transcription to cDNA and to amplify a larger 746 base pair (bp) fragment containing the target region. This step used the TaqPath 1-Step RT-PCR master mix (Applied Biosystems™) and pre-amplification primers. The PCR cleanup step was accomplished by using ExoSAP-IT™ (Applied Biosystems™) to hydrolyze excess primers and nucleotides. The product was transferred to another plate in either one or two wells per samples for either unidirectional or bidirectional sequencing respectively. Here, a second PCR reaction was prepared, using BigDye PCR master mix (Applied Biosystems™) and M13-tailed primers, further amplifying the 443 bp target with added M13 tags for directional sequencing. Next the reagents for cycle sequencing were added directly to the wells, including the BigDye Direct sequencing master mix (Applied Biosystems™) and M13 Forward or M13 Reverse primer to the respective wells to control the direction of sequencing. After cycle sequencing, sample clean-up was accomplished using BigDye XTerminator (Applied Biosystems™) according to the manufacturer's instructions and the samples were analyzed using a 3500 Genetic Analyzer (Applied Biosystems™). Sequences were interpreted for mutations using Variant Reporter Software (Applied Biosystems™). Mutations were recorded and specific mutations for the key variant mutations N501Y, E484K, L452R, and K417N/T were monitored based on the known mutation combinations associated with the Alpha, Beta, Epsilon, and Gamma variants.

Next Generation Sequencing

In a natural progression from the Sanger sequencing surveillance method, AIHG began development of a more comprehensive whole genome sequencing surveillance strategy to monitor the evolution of SARS-CoV-2 variants in the Avera Health footprint. This required the evaluation of a library preparation kit and development of an in-house bioinformatic analysis pipeline. Using the expertise and instrumentation for next-generation sequencing (NGS) available at AIHG, a pilot study of seven samples was successfully processed through library preparation and sequenced on the local MiSeq (Illumina) system. The output from the NGS

instrument was processed through the custom analysis pipeline to produce quality control metrics, consensus sequences, and Pango lineage calls for each sample. As a part of the AIHG effort to provide the health system leaders with ongoing genomic information on the changing presentation of SARS-CoV-2 variants in the community, samples were similarly selected from the positive samples tested at AIHG to sequence on a weekly basis when surveillance efforts were necessitated.

Similar to the Sanger sequencing method, NGS was conducted on a selection of positive samples to generate whole genome sequences as a comprehensive method of genomic surveillance. The library preparation kit used was the CleanPlex® SARS-CoV-2 FLEX Panel (Paragon Genomics). The assay is an amplicon-based design including two primer pools of 343 amplicons to cover the entire genome. Additional Emerging Variant Panel add-on primer pools (Paragon Genomics) were spiked in when available to address the effects of genomic mutations on sequence coverage in SARS-CoV-2 variants. Samples were dual-indexed with i7 and i5 indexes for Illumina platforms. Libraries were pooled and evaluated via electrophoresis for library size and shape as well as the presence of non-specific product using a High Sensitivity DNA Kit (Agilent) run on a 2100 Bioanalyzer (Agilent). The construction of the final library pool was determined via the KAPA Library Quantification Kit (Roche) to ensure an adequate quantity of genomic material is loaded on the NGS instrument for sequencing. Sequences were produced via sequencing by synthesis on a MiSeq (Illumina) instrument in 2x150 base-pair (bp) paired-end runs using a MiSeq reagent kit v2 (Illumina).

Bioinformatic analysis of the resulting FASTQ files was handled on a local Linux workstation by an in-house analysis pipeline - Surveillance Analysis of NGS Data for SARS-CoV-2 Lineage And Sequence Homology (SANDSLASH). The pipeline was developed to take user provided demultiplexed FASTQ files from the MiSeq (Illumina) instrument as input and output QC metrics, consensus sequences, and Pango lineage calls. The pipeline processed the input files through FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC [20] to produce QC metrics. FASTQ files were then converted to unaligned BAM (uBAM) files with Picard (http://broadinstitute.github.io/picard/). The resulting pair of uBAM files per sample were directed to viral-pipelines (https://github.com/broadinstitute/viral-pipelines) for reference-based assembly of the SARS-CoV-2 virus genome. The resulting consensus sequence in FASTA format were subsequently processed via the Phylogenetic

Assignment of Named Global Outbreak Lineages (Pangolin) [21] command line tool to generate Pango lineage calls.

Patient selection for study of surveillance sequencing in 2021-2022

When AIHG initiated a genomic surveillance program in February 2021, diagnostic testing operations had converted to a new diagnostic assay. Collaboration with LumiraDx had resulted in the validation of their Lumira Dx SARS-CoV-2 RNA STAR Complete EUA test for use on the QuantStudio™ 7 Flex (Applied Biosystems). The Lumira Dx SARS-CoV-2 RNA STAR Complete EUA test eliminated the need for sample extraction and introduced rapid amplification methods to AIHG. This alleviated concerns of testing capacity and staffing burden for laboratory operations. Samples continued to be collected for standard of care clinical diagnostic testing for inpatients and outpatients as described above. Samples for genomic surveillance were randomly selected from those samples that tested positive for SARS-CoV-2 using an FDA EUA diagnostic test at AIHG. Assessment of sequencing result success led to integration of minimum sample concentration measured by the AIHG multiplex assay maximum Ct value as an additional selection criterion for each surveillance method (Sanger: Ct max = 30, NGS: Ct max = 26).

South Dakota State Epidemiology Data

Avera Health is based in Sioux Falls, the largest city in SD, with a geographic presence that encompasses the majority of the state making SD state data the most appropriate for comparison to the larger population. Data from the South Dakota Department of Health (SD-DOH) were collected and published by The COVID Tracking Project at The Atlantic from 15 March 2020 to 7 March 2021. The data from The COVID Tracking Project (https://covidtracking.com) includes information of testing, cases, hospitalizations, and deaths within the state. These data were utilized for all analyses examining the relationship with RT-PCR test data from March 2020 to December 2020. To conduct analyses of epidemiological metrics including diagnostic test positivity, hospitalizations, and deaths, the data were summarized by calendar week (Sunday to Saturday) to correct for weekly trends in testing and reporting data. Further, as AIHG data are reported to the SD-DOH, weekly SD state data used for comparison were corrected by subtracting the weekly AIHG test results from the reported weekly state summary data.

Statistical Analyses

First, to explore diagnostic test positivity between AIHG and state of SD reported data in 2020 (23 March to 31 December), the values were compared using the Wilcoxon sign-ranked test. Simple linear regression was applied to assess explained variance for the dependent variable SD weekly positivity using AIHG RT-PCR weekly positivity as a predictor.

Next, to investigate weekly variance in hospitalization count and death count explained by the Ct values produced by the Avera Institute for Human Genetics SARS-CoV-2 Assay in 2020, the data for AIHG diagnostic test Ct values were matched to the weekly count data for hospitalizations or deaths by fitting an exponential model ($Y = ab^X$; where Y = dependent variable: hospitalization or death count, a = y-intercept, b = constant factor, and X = independent variable: mean Ct value) to the data. This was repeated for the hospitalization and death data corresponding to a one-, two-, three-, and four-week delay from the reference week of RT-PCR test data. The resulting series of coefficient of determination values were compared to determine the optimal variance in hospitalization and death counts explained by AIHG RT-PCR weekly mean Ct value.

We further investigated the combination of symptom presentation and testing data within the population of patients tested at AIHG in 2020. To examine the similarities in weekly summarized categorical symptom data we applied agglomerative hierarchical clustering using Ward's minimum variance method. Additionally, we employed exploratory factor analysis with promax oblique rotation on symptom data to identify symptom groups and proportion of variance for each factor. We compared the Ct values, time of sample collection, and age, we compared these values within each of the symptom groups explained by factor loadings using Welch's *t*-test. Each of the factors were used to identify if a patient's symptom entry belongs to a given symptom group. If the entry included any of the symptoms belonging to a symptom group they were declared as belonging to that group.

To examine differences in the mean Ct values between different variant populations during variant transition in the population, we applied the Welch's t-test between samples confirmed by sequencing for the presence (presumably Alpha) and absence of the N501Y mutation in the spring of 2021 or Delta and Omicron variants in the winter of 2021.

Analyses were conducted using the software packages NumPy (1.21.0), SciKit Learn (v 1.0.2), and SciPy (v 1.8.0) implemented in Python (v 3.8.2). All tests were two-sided and statistical significance was considered for a P-value of less than 0.05.

Ethics

The study protocol including the aggregate use of RT-PCR testing measures, clinical measures and patient demographics were approved by the Institutional Review Board (IRB) of Avera, Sioux Falls, South Dakota, United States of America (Human & Health Services Registration #IRB00001096), on 7 August 2021 (IRB protocol #2021.056.HS).

3. RESULTS

AIHG Diagnostic Testing and South Dakota Epidemiology in 2020

The AIHG and corrected SD weekly diagnostic test positivity percentage paired data (median: 13.3% vs 9.4%, respectively) were significantly different as determined by the Wilcoxon sign-ranked test ($p = 6.16 \times 10^{-7}$). A linear regression model demonstrates that variance in SD positivity is greatly explained by AIHG positivity ($R^2 = 0.758$) (Supplemental Figure 2).

Mean Ct compared to the SD hospitalization data of the same week (reference week) and each of the following four weeks was investigated to determine the strongest explained variance via exponential regression. The strongest variance in SD hospitalizations explained by AIHG positivity was the hospitalizations data with a two-week delay after the reference week with a R^2 of 0.856 (Figure 2). In order from reference to four weeks after RT-PCR testing, the coefficients of determination (R^2) are 0.740, 0.829, 0.856, 0.795, and 0.758 (Supplemental Figure 3).

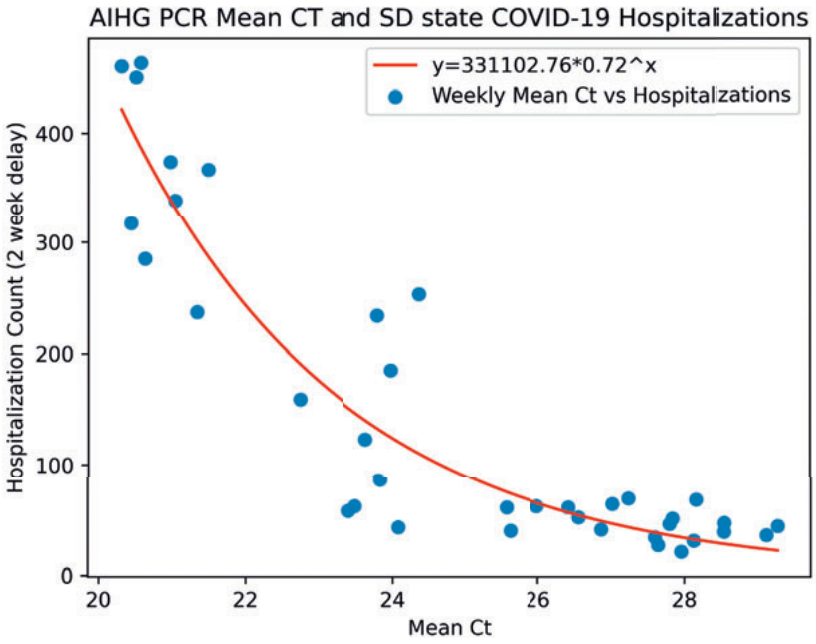


Figure 2: *Weekly AIHG PCR Ct mean data and SD hospitalizations data with a two-week delay after the PCR data were generated (R2 = 0.856).*

Similar to the analysis of hospitalizations, the RT-PCR data from tests performed at AIHG were investigated for correlation with the weekly COVID-19 related deaths in SD. Likewise, the mean Ct value from SARS-CoV-2 RT-PCR testing was compared to SD death data for the same week (reference week) and each subsequent week for 4 weeks. The strongest variance in SD COVID-19 related deaths explained by AIHG positivity was observed with a four-week delay in weekly deaths following the PCR testing (Figure 3). In order from reference to four weeks after PCR testing, the R^2 values are 0.525, 0.616, 0.724, 0.797, and 0.854 (Supplemental Figure 4).

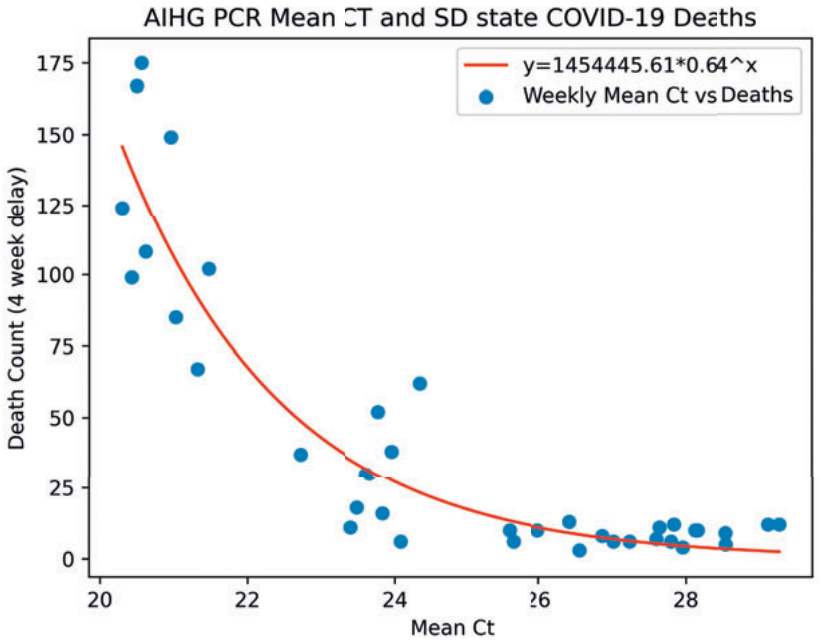


Figure 3: *Weekly AIHG PCR Ct mean data and SD deaths data with a four-week delay after the PCR data were generated (R2 = 0.854).*

AIHG Clinical Presentation in 2020

Data for each primary COVID-19 symptom were examined for changes in prevalence at the various timepoints (Figure 4). Cough, fatigue, and shortness of breath were the leading symptoms for individuals who had symptoms documented in the EMR, with these becoming substantially more prevalent following sample collection. Symptoms affecting the gastrointestinal tract (nausea and diarrhea) were more prevalent closest to the time of testing.

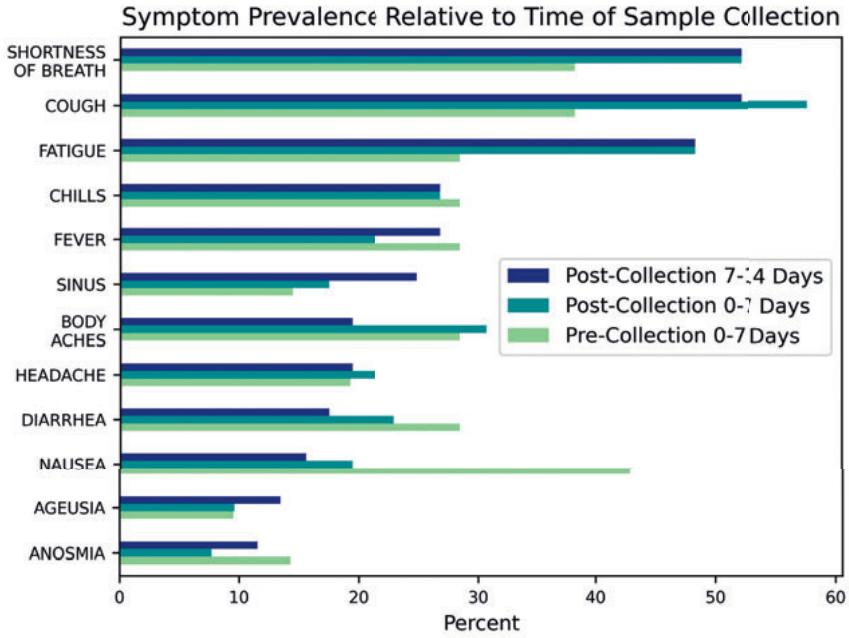


Figure 4: *Prevalence of primary COVID-19 symptoms in patients (N = 101) with positive diagnostic PCR test results. Prevalence is binned into 3 different 7-day windows: -7 to 0 days, 0 to 7 days, and 7 to 14 days where day 0 is the day of sample collection.*

Agglomerative hierarchical clustering of patient symptoms demonstrates that primary clustering of patient symptoms differentiates relatively high and low symptom prevalence (Figure 5A). In addition, hierarchical clustering reveals a clustering of symptoms affecting similar systems, as supported by correlations between the symptoms (Figure 5B) including anosmia and ageusia (0.87), nausea and diarrhea (0.57), and cough and shortness of breath (0.52).

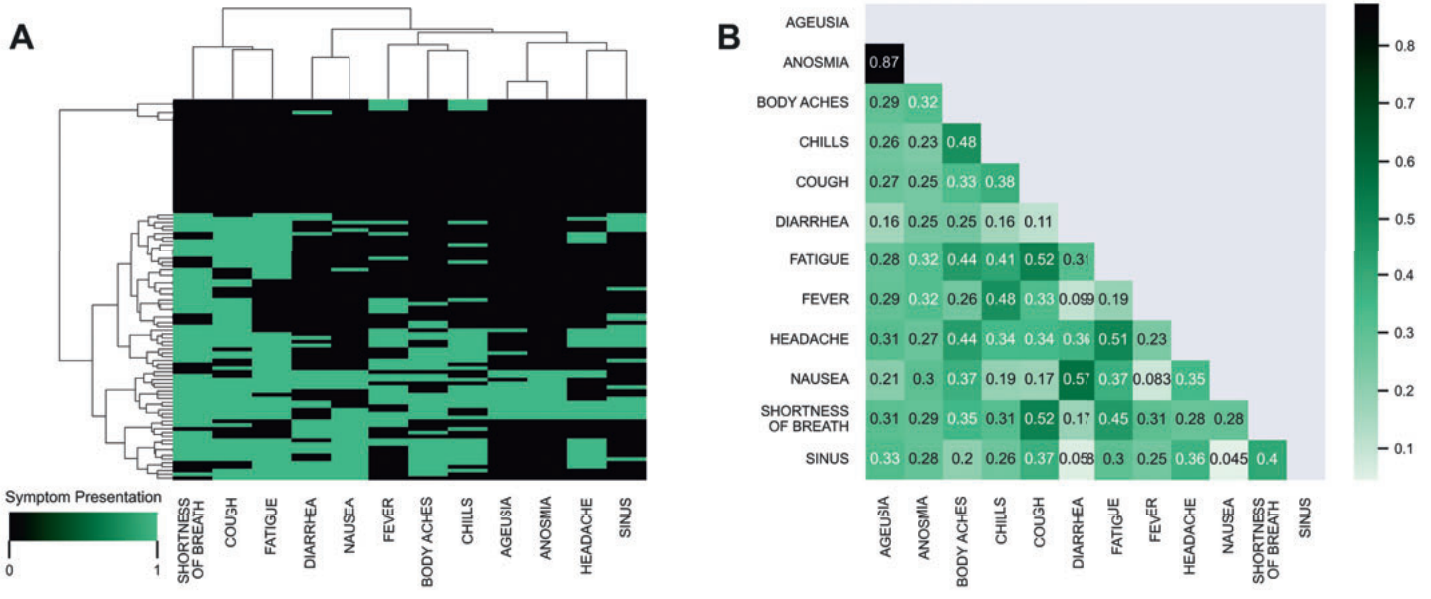


Figure 5: *A) Heatmap of patient (N= 101) symptoms within 7 days prior and 14 days after sample collection. For each symptom, a value of 0 indicates absence where a value of 1 indicates presence. B) Correlation matrix heatmap of symptom combinations.*

Further examination through exploratory factor analysis with promax oblique rotation revealed three relevant factors identified in this dataset (Table 2) and the symptoms correlated closely with the three factors. Factor 1 explained the greatest proportion of variance (0.228) by chills, cough, fatigue, fever, headache, shortness of breath, and sinus symptoms. The factor 1 group also were older compared to participants that did not present with factor 1 symptoms (mean: 65.2 vs 52.7 years, p = 0.010). Factor 2 (proportion variance = 0.144) presented substantial factor loadings by anosmia (loss of smell) and ageusia (loss of taste). Participants with weekly symptom entries that included factor 2 symptoms trended towards both a lower age (mean: 56.7 vs 62.7 years, p = 0.165) and lower Ct value (mean: 19.9 vs 22.4, p = 0.111) at time of sample collection. In addition, time from collection for the entry of COVID-19 symptoms in the EMR was significantly longer for both factor 1 (mean: 7.9 vs 4.8 days: p = 0.026) and factor 2 (mean 9.1 vs 6.7 days: p = 0.035). Lastly, factor 3 (proportion variance = 0.128) had substantial factor loadings for the gastrointestinal symptoms of nausea and diarrhea. Together the three factors demonstrate variation in measures related to the course of disease in the patient population, including Ct values, and explain 50.0% of the total variance among patient symptoms entered in the EMR.

Surveillance sequencing in 2021-2022

While RBD surveillance sequencing was in operation at AIHG from February to June 2021, 340 samples were successfully analyzed for key mutations and 175 presented with genomic variation resulting in the N501Y substitution (point substitution A23063T). The presence of N501Y in the sample population initially showed a substantial rise in prevalence followed by a slow increase over several weeks before establishing complete dominance (Figure 6). A portion of sequenced samples (n = 193) were processed on the multiplex Avera Institute for Human Genetics SARS-CoV-2 Assay resulting in an N501Y mean Ct of 20.892 (SE = 0.323) and non-N501Y mean Ct of 20.638 (SE = 0.461) which is not significantly different (p = 0.652).

Table 2: *Factor analysis of symptom data for each patient (N = 101). The factor loadings for three factors are provided for each symptom. Factor loadings highlighted in red indicate inclusion in the symptoms defining the factor. Age (years), Ct value, and time of symptoms from collection (days) present mean values and standard deviation (SD) of symptom entries featuring or not featuring the factor defining symptoms. Welch's t-test was used to compare between entries with or without symptoms for each factor for age, Ct value, and time from collection.*

	Value Description	FACTOR 1	FACTOR 2	FACTOR 3
Symptom				
AGEUSIA	Factor Loading	0.07630	0.84879	0.00741
ANOSMIA	Factor Loading	-0.07237	0.98144	0.14217
BODY ACHES	Factor Loading	0.44289	0.00872	0.27408
CHILLS	Factor Loading	0.64666	-0.05588	0.00520
COUGH	Factor Loading	0.79605	-0.09785	-0.09405
DIARRHEA	Factor Loading	-0.11088	0.03941	0.72636
FATIGUE	Factor Loading	0.58098	-0.07581	0.28421
FEVER	Factor Loading	0.49645	0.11947	-0.12665
HEADACHE	Factor Loading	0.42328	-0.00297	0.31913
NAUSEA	Factor Loading	-0.09410	0.05071	0.82634
SHORTNESS OF BREATH	Factor Loading	0.62039	0.00230	0.02725
SINUS	Factor Loading	0.55129	0.09970	-0.15811
Proportion Variance	Proportion	0.228	0.144	0.128
Age (in years)	P value	0.010	0.165	0.180
Entries featuring symptoms	Mean (SD)	65.162 (14.3)	56.667 (14.6)	64.829 (14.7)
Entries without symptoms	Mean (SD)	52.704 (22.1)	62.733 (17.9)	60.242 (18.7)
Ct Value	P value	0.893	0.111	0.392
Entries featuring symptoms	Mean (SD)	22.089 (5.8)	19.942 (5.1)	21.337 (5.8)
Entries without symptoms	Mean (SD)	21.890 (6.8)	22.401 (6.2)	22.406 (6.2)
Time From Collection for First Entry (in days)	P value	0.515	0.187	0.478
Entries featuring symptoms	Mean (SD)	5.162 (4.6)	6.267 (3.9)	4.457 (4.6)
Entries without symptoms	Mean (SD)	4.296 (6.3)	4.698 (5.2)	5.182 (5.3)
Time From Collection for Last Entry (in days)	P value	0.026	0.035	0.833
Entries featuring symptoms	Mean (SD)	7.946 (4.8)	9.133 (3.4)	6.943 (5.3)
Entries without symptoms	Mean (SD)	4.778 (6.5)	6.744 (5.7)	7.182 (5.6)

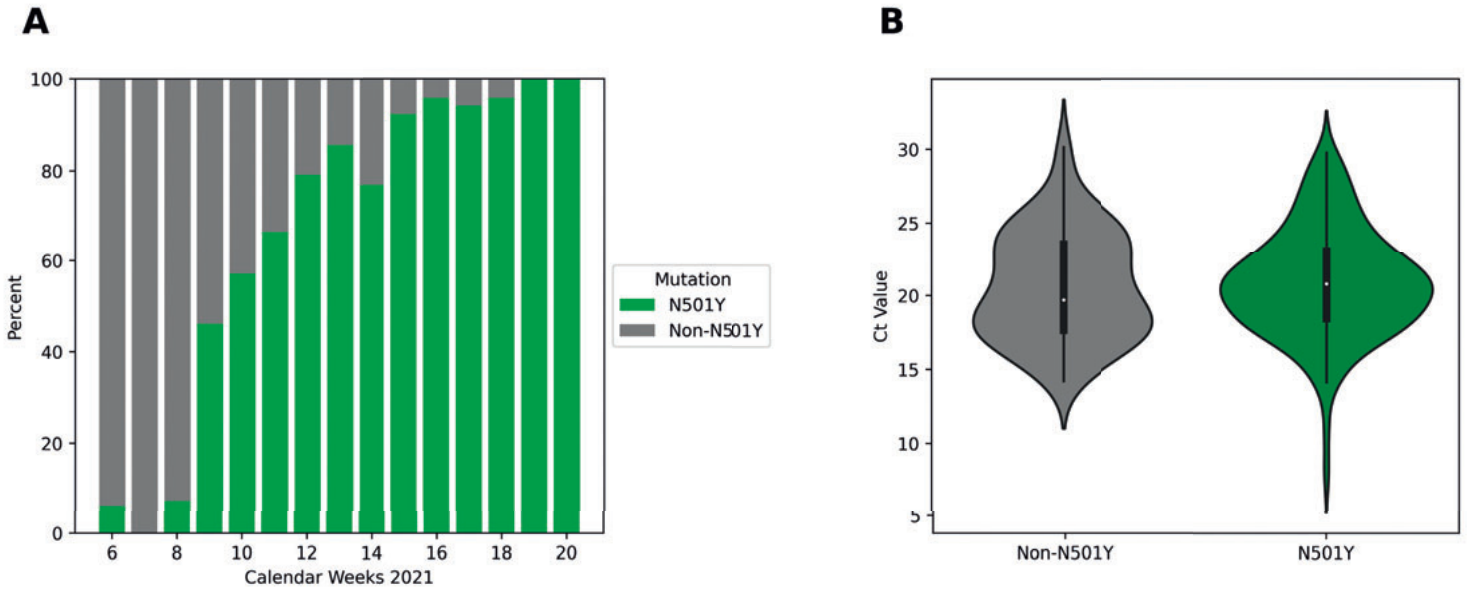


Figure 6: *A) Receptor binding domain targeted sequencing surveillance data produced by the AIHG from February 2021 to May 2021 (N = 356). Prevalence of N501Y substitution binned by the calendar week of the testing date. B) Violin plot of Ct values of the AIHG SARS-CoV-2 multiplex assay for samples based on N501Y substitution status by Sanger sequencing (mean: N501Y = 20.892, Non-N501Y = 20.638; Welch's t-test: p = 0.652).*

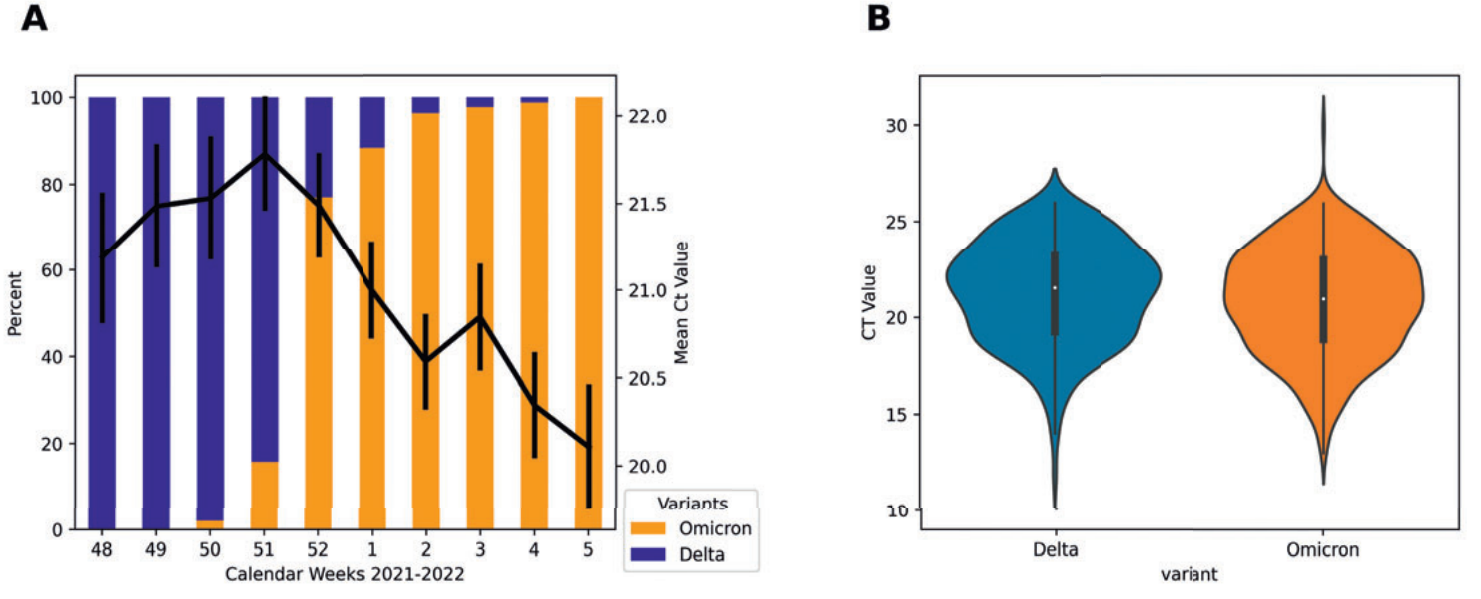


Figure 7: *Whole genome sequencing surveillance data produced by the AIHG from December 2021 to February 2022. A) Samples (N = 698) are binned by the calendar week of the sample collection date. The average Ct value of the sequenced samples produced by the AIHG SARS-CoV-2 multiplex assay are shown by the black line graph with error bars demonstrating the standard error of the mean. B) Violin plot of Ct values of the AIHG SARS-CoV-2 multiplex assay for samples identified as Delta (mean: 21.279) and Omicron (mean: 20.834) SARS-CoV-2 variants by whole genome sequencing (Welch's t-test: p = 0.0374).*

After the Delta variant became the dominant variant of SARS-CoV-2 in the summer of 2021, there were several months of no significant changes to variant prevalence. The AIHG whole genome sequencing surveillance program was brought into routine operation in November 2021 to begin testing a random selection of cases each week to monitor the emerging Omicron variant. As expected, the first complete batches of SARS-CoV-2 whole genome sequencing confirmed positive samples were entirely comprised of the Delta variant. Following identification of the first Omicron variant sample in mid-December, there was a dramatic shift to Omicron dominance over the two subsequent weeks (Figure 7A). Samples from individuals who were confirmed to carry Omicron have a similar distribution (Figure 7B) but a significantly lower mean Ct value compared to Delta samples (20.834 vs 21.279 respectively; Welch's *t*-test p = 0.0374) when tested using the multiplex Avera Institute for Human Genetics SARS-CoV-2 RT-PCR assay.

4. DISCUSSION

Throughout the progression of the COVID-19 pandemic, there has been an increasing demand for laboratory services to provide testing to not only patients in acute care, but also large-scale screening efforts to mitigate spread of the virus. Laboratory diagnostics have provided a valuable tool for both patient care and epidemiologic monitoring of COVID-19. As a molecular genetics institute, AIHG laboratory operations includes both clinical and research pillars, affectionately referred to as "cli-search". As a molecular genetics laboratory within a healthcare system, it was essential that the available resources be implemented for genetic interrogation of biological samples for the virus and its variants. As a result of the decision of the U.S. Department of Health and Human Services to enable the FDA to provide EUA of diagnostic tests, laboratories like AIHG were able to produce and validate LDTs in response to changes in the testing demands of the communities they serve [8, 9]. Here we present findings from retrospective analysis of SARS-CoV-2 diagnostic and surveillance testing conducted at AIHG over the last two years and discuss applications of diagnostic testing in response to future epidemics.

Findings from this retrospective analysis of COVID-19 diagnostic testing in 2020 have demonstrated that changes in the presentation of Ct value for cases confirmed by RT-PCR has a strong relationship with regional diagnostic positivity percentage

and future demand for healthcare services. As presented in our analysis of weekly diagnostic testing Ct values in our laboratory compared to state epidemiology data in SD, decreases in the mean Ct value demonstrate a very strong relationship with a two-week delay in hospitalization counts (R^2 = 0.856). Similarly, the weekly mean Ct value also presents a strong exponential relationship with a four-week delay in death data (R^2 = 0.854). This change indicates an increased prevalence of individuals presenting to test collection closer to the time of peak infection with high viral loads prior to a rise in demand for healthcare services.

Previous studies of the ancestral SARS-CoV-2 lineage have demonstrated a viral shedding time of 10 to 30 days depending on several factors related to host immune status including age, symptoms, comorbidities, and treatment [22]. Previous reports have also indicated correlation between symptom onset and low Ct values, presumably associated with increased viral load and infectiousness [23]. Our findings showed that spread of viral particles during the incubation period was followed by a rise in community infection over the following weeks. Cases identified through routine screening are likely to be mild or asymptomatic at the time of sample collection as the disease has not presented in a manner that necessitates seeking out care. Consequently, asymptomatic cases at time of collection have previously been shown to present with low viral loads and high Ct values [23]. This seemingly indicates that consistent population screening serves as a baseline and increases in testing of symptomatic cases, due to a rise in community infections, results in a testing population shift toward lower Ct values. Therefore, we propose mean Ct value may have important public health applications in evaluating if diagnostic screening efforts are sufficient based on the current virus dynamics in the community.

Among the patients tested at AIHG in 2020, we examined clinical presentation via COVID-19 symptoms documented in the electronic medical record (EMR). As expected, co-presentation of symptoms was most prevalent in symptoms affecting the same systems, such as sensory symptoms (ageusia and anosmia) as well as gastrointestinal symptoms (diarrhea and nausea). This finding was recapitulated in the factor analysis which grouped these symptoms into unique factors to explain variation in the dataset. Symptoms common to respiratory viral infection such as shortness of breath, cough, and fatigue were the most prevalent symptoms reported following sample collection and present with a significantly increased age demonstrated by factor 1. It is of interest that both factor 1 (respiratory viral

infection symptoms) and factor 2 (sensory symptoms) present with a significantly longer time of symptom documentation in the EMR compared to those without these symptoms and may be indicative of a longer course of severe disease in these patients. Examining patient population symptoms in relation to diagnostic RT-PCR testing Ct values provides insight into the timing of severe symptoms relative to the course of disease. For example, patients with ageusia or anosmia (factor 2) in our sample population are generally younger, present with lower Ct values at the time of testing and have a longer duration of symptom documentation in the EMR. In contrast, patients with GI symptoms (e.g. nausea and diarrhea) were more prevalent the week prior to sample collection and present a tendency towards older aged patients based on comparison of factor 3, warranting further investigation. Applying these methods to a larger number of samples and access to additional details regarding timing of symptoms and comorbidities would make this a useful approach for further distinguishing different groups of symptoms and clinical presentations among patients. Additional analysis including clinical outcomes and treatments may also be able to indicate effective treatment strategies for different clinical presentations of COVID-19.

SARS-CoV-2 variants of interest (VOI) and variants of concern (VOC) have since diversified the presentation of SARS-CoV-2 across the globe, and genomic surveillance has provided information on the trends of mutations and variants. The highly transmissible SARS-CoV-2 VOC have presented with selective advantages resulting in enhanced characteristics capable of outcompeting other circulating strains [24]. We present findings from investigation of diagnostic real-time RT-PCR testing metrics as they relate to the changes in SARS-CoV-2 variant prevalence over two critical timepoints of variant transition. Among these is the N501Y mutation, specifically located in the receptor binding domain of the spike protein, which increases the affinity for the human angiotensin-converting enzyme-2 (ACE2) receptor [25, 26]. However, our study investigating the Ct values of sequenced samples during the change from ancestral SARS-CoV-2 to N501Y (presumably Alpha) variants from February to June 2021 does not indicate that this results in a significant change in SARS-CoV-2 genomic material at time of collection. In contrast, the transition from Delta to Omicron from November 2021 to January 2022 presents with a decline in mean Ct values and a significant lower Ct among Omicron cases. This contradicts findings from others indicating no difference in Omicron viral load compared to Delta

[27, 28]. Interestingly, other studies have found symptomatic infection of Omicron has a shorter incubation period than Delta, with less debilitating symptoms and a shorter duration [29].

Based on these findings we propose that the community measure of Ct value indicates changes in timing of sample collection relative to the course of disease within the testing population. As samples in our study were collected from patients for diagnostic testing, the patients first required a test order for sample collection. Previous findings also indicate that symptomatic patients, especially with increased symptom burden, are most likely to test positive via PCR [30]. An increased proportion of individuals presenting to sample collection with symptoms or suspected infection closer to peak infection would result in lowering the mean Ct value [31]. Noting the decrease in cases in South Dakota during the time N501Y established dominance in the summer of 2021 and substantial rise in cases during the transition to Omicron at the beginning of 2022, the trajectory of case volumes in the community may have significantly contributed to the presentation of Ct values. In this context, the individual sample Ct values as a proxy for viral load may vary for a variety of reasons, including sample timing or an increased risk of severe disease [32]; however, decrease in aggregate Ct values may be caused by a population shift and necessitates further investigation as a potential indicator of increasing case numbers. While genomic surveillance to examine virus variants provides tangible evidence of virus changes in response to selective pressures, the real-time data provided by diagnostic tests may provide an early indicator that should be considered in monitoring the trajectory of COVID-19 cases in a testing population.

While we were able to differentiate clinical presentations of COVID-19 and observe trends in the community epidemiology data, there are still study limitations that must be considered. First, AIHG testing for SARS-CoV-2 was initiated as the lone provider of diagnostic testing within the Avera Health system but as resources became available to provide testing in on-site clinical laboratories AIHG testing became primarily used for routine testing due to the high-throughput platform. As a result, few individuals who actively presented with severe disease as recorded in the EMR had their samples processed at AIHG. The limited availability of records in the EMR detailing specific symptoms and their timing further restricted the study to a small number of individuals. We presume that small sample size and variation in documentation practices of symptom data may have limited findings in our analysis

of clinical presentation. For future emerging pathogens, we recommend strategic planning to establish systematic documentation of clinical presentation to aid in real-time investigation of trends in symptomology and relation with diagnostic laboratory results. Another limitation is vaccination status at the time of collection. In the U.S., the COVID-19 vaccines became widely available in 2021 and therefore were not considered to influence the results of 2020 data. Vaccination status is not considered to be a confounder in genomic surveillance results, but previous studies have examined vaccination status in relation to clinical severity with SARS-CoV-2 variants [33]. Timing of infection relative to vaccination, the number of doses, and effectiveness of the immune response to a specific variant may all influence the viral load estimated by RT-PCR Ct values. Findings of significantly lower Ct values in Omicron cases compared to Delta may be due to reduced vaccine efficacy in this SARS-CoV-2 variant. Epidemiologic study of vaccination status and novel variant transmission within a community is an area for further investigation.

The availability of a diverse portfolio of advanced molecular genetics techniques, technical proficiency, and “cli-search” experience provided the tools needed for AIHG to rapidly respond to the needs of the community through development of SARS-CoV-2 diagnostic testing and genomic surveillance. Further study of trends in the diverse symptom presentation of COVID-19 will provide valuable information to aid in triage strategies and follow-up care for patients. In addition, while there is insufficient evidence to support use of Ct values for individual clinical decisions regarding COVID-19 [34], these findings provide support for further applications of real-time RT-PCR diagnostic testing in public health monitoring during epidemics of highly transmissible respiratory viruses such as SARS-CoV-2.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ACKNOWLEDGEMENTS

We would like to thank the South Dakota Public Health Laboratory and especially Director Dr. Timothy Southern for providing the samples and support for comparison studies to validate our SARS-CoV-2 diagnostic testing at the Avera Institute for Human Genetics. We also would like to thank the staff of Avera Translational Oncology, Avera McKennan Hospital Histology, and ALUMEND for providing staffing support to meet

the demand for diagnostic testing early in the pandemic. These efforts were essential to supporting AIHG SARS-CoV-2 diagnostic testing and providing the best possible care to the community during this challenging time.

Finally, we would like to especially recognize the late Dr. Bruce R. Prouse, an AP/CP certified pathologist with Physicians Laboratory and Avera McKennan Hospital. In his role as Clinical Vice President of the Avera Laboratory Service Line, he provided critical leadership from the beginning and throughout the pandemic to ensure COVID-19 testing was available to meet the high demand coming into our labs from the hospital, clinics, and the community. We are grateful for his hard work, commitment, and care. He will be truly missed.

REFERENCES

1. Coronaviridae Study Group of the International Committee on Taxonomy of, V., *The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.* Nat Microbiol, 2020. **5**(4): p. 536-544.
2. Del Rio, C. and P.N. Malani, *COVID-19-New Insights on a Rapidly Changing Epidemic.* JAMA, 2020. **323**(14): p. 1339-1340.
3. Paules, C.I., H.D. Marston, and A.S. Fauci, *Coronavirus Infections-More Than Just the Common Cold.* JAMA, 2020. **323**(8): p. 707-708.
4. Zhou, P., et al., *A pneumonia outbreak associated with a new coronavirus of probable bat origin.* Nature, 2020. **579**(7798): p. 270-273.
5. Yan, Y., et al., *The First 75 Days of Novel Coronavirus (SARS-CoV-2) Outbreak: Recent Advances, Prevention, and Treatment.* Int J Environ Res Public Health, 2020. **17**(7).
6. Rajan, S., J. Cylus, and M. McKee, *Successful find-test-trace-isolate-support systems: how to win at snakes and ladders.* Eurohealth, 2020. **26**(2).
7. Del Rio, C. and P.N. Malani, *2019 Novel Coronavirus-Important Information for Clinicians.* JAMA, 2020. **323**(11): p. 1039-1040.
8. Shuren, J. and T. Stenzel, *Covid-19 Molecular Diagnostic Testing - Lessons Learned.* N Engl J Med, 2020. **383**(17): p. e97.
9. Ravi, N., et al., *Diagnostics for SARS-CoV-2 detection: A comprehensive review of the FDA-EUA COVID-19 testing landscape.* Biosens Bioelectron, 2020. **165**: p. 112454.
10. da Silva, S.J.R., et al., *Clinical and Laboratory Diagnosis of SARS-CoV-2, the Virus Causing COVID-19.* ACS Infect Dis, 2020. **6**(9): p. 2319-2336.
11. Kevadiya, B.D., et al., *Diagnostics for SARS-CoV-2 infections.* Nat Mater, 2021.
12. Finnicum, C.T., et al., *Relative Telomere Repeat Mass in Buccal and Leukocyte-Derived DNA.* PLoS One, 2017. **12**(1): p. e0170765.
13. Johnson, B.N., et al., *Male microchimerism in females: a quantitative study of twin pedigrees to investigate mechanisms.* Hum Reprod, 2021. **36**(9): p. 2529-2537.
14. Peters, H.E., et al., *Low prevalence of male microchimerism in women with Mayer-Rokitansky-Kuster-Hauser syndrome.* Hum Reprod, 2019.
15. Lu, X., et al., *US CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome Coronavirus 2.* Emerg Infect Dis, 2020. **26**(8).
16. Rappo, U., et al., *Impact of Early Detection of Respiratory Viruses by Multiplex PCR Assay on Clinical Outcomes in Adult Patients.* J Clin Microbiol, 2016. **54**(8): p. 2096-103.
17. Davies, N.G., et al., *Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England.* Science, 2021. **372**(6538).
18. Zhang, W., et al., *Emergence of a Novel SARS-CoV-2 Variant in Southern California.* JAMA, 2021.
19. Mascola, J.R., B.S. Graham, and A.S. Fauci, *SARS-CoV-2 Viral Variants-Tackling a Moving Target.* JAMA, 2021.
20. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-8.

21. Rambaut, A., et al., *A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology.* Nat Microbiol, 2020. **5**(11): p. 1403-1407.

22. Yan, D., et al., *Characteristics of Viral Shedding Time in SARS-CoV-2 Infections: A Systematic Review and Meta-Analysis.* Front Public Health, 2021. **9**: p. 652842.

23. Salvatore, P.P., et al., *Epidemiological Correlates of Polymerase Chain Reaction Cycle Threshold Values in the Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2).* Clin Infect Dis, 2021. **72**(11): p. e761-e767.

24. Hirabara, S.M., et al., *SARS-COV-2 Variants: Differences and Potential of Immune Evasion.* Front Cell Infect Microbiol, 2021. **11**: p. 781429.

25. Starr, T.N., et al., *Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding.* Cell, 2020. **182**(5): p. 1295-1310 e20.

26. Liu, H., et al., *The basis of a more contagious 501Y.V1 variant of SARS-CoV-2.* Cell Res, 2021. **31**(6): p. 720-722.

27. Laitman, A.M., et al., *The SARS-CoV-2 Omicron Variant Does Not Have Higher Nasal Viral Loads Compared to the Delta Variant in Symptomatic and Asymptomatic Individuals.* J Clin Microbiol, 2022. **60**(4): p. e0013922.

28. Fall, A., et al., *The displacement of the SARS-CoV-2 variant Delta with Omicron: An investigation of hospital admissions and upper respiratory viral loads.* EBioMedicine, 2022. **79**: p. 104008.

29. Menni, C., et al., *Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study.* Lancet, 2022. **399**(10335): p. 1618-1624.

30. Wohl, D.A., et al., *COVID-19 symptoms at time of testing and association with positivity among outpatients tested for SARS-CoV-2.* PLoS One, 2021. **16**(12): p. e0260879.

31. Walsh, K.A., et al., *SARS-CoV-2 detection, viral load and infectivity over the course of an infection.* J Infect, 2020. **81**(3): p. 357-371.

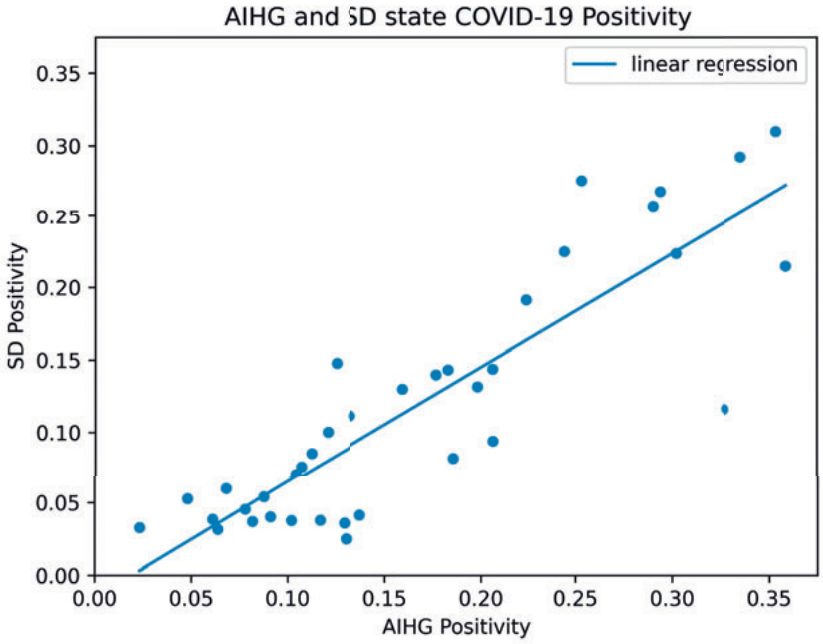
32. Fajnzylber, J., et al., *SARS-CoV-2 viral load is associated with increased disease severity and mortality.* Nat Commun, 2020. **11**(1): p. 5493.

33. Lauring, A.S., et al., *Clinical severity of, and effectiveness of mRNA vaccines against, covid-19 from omicron, delta, and alpha SARS-CoV-2 variants in the United States: prospective observational study.* BMJ, 2022. **376**: p. e069761.

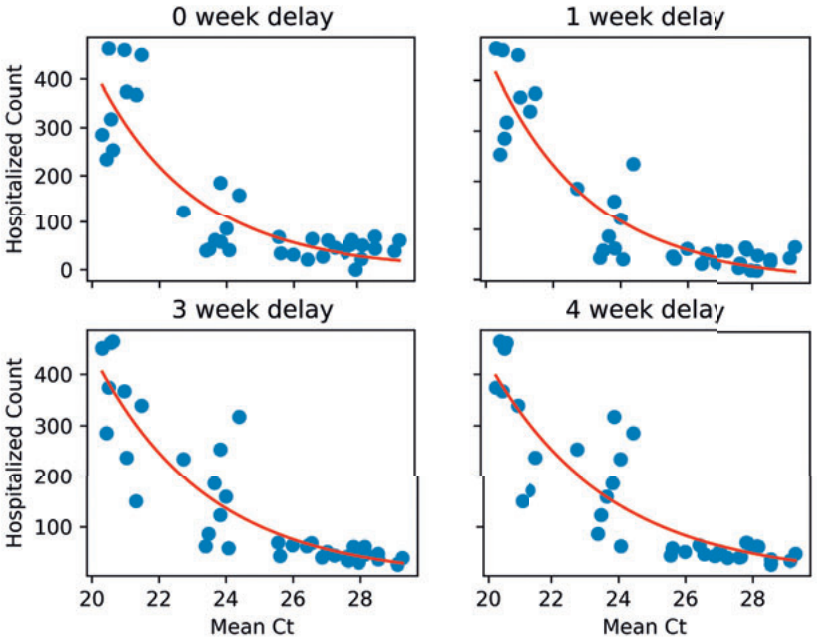
34. Evans, D., et al., *The Dangers of Using Cq to Quantify Nucleic Acid in Biological Samples: A Lesson From COVID-19.* Clin Chem, 2021. **68**(1): p. 153-162.

SUPPLEMENTARY MATERIALS

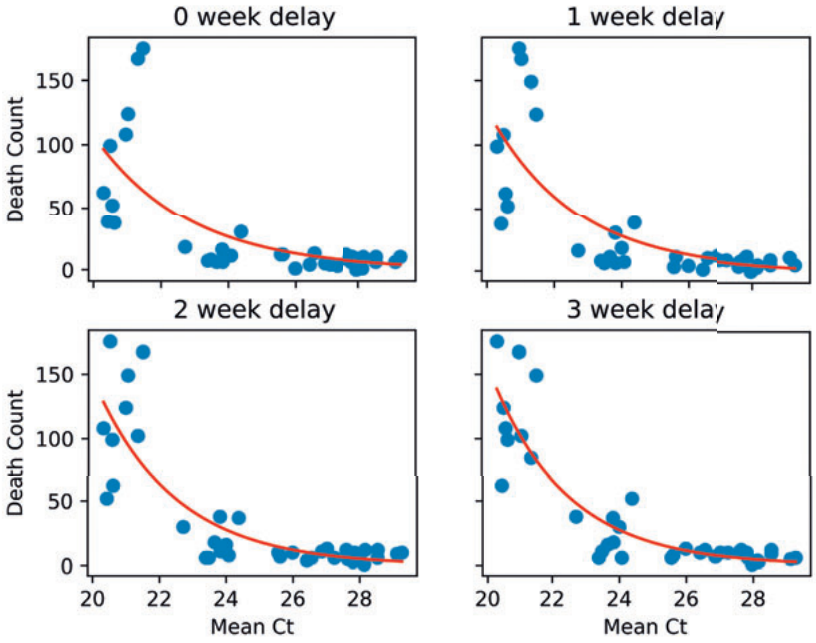
Supplemental Figure 1: *Sanger sequencing nested PCR annotation on SARS-CoV-2 virus genome, base positions 22621 - 23460. The binding sites for the initial PCR reaction are highlighted in pink, and similarly highlighted in blue are the sequences corresponding to the primers of the second PCR reaction. Highlighted in green is the sequenced region of interest.*



Supplemental Figure 2: *Scatter plot and linear regression of the AIHG and SD state weekly positivity (R^2 = 0.758).*



Supplemental Figure 3: *Weekly AIHG PCR Ct mean data and SD hospitalizations data with a zero-, one-, three-, and four-week delay after the PCR data were generated (R^2 = 0.740, 0.829, 0.795, and 0.758, respectively).*



Supplemental Figure 4: *Weekly AIHG PCR Ct mean data and SD deaths data with a zero-, one-, two-, and three-week delay after the PCR data were generated (R^2 = 0.0.525, 0.616, 0.724, and 0.797, respectively).*

Supplemental Table 1: *Details pertaining to the components and reporter dyes of the probes included in the AIHG SARS-CoV-2 multiplex PCR assay.*

Manufacturer	Component	Oligonucleotide Sequence
Integrated DNA Technologies	N1 (FAM) Probe	5'-FAM-ACC CCG CAT TAC GTT TGG ACC-BHQ1-3'
	N1 Forward Primer	5'-GAC CCC AAA ATC AGC GAA AT-3'
	N1 Reverse Primer	5'-TCT GGT TAC TGC CAG TTG AAT CTG-3'
	N2 (SUN) Probe	5'-SUN-ACA ATT TGC CCC CAG CGC TTC AG-BHQ1-3'
	N2 Forward Primer	5'-TTA CAA ACA TTG GCC GCA AA-3'
	N2 Reverse Primer	5'-GCG CGA CAT TCC GAA GAA-3'
	RNase P (ATTO 647) Probe	5'-ATTO647-TTC TGA CCT GAA GGC TCT GCG CG-BHQ-1-3'
	RNase P Forward Primer	5'-AGA TTT GGA CCT GCG AGC G-3'
	RNase P Reverse Primer	5'-GAG CGG CTG TCT CCA CAA GT-3'

Supplemental Table 2: *This table details the possible outcomes from the AIHG SARS-CoV-2 Assay and the corresponding interpretation.*

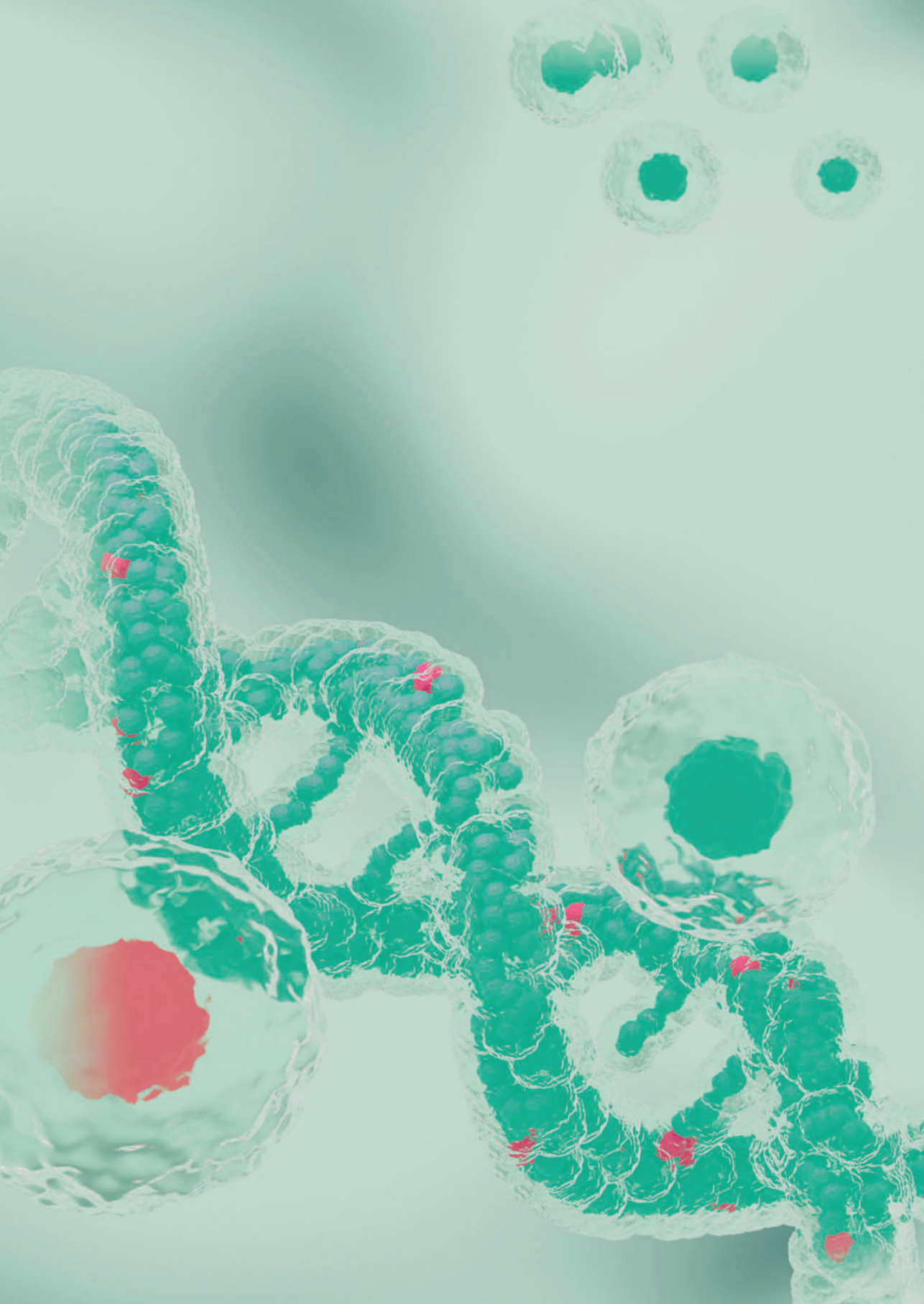
N1	N2	RP	Result Interpretation
+	+	±	SARS-CoV-2 detected
If only one of the two targets is positive		±	Inconclusive Result
-	-	+	SARS-CoV-2 not detected
-	-	-	Invalid Result

Supplemental Table 3: *Quality control materials and the expected results. Failure in a control is an indicator of a problem in the process described under the header "Used to Monitor".*

Control Type	Used to Monitor	N1	N2	RP	Expected Ct Values
Positive	Substantial reagent failure including primer and probe integrity	+	+	+	< 40.00 Ct
Negative	Reagent and/ or environmental contamination	-	-	-	None detected
Extraction	Failure in lysis and extraction procedure, potential contamination during extraction	-	-	+	< 40.00 Ct

Supplemental Table 4: *Primers for Sanger sequencing of the SARS-CoV-2 spike protein receptor binding domain located on the S gene. Note: M13 sequence tags are highlighted in red text.*

Sequence	PCR Reaction	Direction
CAACTGTGTTGCTGATTATTCTGTC	PCR 1	Forward
CCTGGTTAGAAGTATTTGTTCCTGGTG	PCR 1	Reverse
TGTAAAACGACGGCCAGTGTGATGAAGTCAGACAAATCGCTCC	PCR 2	Forward
CAGGAAACAGCTATGACCAACACCTGTGCCTGTTAAACCATTG	PCR 2	Reverse



6

SUMMARY AND DISCUSSION

SUMMARY AND DISCUSSION

The central objective of this thesis is to investigate exposures to genetically foreign material through quantitative measures of low concentration genomic material in biological samples. I begin with examining the current knowledge on the human health implications of natural chimerism, a biological manifestation resulting in an individual with genetically distinct cell populations originating from two or more unique zygotes. To further examine the chimerism phenomenon, I conduct two studies in which I implement a high-sensitivity assay for the relative quantification of Y chromosome through quantitative polymerase chain reaction (qPCR) in blood samples collected from females. First, I investigate the prevalence of male microchimerism in a case-control study of Mayer-Rokitansky-Küster-Hauser syndrome. This is followed by a study of female members of twin families to investigate the presentation of microchimerism both within and across generations. During my Ph.D. trajectory, research projects were interrupted due to the emergence of the Coronavirus Disease 2019 (COVID-19) pandemic, but new opportunities also arose. The prior development and implementation of a high sensitivity qPCR assay for detection of unique molecular genetic biomarkers for microchimerism proved to be of great value for the design of molecular diagnostic tests necessary to rapidly address the growing public health needs. Applying previous research and clinical laboratory experience, I was able to translate this methodology for application in the development, validation, and operationalization of a diagnostic real-time reverse transcription PCR (RT-PCR) assay to detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

The studies described in this thesis collectively explore quantitative methods to detect low concentrations of foreign genomic material and provide valuable details about the biological implications concerning the larger population.

In this chapter, I present a summary of the findings for my research which are described in detail in each of the preceding chapters. This is followed by a general discussion of human chimerism, applications of nucleic acid quantitation, and future perspectives on quantitative measures of genomic material in research and healthcare.

CHAPTER 2: CHIMERISM AND IMPLICATIONS ON HEALTH AND BEHAVIOR

The general understanding of human heritability is centered on the knowledge of human reproduction, where the fusion of haploid gametes produces a diploid zygote resulting in offspring that share one set of alleles with each parent. An individual embryo will proceed to develop a diverse array of tissues and organs, generally assumed to be comprised of cell populations that all originate from the same common zygote. Some individuals, however, do not meet this assumption as some of their tissues are composed of two genetically distinct cell populations. While there are three primary forms of chimerism, the exchange of cells via the placental vasculature has been suggested as the leading cause of natural chimerism and commonly results in < 1% concentration of donor cells within the host, termed microchimerism.

Chapter 2 provides an overview of chimerism through a systematic review of the literature to date of publication. The primary purpose of this chapter is to provide a comprehensive examination of chimerism literature, from which I compile findings related to health implications and also derive observations for future investigation. I begin by summarizing the various methodologies described in the literature and examining the relative merits for the application of each method in the detection and study of chimerism. Historically, chimerism had been discovered accidentally due to discrepancies in clinical testing. Heterogeneous cell populations in collected samples were first determined through less sensitive methodologies like routine blood antigen typing in clinical laboratories when the sample contained a relatively high concentration of the minor cell population (e.g. >10% in blood antisera testing) [1]. More recent studies on chimerism were able to achieve detection of lower concentrations of foreign DNA by applying molecular testing methodologies such as fluorescent labeling and nucleic acid amplification to enhance sensitivity. Techniques like real-time quantitative PCR and flow cytometry have provided essential tools to better examine this phenomenon in low concentrations of heterogenous samples.

To understand chimerism in the context of disease or other associated health disorders, studies applying the techniques for detection of chimerism have investigated this phenomenon in both animals and humans. Developmental disorders have been a consideration for chimerism research due to timing of transplacental exposure to foreign cells and sensitive nature of fetal development. Studies in humans as well as studies of nonhuman mammals including primates

and murids, have provided evidence to support the occurrence of pregnancy-related microchimerism in both the mother and offspring. Furthermore, chimerism between twins can be found in the peripheral blood of cattle with freemartin syndrome, a disorder resulting in an infertile female [2]. The freemartin phenotype in cattle has been described as the result of anastomoses of placental vasculature resulting in intrauterine transfusion between dizygotic opposite-sex twins [3]. Peripheral blood exchange between the developing twins also results in female exposure to male anti-Müllerian hormone during early development contributing to the pathogenesis of freemartinism. In addition, the female co-twin is exposed to male cells resulting in XX/XY (two cell populations with different sex chromosomes) chimerism [4]. Interestingly, disorders of sexual development with similarity to the freemartin phenotype have also been described in several other mammalian species [5]. In addition, reports of human disorders of sexual development have revealed XX/XY karyotypes in cases of true hermaphroditism. I conclude that despite numerous hypotheses presented, there is still much unknown about the mechanisms of mixed cell populations in the etiology of true hermaphroditism in humans, partially due to the rarity of the condition.

Further studies have examined chimerism in relation to complex disease such as autoimmune disease and cancer. The evidence of transplacental chimerism as a consequence of pregnancy in females suggests that the chimerism may contribute to diseases found to be more prevalent in women. Numerous autoimmune diseases have a female predilection, in addition autoimmune conditions including systemic sclerosis and Sjögren's syndrome have a similar clinical presentation to chronic graft versus host disease [6-8]. These observations have supported the hypothesis that exposure to foreign cells in chimerism, primarily as a result of pregnancy, is a component of autoimmune disease etiology. Studies reported in the literature have detailed detectable chimerism in association with several rheumatic and thyroid autoimmune diseases. Similarly, juvenile autoimmune conditions including juvenile idiopathic inflammatory myopathies and type 1 diabetes also demonstrate associations with detectable maternal microchimerism. I outline how the commonality of chimerism prevalence in several studies of diverse autoimmune conditions continues to encourage further investigation of this pathogenic hypothesis for chimerism in health. However, in my review I also found that there are observations from other studies which contradict this pathogenic theory of

allogeneic cell populations in chimerism, primarily in studies of chimerism in cancer diagnoses. Studies investigating prevalence of male microchimerism in the tissues of women with papillary thyroid cancer have indicated a more attenuated disease and better outcomes when compared to women with the disease but without male microchimerism [9]. Similarly, findings from studies of breast cancer reveal reduced prevalence of chimerism in the normal breast tissue of women with invasive breast cancer diagnosis [10] and in peripheral blood of breast cancer cases. As a result, I conclude that these findings in cancer patients with chimerism suggests that these cells could also serve as sentinels mitigating tumor progression, thereby improving prognoses and clinical outcomes.

The outcomes of the review in Chapter 2 demonstrate that the implications of chimerism for human health may be remarkably complex. This underlying complexity is illustrated by the diverse presentation of acquired allogenic cells in unique clinical scenarios which are ostensibly mediated by the tissue microenvironment and host immune response to foreign cellular material. It also shows that there is a need to better understand the cellular mechanisms and interactions between the cell populations which are involved in producing the protective and pathogenic effects implicated in studies of chimerism. Furthermore, longitudinal studies of chimerism are needed to better understand the timing of chimerism onset, changes to microchimerism burden over the lifetime, and dynamics of chimerism concentration during critical times of biological, developmental, and behavioral change.

CHAPTER 3: MALE MICROCHIMERISM AND MAYER–ROKITANSKY–KÜSTER–HAUSER SYNDROME

Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome is a disorder of sexual development defined by failure of Müllerian structure development, resulting in vaginal and uterine agenesis. MRKH syndrome is considered rare, affecting approximately 1 in 5000 females [11], presents clinically with a normal female (46,XX) karyotype and occasionally may present with additional renal or skeletal malformations [12]. The MRKH syndrome presents an analogous clinical picture to the freemartin phenotype described in cattle and other mammals. Indeed, human twins can also exhibit intrauterine cell trafficking [13, 14] and a case report indicated XY cells from a male co-twin in the blood of a girl with aplasia of the uterus [15]. Based on the similarity between these biological conditions, we proposed that a

similar mechanism of intrauterine exposure to anti-Müllerian hormone (AMH) results in the pathogenesis of MRKH syndrome and that this would be detectable through persistent male microchimerism in adult patients.

A study is presented in Chapter 3 with the purpose of examining the prevalence of male microchimerism as an indicator of prior exposure to male blood during development in females with MRKH syndrome. For this observational case-control study we recruited women diagnosed with MRKH from the Dutch patients' association of women with MRKH and included controls with no history of cancer or pregnancy from the Amsterdam University Medical Center. For both qualitative and quantitative assessment of male microchimerism, I developed and implemented a real-time qPCR assay to measure the total genome mass and male genome mass using assays which target the *β-globin* (cellular control) and *DYS14* (male target) genes respectively.

Among the women tested, 11.3% had quantifiable measure of *DYS14*, indicative of male microchimerism. However, the prevalence of male microchimerism was significantly greater in the control group compared to the MRKH. Similarly, the mean concentration in male genome equivalents (GEq) was significantly higher in the control group compared to MRKH cases. Contrary to our hypothesis, these findings do not lend support to male blood exposure as a source of AMH during early embryological development and as a factor promoting MRKH pathogenesis. As the significant difference favors the control group, this presents an intriguing finding in relation to the prevalence of human microchimerism. A 17.2% prevalence of male microchimerism among control participants with no history of pregnancy is in line with previous findings of 13.6% in adolescent girls [16] and 13.3% in healthy null gravid women (that is, women who were never pregnant) [17].

To investigate these findings of male microchimerism in more detail, I also explored additional sources of chimerism that have been proposed, which include older male siblings [16, 18], sexual intercourse [16], and blood transfusions [16, 19]. To consider these additional sources, we performed post-hoc analysis of male microchimerism status relative to having an older brother, blood transfusion, and reported history of sexual intercourse. None of these proposed sources of male microchimerism explained the difference in the distribution of individuals based on detectable male

microchimerism, although there was a small proportion of individuals who tested positive for male microchimerism in each group.

A significantly lower prevalence of measurable male genome in women with MRKH syndrome in comparison to a healthy control population refutes the hypothesis of increased male microchimerism in women with MRKH as a result of exposure to male blood in early gestation. The significantly greater prevalence of male microchimerism in the control group may imply that there were differences in exposure to male cells through unrecognized pregnancies or additional environmental exposures in this group. This study examined male microchimerism in a cohort of adult women who may have been exposed to several other sources of male cells throughout their lifetime. It is likely that microchimerism is a consequence of several unique exposures to sources of male genome including transfusion history, sexual intercourse, and older male siblings in this population. The fact that I could detect male microchimerism in my study populations provides additional support for male genome prevalence in null gravid women and emphasizes the need for future studies to continue unraveling the conditions that promote persistent chimerism in humans.

CHAPTER 4: MALE MICROCHIMERISM PREVALENCE IN TWIN FAMILIES

The bi-directional exposure to allogeneic cells during pregnancy for both the mother and fetus has invoked many questions regarding exposures promoting persistent chimerism. There is potential for all humans to have early life exposures in utero that facilitate the onset of chimerism, while pregnancy in women is an additional known source later in life [20-22]. This chimerism between generations has been supported by measures of male cells in studies of maternal blood following pregnancy with a male fetus [23, 24]. Multiples also present a unique potential for intrauterine exposures with their co-twin, resulting in chimerism of cell populations within the same generation. In a seminal study, van Dijk et al. reported 8% prevalence of blood chimerism in 552 dizygotic twin pairs via detection of red blood cell antigens [14]. These findings have indicated several potential sources for donor cells within families, including mothers, children and siblings as well as co-twins in cases of dizygotic twins. There is interest in exploring transmission patterns of chimerism in families to understand the most substantial contributors to persistent chimerism.

In Chapter 4, I expanded upon the study of twins introduced by van Dijk et al. [14] by investigating patterns of microchimerism prevalence in two generations with a twin-

family study design. Selection criteria for inclusion in this study were Netherlands Twin Register adult female participants who had donated peripheral blood samples and included monozygotic twins (MZ), dizygotic same-sex twins (DZss), dizygotic opposite-sex twins (DOS), in addition to the mother and a non-twin sister of the twins. The resulting 446 participants from 152 families were tested for male microchimerism using the *β-globin* and *DYS14* qPCR assay I developed and also applied in the study described in Chapter 3.

In the samples tested, I measured detectable male microchimerism in 120 out of the 446 women (26.9%). We first examined the findings in groups of participants based on their relationship to the twins of the pedigree. In the subset of samples that revealed detectable male microchimerism, participant groups presented with similar median concentrations of male genome. Similar findings of male microchimerism prevalence were found for females from same sex DZ (23.5%) and opposite sex twins (27.4%), and singleton sisters (25.4%), while MZ twins (16.3%) presented with the lowest overall prevalence. Our primary hypothesis for this study was that the shared environment of DOS twins in utero would be related to increased prevalence of male microchimerism in the female co-twins, but there was not a significant difference in detectable male microchimerism in females of DOS twins compared to DZss twins. Notably, the mothers of twins group presented with the greatest prevalence of male microchimerism (38.7%). Further investigation revealed a positive relationship between age and microchimerism.

Our study also aimed to understand how additional potential sources of male microchimerism may influence the occurence of detectable male genome in the study participants. I examined participants with male offspring, an older brother, or both which did not reveal any significant differences compared to other women for prevalence of male microchimerism. I investigated the influence of microchimerism across generations and found that offspring from mothers with detectable male microchimerism generally have a greater prevalence of male microchimerism (26.8%) compared to the offspring of mothers with no detectable chimerism (17.6%). I also examined the correlation between MZ twin pairs and first-degree relatives which did not reveal any significant differences based on familial resemblance.

The findings from this study provide important context to the current understanding of chimerism transmission within families. Microchimerism appears with notable

frequency since just over a quarter of females in this population had detectable male microchimerism in their peripheral blood samples. Despite our hypothesis due to the shared timing and proximity to their male sibling in utero, female co-twins of DOS twins are not at any greater risk of presenting with male microchimerism. The study of several pedigrees has resulted in an examination of two generations for chimerism to understand the role that age, time, and additional life experiences may play in the occurrence of microchimerism. Comparison of participants in this twin pedigree study revealed age as a significant factor in microchimerism prevalence. Additionally, outliers with increased male microchimerism concentration revealed a significantly lower age compared to other male microchimerism positive participants. These findings suggest that persistent male microchimerism may be optimally preserved at a low concentration, as I observed in the majority of microchimerism positive samples. The outcomes of this study may be extended to further extrapolate how additional experience and changes in biological homeostasis with increasing age may facilitate long-term risks associated with prolonged exposure to foreign allogeneic cell populations.

CHAPTER 5: DETECTION AND SURVEILLANCE OF SARS-COV-2

When the coronavirus disease 2019 (COVID-19) pandemic became a concern for local public health, my experience with developing the assay for the detection of male microchimerism in female samples directly informed the development and methodology of a molecular diagnostic assay for direct detection of SARS-CoV-2 in upper respiratory samples. The Covid-19 test which we thus developed in the laboratory at the Avera Institute for Human Genetics (AIHG) was implemented in routine clinical diagnostics for the Avera Health system as samples were delivered to AIHG from across the region for analysis.

In Chapter 5, I utilized the results for nearly 75000 patients tested at AIHG in 2020 using the diagnostic real-time reverse transcription PCR (RT-PCR) assay to determine the additional epidemiological value for the monitoring of local dynamics and community transmission.

While the test results for clinical diagnostic application were reported as exclusively qualitative results (e.g. detected or not detected) the real-time RT-PCR test also produces cycle threshold (Ct) values, a measure inversely proportional to the amount of template material introduced into the reaction. Our primary aim here was to

investigate how demand for healthcare services can be predicted by changes in Ct value in the community over time. We found that the mean weekly Ct values produced at AIHG were not highly predictive of epidemiological measures during the reference week (that is, the week in which the samples were tested) but were most predictive for hospitalization count two weeks following testing. Similarly, the AIHG mean Ct values per week optimally predicted patient death count in South Dakota with a four-week delay after testing.

In general, the symptoms of COVID-19 are similar to other respiratory diseases which includes symptoms like cough and fever along with more unique symptoms including ageusia (loss of taste) and anosmia (loss of smell). However, there is significant variation in both the clinical presentation and severity of COVID-19 between individual patients, from asymptomatic infection to severe disease and death. To contribute to the understanding of clinical variation between patients, we investigated variation in symptom presentation over a three-week span around the testing date. The symptoms reported by each patient were partially explained (total variance explained = 50.0%) by three factors which could be categorized based on symptom groups, including symptoms common to respiratory viral infection, sensory symptoms, and gastrointestinal symptoms. Patients displaying each of the symptom groups were examined for differences in age, Ct value, and timing of symptom relative to collection. Interestingly, individuals who present with sensory symptoms were generally younger with samples producing a lower Ct value.

The emergence of SARS-CoV-2 variants of concern (VOC) in late 2021 prompted AIHG to develop and integrate a genomic surveillance program to monitor the variants present in the testing population. I further inquired if Ct values of unique variants differed in the population due to biological differences in transmissibility or virulence. The transition from the ancestral strain of SARS-CoV-2 to strains presenting with N501Y (presumably Alpha) in the spring of 2021 did not reveal any significant variation in mean Ct value. In contrast, the Omicron VOC had a significantly lower mean Ct compared to Delta during transition in the winter of 2021. These findings imply that observed differences in Ct values of SARS-CoV-2 variants within a community spanning a short window of variant transition may be reflective of viral dynamics within the community, in contrast to exclusively a product of biological differences of SARS-CoV-2 variants.

In conclusion, Chapter 6 illustrates how low-level detection of genomic biomarkers facilitated the development of a diagnostic assay in response to the COVID-19 pandemic. Changes in the Ct values of a community reflect changes in the transmission of the virus in the testing population and has implications in associated hospitalizations and deaths associated with infection. The findings described in this chapter further support the examination of real-time PCR data from routine screening for highly transmissible viruses to provide real-time data on the dynamics of local community transmission.

GENERAL DISCUSSION

This thesis details investigations on the exposure and presentation of small populations of non-inherited genetic material from foreign sources. First, I examine the chimerism phenomenon as a naturally occurring persistent exposure to foreign genetic material in humans, discussing current details on the prevalence and clinical implications. Next, I detail the diverse and growing applications of molecular testing and nucleic acid quantification. Lastly, I conclude by providing insight into the future of molecular quantitative measures in biomedical research and healthcare.

CHIMERISM AND HUMAN HEALTH

The occurrence of chimerism in humans has been under investigation for several decades to better understand the etiology of the condition as well as the pathological implications. A product of exposure to cells derived from another unique zygote, the chimera has presented a conundrum for health sciences. Cases of chimerism continue to go unnoticed and are only detected through investigation of testing discrepancies. While blood type discrepancies in both animals and humans provided the first indication of chimerism as a biological phenomenon, further discrepancies observed in forensics and parental (paternity/maternity) testing have illuminated the relevance of this condition in examining biological relationships. In one reported case, a baby was born with an incongruent blood type (AB+) compared to the parents (A+ and A-) and the baby shared 25% DNA homology with the father, consistent with an uncle relationship [25]. Humans are diploid organisms, thus genetic similarity between parent and offspring is expected to be 50% as half of the alleles are inherited by each parent. Further investigation confirmed the father to be a chimera with a

minor genome of a vanishing twin. Other cases have investigated women whose genetic profile did not correlate with that of their children, only to later discover that they are tetragametic chimeras who gave birth to children with half of their genetics obtained from the mother's vanishing twin [26, 27].

Nonetheless, these extreme cases of tetragametic chimerism are generally rare and the most prevalent cases of chimerism have been found to present in the recipient with a small concentration of the acquired allogeneic minor cell population. Exchange of cell populations across the placenta during pregnancy can result in both fetal and maternal microchimerism. Furthermore, studies in twins have revealed chimerism to be a generally common occurrence. Pivotal research revealed 8% chimerism in dizygotic (DZ) twins [14] and several reported cases of DZ twins as a result of artificial reproductive technology presenting with monochorionic placentas [13]. This emphasizes the growing need to further understand the biological implications of this condition both in twins and the general population. Studies of both null gravid and adolescent females have continued to provide evidence that history of pregnancy is not fundamental to becoming a carrier of male microchimerism [16, 17]. As I have shown, further evidence of this comes from study of adult females with no history of pregnancy who present with a significantly greater prevalence of male microchimerism compared to MRKH cases. The mystery behind the source of these foreign cell populations continues to plague researchers seeking to understand the mechanisms of chimerism epidemiology and associated clinical implications. Despite efforts to distinguish the different possible sources of male microchimerism in female participants there has been challenges in tracing a clear common pattern of inheritance.

Without a clear pattern of inheritance, the variance of natural microchimerism in the population may have similarities to variance observed in complex diseases. The field of genomics research includes untangling the complex relationship between genetics and environment in complex human diseases, often by considering the relative contributions of genetics and environment in the explanation of trait variation. By examining concordance in twins, assuming equal environments, a trait with greater similarity measured between MZ twins in relation to DZ twins implies that there is a proportion of the overall variance for the trait that is heritable. Additionally, both MZ and DZ twin pairs often share very similar environmental conditions during development as they are commonly reared in the same living

conditions referred to as common environment, making DZ twins the ideal 'control' group when interpreting the resemblance of MZ twin pairs for a trait. Individual twins will also have experiences and exposures unique to each co-twin. These unique environments are observed within the classical twin study by observing differences within MZ twin pairs.

Microchimerism as a condition resides at a unique intersection of the genetic and environment categories of variance, the condition itself may be dependent on both the genetic relationship between two cell populations as well as the environment which initiated the initial exposure and promoted persistence. Examining twin and their family members across two generations, I did not discover a significant difference in chimerism prevalence or concordance among MZ and same-sex DZ twins. Although this suggests at best a limited heritability for chimerism, the more complex relationship of the genetics between host and donor cells requires further study and larger sample size to establish the genetic contribution to the development and maintenance of persistent chimerism. Considering a possible liability threshold model for persistent chimerism, these genetic and environmental factors may collectively contribute to the individual's liability for allogeneic cell populations to graft and persist. Similarly, exposure to small concentrations of donor cell populations in the host may induce a multitude of different biological responses with diverse implications on disease liability, contributing to the complexities around associations with various pathological manifestations and diversity in their clinical presentation.

Prior to the pedigree study in Chapter 4, researchers had yet to examine this variance in chimerism as a phenotype. Interestingly, a study examining survival in Danish women aged 50-64 years found a remarkably high 70% prevalence of male microchimerism [28], compared to 38.7% I observed among the mothers of twins with a median age of 59. Studies of population structure examine genetic variation to explain genetic differences among a population and one such study in Denmark found substantial homogeneity among the Danish population irrespective of geographic location [29]. Reduced population structure as measured by genetic variation may provide an advantage for persistent microchimerism within a population if it promotes increased allele sharing between host and donor or reduces allosensitization, thus reducing the potential for inducing an alloimmune response. In this light, specific combinations of alleles, most likely related to the immune system and immunoregulation, may facilitate persistent microchimerism. This is supported

by evidence of transplantation success in HLA matching of host and donor for organ transplantation. Indeed, comparison of shared and non-shared alleles between host and donor may provide the most informative evaluation of chimerism, similar to studies of non-inherited genetics [30].

As I have summarized from the literature of chimerism in Chapter 2, there is a predominant interest in establishing the role of chimerism in disease etiology. Many studies of complex diseases have used genome-wide association studies (GWAS) as a study design to identify genes associated with particular traits. This design has been used extensively in over 5800 publications producing over 390000 associations to date (https://www.ebi.ac.uk/gwas/). In some studies, with very large sample sizes, genome-wide association studies (GWAS) and meta-analyses for complex phenotypes have been able to establish numerous significantly associated variants across the genome. Only recently, the largest GWAS meta-analysis on height found 12111 significantly associated genetic variants in 7209 associated loci which were able to recover the entire heritability [31]. While the heritability for height could be completely recovered, the heritability of complex diseases may have additional underlying complexities contributing to variance. It Is possible that examining the relationship between host and donor genetics may help explain some of the missing heritability in complex diseases. I propose that the interaction between allogeneic cells with different genotypes will result in host and donor dynamics which contribute to disease pathogenesis. While the effects of chimerism on the individual are environmental as they reside outside of an individual host's genome, there are indeed genetic and host-donor genetic interactions which may contribute to lasting effects in individuals with pathological implications as observed in studies of chimerism in human disease. Similar to the GWAS hypothesis of "common disease, common variant" [32], I hypothesize that the seemingly high frequency of chimerism would also behave as a common environmental variable in complex diseases.

In Chapter 4, the study of chimerism among twin pedigrees revealed an incomplete concordance of male microchimerism between MZ twins. The observation of incomplete concordance in MZ twin pairs is a common phenomenon in disorders studied in MZ twins, rarely demonstrating a concordance greater than 50% [33]. As such, environmental differences as a result of prior exposure to allogeneic cell populations in chimerism may further contribute to the variation observed in some complex diseases via either pathogenic or protective mechanisms. If male

microchimerism is transferred through the mother from another distant male relative, such as an older male sibling, the offspring would require initial exposure to these male cells to facilitate microchimerism onset. With the small population of cells exchanged, it is likely that it is not a universal occurrence for offspring to be exposed to the same cell populations. In addition, further studies of transplantations have observed better outcomes in recipients when they receive tissues with HLA incompatibility matching their non-inherited maternal HLA antigens for transplantation compared to matching the recipient's non-inherited paternal HLA antigens [34, 35]. As HLA genes encoding the HLA antigens are polymorphic and comprised of many alleles, tolerance of non-inherited maternal antigens further expands the potential alleles that one may be exposed to that may successfully graft and persist. This could further explain how a homogeneous population may increase prevalence of microchimerism. Immune tolerance may be a key regulator of microchimerism prevalence within a population, largely dictated by the amount of genetic variation.

Chimerism is a condition which has a long history of being induced via medical interventions where tissues and even entire organs that are diseased or failing can be replaced via the tissues or organs of another individual. These processes of transplantation and transfusion have been well studied for implementation in clinical practice since the early 20th century. In response to the challenge of transplant rejection, a pioneer in transplantation, Erwin Payr, presented the theory that transplantation could be improved "by artificially reducing the biochemical difference between individuals" [36]. This laid the groundwork for the study of the complex interaction between host and donor biology which enhanced transplantation success and led to the practice becoming established as a treatment option in the 1960s.

During Georg Schöne's research into immunology to treat cancer in the early 20th century, he first described "transplantation immunity" which caused the rejection of transplants via the body's own immune system [36]. Indeed, it has also been investigated if donor cells constituting chimerism are involved in the biology surrounding cancer [37]. Studies in women have found chimerism to be less prevalent in both the peripheral blood and breast tissue of breast cancer patients compared to healthy controls, which has supported the theory that chimerism may be protective [10, 38]. In addition, women with pregnancies complicated by preeclampsia, which

increases cell trafficking between mother and fetus, were found to be at a reduced risk of cancer mortality but had an increase in cardiovascular mortality [39]. Breast cancer risk has also been demonstrated to be reduced in mothers of twins [40, 41] which may be a consequence of high prevalence of microchimerism in mothers of twins as shown in the study described in Chapter 4, though a definitive mechanism has yet to be determined. This duality in relation to health and disease suggests that chimerism may play the role of a biological double agent. These cells appear to instigate a response from the host immune system resulting in autoimmune disease and provide protection through suppression of cancerous cell proliferation. My conjecture is that the protective effect of chimerism on cancer in certain tissues may be maintained by keeping the immune system at a heighted state of activity by low levels of foreign donor cells inadvertently interacting with the host immune system. As such, the immunosurveillance activity of the immune system may aid in the culling of pre-cancerous cells before the onset of disease.

This theory would further support the pathological hypothesis of chimerism as a contributor to the etiology of autoimmune disease. A recent study found that women with microchimerism containing a rheumatoid arthritis protective 5-amino acid motif (DERAA) were 17.1 times more likely to develop rheumatoid arthritis, indicating a pathogenic role for exposure to the foreign antigen [32]. By maintaining a persistent low-level of chimerism in the host, it is likely that continued environmental exposures to pathogens as well as biological changes associated with time and aging may indeed contribute to a change in tolerance to the graft cell population. Perhaps when the cell population crosses a theoretical threshold of prevalence, the host immune system becomes actively responsive to the interlopers and mounts an adaptive immune response against these allogeneic human cells. Due to the structural similarity to other host cells, particularly with cells obtained from close family members, this may induce an autoimmune response to eliminate the threat of the donor cells in the host. Further investigation of the longitudinal change in chimerism prevalence before and after the onset of autoimmune disease is needed to examine if this is indeed a response targeting the foreign donor cells and to understand the underlying mechanisms of this relationship.

Examination of male cells presents limitations in this context as chimerism is not sex specific but rather defined based on the zygote from which the cells derived. It is of note that the H-Y antigen is a minor histocompatibility antigen present on the nucleated cells of males but not that of females as the gene encoding the antigen is located on the Y chromosome [42]. While this antigen is classified as a minor histocompatibility antigen, it is possible that it may have implications in the triggering of the immune system in cases of autoimmune disease in women. Hence, the detection of male microchimerism in females may provide a stronger indicator for disease risk based on an immune response, despite not measuring all instances of human chimerism in these patients, including XX/XX chimerism. Many studies have utilized the Y chromosome as a unique indicator of chimerism in both humans and animals. The Y chromosome has inherent value as a unique marker capable of distinguishing between cell populations native to the host and those which originate from another zygote. However, this also is a limitation as it only indicates if the cells identified are of male origin and does not provide further context to whom these cells may have originated from. There is need for future work to distinguish the specific source of the donor cells found in a chimera. Development and implementation of such methods will facilitate future investigation of the relationship between host and donor genotypes as well as mechanisms of cell exposure, grafting, and persistence.

MOLECULAR DIAGNOSTICS AND NUCLEIC ACID QUANTIFICATION

Like the bond between twins portrayed in literature and film, clinical and research laboratories present a duality, each serving a unique purpose in delivering scientific data. A research laboratory is rich with the knowledge of and experience in using modern tools for deriving innovative solutions to address complex questions, producing new assays and techniques relevant to specific studies. Meanwhile, clinical laboratories assist in patient assessment and diagnosis of clinical pathology by specializing in the application of reliable high complexity test methods and delivery of patient test results. While the fundamental principles behind the molecular methodologies used in both clinical and research laboratories remain consistent, there are substantial differences related to the application in unique laboratory environments. Research laboratories apply molecular testing methods to investigate the trends and characterization of associated biological mechanisms in research populations for various traits. In clinical laboratories, molecular test data is valued for aiding in disease diagnosis and informing treatment plans for individual patients

based on associations derived from research. In addition, the services provided by healthcare professionals in clinical laboratories are time-sensitive and necessitate optimization of methods developed in research laboratories to meet the high-throughput workflow requirements in the clinical laboratory environment. Together, these two categories of laboratories have a symbiotic relationship, dependent on each other for advancing science in healthcare.

Throughout my career, I have applied molecular methods in both a clinical laboratory setting as a Medical Laboratory Scientist and in a research laboratory setting at the Avera Institute for Human Genetics (AIHG) during my Ph.D. trajectory. I have observed how these scientific tools provide immeasurable value to our understanding of genetics while also inspiring new and diverse applications in the detection of clinically relevant genetic variants and biomarkers. Molecular diagnostics has been fundamental to facilitating modern clinical decisions through increased sensitivity and throughput of advanced testing methodologies to detect biomarkers in the genetic code of various organisms [43, 44]. The development of diagnostic tests to detect infectious diseases has revolutionized the landscape of clinical microbiology by streamlining testing workflows, enhancing sensitivity, and reducing time to diagnosis. As clinically relevant microbes are often found at low concentrations in samples collected from patients, these tests rely on the detection of unique biomarkers specific to the genome of the organism of interest. The ability to rapidly detect as little as a single copy of nucleic acid material from a pathogenic organism in a background of human material within a sample has had profound implications on healthcare quality when early detection is vital.

During my Ph.D. training, I was able to contribute directly to the relationship between research and clinical laboratories. After I developed and implemented PCR assays for the study of male microchimerism in females, there was growing demand for molecular diagnostic solutions for the detection of SARS-CoV-2 infection for clinical diagnosis and to assist public health officials in controlling spread of the virus within the surrounding communities. The knowledge of highly specific tests for the detection of low concentrations of biomarkers in the study of microchimerism was directly translatable to the detection of the SARS-CoV-2 genome. In a coordinated effort, the research staff at AIHG applied this prior experience to develop a RT-PCR test method for the detection of low concentrations of viral genome in patient upper respiratory samples. In addition, clinical laboratory personnel provided insight for

validation plans, optimized the testing workflow for high throughput testing, and coordinated logistics for delivery of both testing samples and results. As a result of this combined effort, the AIHG laboratory produced molecular diagnostic tests that resulted in delivery of hundreds of thousands of test results for patients. Having previous experience with PCR molecular genetic tests for low concentrations of minor genomic populations was essential for the rapid development, validation, and integration of this diagnostic test when patients and physicians needed a solution to reduce time to diagnosis.

Similarly, clinical laboratories are specially focused on the timely delivery of clinically relevant patient results and are generally not equipped to address additional inquiries into the genomic surveillance of emerging pathogens. Epidemiologists rely on available data to produce meaningful information to trace the dynamics of pathogens within communities. When this information relies on the test data from molecular diagnostics to evaluate changes over time, the use of test positivity as a single dimension of the data provides limited perspective. Extensive expertise in sequencing techniques for research applications and samples available on-site for SARS-CoV-2 testing made it possible for AIHG to evaluate samples for virus variants and examine population changes in variant prevalence while minimizing the delay in data availability. This produced clear and timely evidence of genetic changes to the virus biology within the community. Having sequencing capabilities available as a part of the research repertoire at Avera Health provided a further advantage for the healthcare system to monitor the changes in local SARS-CoV-2 variant composition in the testing population. Without timely insight into the trends and dynamics of transmissible pathogens within a community, it is increasingly challenging to both adequately manage policy changes to address public health and manage the resources needed for testing and treatment. However, there is inequal access to genomic surveillance across the globe, leaving some with significant lag time in data to anticipate changes in the virus dynamics and associated healthcare needs [45]. By increasing the availability of real-time measures of viral dynamics in communities with limited resources, it likely possible to improve mitigation practices and public health management.

There have been several hypotheses which consider that the measure of SARS-CoV-2 virus particle burden (i.e. viral load, the concentration of viral particles) in an individual may be closely linked to their disease severity [46, 47]. This is reasonable given that other viruses such as influenza and HIV have frequently demonstrated an

association between viral load and disease severity [48-52]. Furthermore, others have suggested that differences in the biology of virus variants may result in changing viral loads, with the viral load increasing as a product of improved virus transmissibility [53, 54]. However, recent evidence has suggested that transitions in SARS-CoV-2 variant prevalence may not be linked to clear changes in the concentration of virus present in collected samples [55]. Amplification of the material of interest via real-time PCR not only provides a qualitative test of the virus genome in the sample but also can provides quantitative insight into the relative concentration of foreign genomic material in the original sample. This information has been applied in diagnostic tests of blood-borne viruses such as hepatitis and HIV viruses collected in blood specimens [51, 52, 56, 57]. Quantitative measures have greater considerations compared to qualitative tests including pre-analytic (e.g. collection quality) and analytic (e.g. test performance) variables [58]. Variation in these categories may produce erroneous results and inconsistency in clinical utility of the data generated for individual patients. For example, homogenized liquid samples such as peripheral blood will provide a generally consistent sample whereas swab collections may have substantial variation in collection quality (e.g. difference in deep nasopharyngeal and anterior nares collections as well as differences in the vigor and tenacity of the collection). However, despite variation that may present in individual samples, I detail in Chapter 5 that when quantitative measures derived during routine testing in the population are taken together and examined for a cumulative change in virus magnitude it is possible to observe changes related to the testing population. Adding quantitative measures as an additional dimension to assess data available from clinical laboratories may provide insight into changes in community disease burden and associated public health needs.

The quantitative data produced by real-time PCR assays for SARS-CoV-2 has yet to support their use in clinical decision-making. However, despite the caveats to these measures due to collection method and analytical variation, Ct value has been applied as a criterion for informing return to work policies. The findings I present in this thesis provide a basis for examining magnitude of change in a population as a measure for laboratories to monitor local change in transmission dynamics. Given the location of South Dakota in the Midwest and a lower population density compared to the coastal states, the timing of surges in case volume did not always reflect the rest of the nation likely due to spatiotemporal variation between geographic

regions of the nation [59]. However, despite the substantial variation in symptoms among the patients diagnosed with COVID-19, examination of variance in symptom presentation was able to reveal remarkable commonalities among patients with different symptom groups in the patient population. Further studies examining larger numbers of patients may be able to provide valuable information about the symptom presentation and population specific variation in clinical presentation among patients, better equipping public health agencies to address the specific needs of the community. Thus, by examining data as it relates to specific testing populations, we can produce more precise observations on virus transmission and reduce reliance on generalizations by adequately accounting for variation in environment and population characteristics. The means to examine this information at the local level would be valuable for providing specific targeted information to leaders in public health as it relates to their immediate jurisdiction.

FUTURE PERSPECTIVES

The knowledge and application of genetics in both health and disease continues to expand at an appreciable rate. Through continued advancements in the understanding of human genetic diversity and complex traits, the field of genomics now extends beyond the detection of monogenic Mendelian diseases. Further applications of genomic techniques have produced opportunities to better address healthcare needs by developing sensitive diagnostics for the identification of pathogens, and by improving time to diagnosis and treatment strategies. While the field of genomics has found its place as a gold standard in many clinical applications, innovation in the field continues to pursue opportunities to improve clinical care.

An emerging field of clinical application for genetics in clinical care is the use of a patient's unique genetic composition to make informed decisions, aptly labeled as personalized medicine. Two recent innovations to enhance personalized medicine include pharmacogenomics and polygenic risk scores. Through pharmacogenomics a patient's genotype can be used to establish their molecular proficiency for metabolizing specific medications to determine which medication and dose may be most appropriate for their care [60]. This is expediting the process for getting patients on the most appropriate medication, reducing unnecessary stress and suffering for the patient. In addition, there is growing interest in the application of genotypes to provide a predictive measure of an individual's genetic predisposition to a given trait

through polygenic risk scores. This quantitative measure of genetic risk combines the effect size of associated alleles with an individuals inherited genotype to establish the individual's risk for developing the condition relative to the population [61]. As a natural progression to discussing family history alone, polygenic risk scores may provide physicians with an objective measure of genetic risk for complex diseases including cancer and heart disease and encourage increased preventative screening for those at increased risk [62]. This provides a particular advantage for individuals who may not have knowledge of predispositions from their biological ancestors, with polygenic risk scores providing a valuable means of understanding their genetic risks [63]. Integration of additional genomics test methods into routine practice will provide tools to be able to provide the best care specific to the needs of the patient.

However, as studies in numerous traits and diseases have illustrated, genomics alone does not provide a comprehensive explanation of variation in clinical presentation and outcomes among patients. My work in this dissertation contributes new observations of the chimerism phenomenon, supports findings of microchimerism in null gravid women, and illustrates microchimerism persistence related to age. I highlight in this thesis the importance of exposures to foreign genomes in our environment, presenting another layer of complexity to our understanding of health and disease. Both the principal exposure and specific combinations of major and minor cell populations in an individual may have broad implications for a variety of health-related traits. As previously demonstrated by observed significance in a study of non-inherited parental alleles, genetics can also contribute to environment variance [30] which in chimerism may present as non-shared alleles in the allogeneic minor cell population. The genetic composition of both the host and pathogen may be able to explain variation in disease pathology between individuals. Likewise, exposure to foreign genetic material in the environment may facilitate diverse biological implications in complex traits as a result of persistent exposure to non-inherited alleles in chimerism. Future investigation of exposure to foreign genetics in complex traits should examine associations of both non-inherited maternal alleles in offspring and non-shared partner alleles in women to provide insight into regions of genetic discordance contributing to disease variance. I perceive the isolation of allogeneic minor cell populations and examination of donor genetic characteristics in cases of chimerism as a critical step for investigations of non-shared alleles across the genome and defining the mechanisms associated with health

and disease. The added information by small populations of allogeneic cells may be able to provide additional information on the complexities of allograft tolerance or rejection. One such study examining host and donor pairs in hematopoietic stem cell transplantation determined that a female sibling donor with a history of pregnancy is at a lower risk for graft versus host disease compared to male unrelated donors [64]. With added information pertaining to minor histocompatibility antigens and triggers resulting in immune responses we may be able to further improve patient health care through transplantation medicine and evaluation of microchimerism associated pathologies.

The dynamics of SARS-CoV-2 and chimerism may display similar common relationships regarding characteristics of transmission and time. Both of these phenomena are of interest for understanding implications in individual health and transmission in a population. Such studies of health and transmission require examination through different levels of magnification, including both the study of cell concentration in an individual as well as groups of affected people in a population. A common observation through these studies of genomic concentrations is that timing of sample collection is of critical importance and must be considered for future investigation of these phenomenon to derive underlying complexities. Current studies of chimerism rely on implied persistence associated with disease pathology which induces complexities when using detection of chimerism as a biomarker for prior exposures. Utilization of microchimerism as a biomarker, as applied in Chapter 3, is valuable as a long-term marker of prior exposures; however, this only provides an examination of persistent chimerism and does not account for effects from chimerism exposures which may have been transient. In study of two generation twin families, I was able to observe age-related association in chimerism prevalence but no discernable common source for male cells. I believe that longitudinal studies examining changes in allogeneic minor genomic concentrations over time will provide insight into biological changes related to populations of foreign genetic materials. This introduces an inherent challenge for researchers to devise studies to address these outstanding questions. Studies in younger cohorts are needed to illustrate the implications and dynamics of host response in early life, especially in chimerism as it relates to the juvenile autoimmune diseases described in Chapter 2 and developmental disorders as in Chapter 3. Increased importance in well-organized research programs with meticulously kept demographic and phenotypic records with longitudinal specimen collection and biobanking will

provide invaluable tools for understanding contributors to human health across the lifespan. For this reason, future studies of chimerism in established twin registers will be valuable as these have often carefully collected biological samples from twins and their family members as well as curated datasets of phenotype data from follow-up questionnaires which can be used to investigate potential sources of chimerism and association with future disease pathogenesis.

The landscape of scientific inquiry into human genetics continues to shift as more information and knowledge become attainable. Quantitative measures of genetic variation have become a primary focus of many scientists and clinicians to provide better tools for understanding progression of human health and disease. Through the collective study of the general population, disease cohorts, and nuclear families it is possible for researchers to extrapolate meaningful conclusions on health and disease from the complexities of the human genome and further understand the implications of our unique environments. Future studies will rely on multi-omics approaches to investigate the vast network of relationships driving human health, including genomics, transcriptomics, metabolomics, and proteomics. Investigation of these high-dimensional studies between genetically different cell populations will be fundamentally complex but potentially rich with information pertaining to underlying cellular mechanisms between host and donor. This may provide insight into the variation observed in these studies of chimerism as well as the differences in disease presentation found in COVID-19. Furthermore, epigenome studies will be instrumental in understanding the lasting effect of these exposures. Recent discovery of a unique MZ twin epigenetic signature highlights that events in fetal life can have a lasting influence [65]. Future studies may be able to extrapolate if there are significant epigenetic implications to genetic exposures across an individual's lifespan.

Finding answers to the complexities of human disease will depend on identifying the diverse contributors to disease pathogenesis including the contribution of non-inherited genetics to an individual's environment. The observations made throughout this dissertation further illustrate the prevalence of exposures to foreign genetic material and the potential of quantitation via molecular genomics tools to help understand the characteristics of transmission in the population. Like polygenic risk scores, the potential of quantitative measures includes additional insight which may contribute to developing better screening tools and interventions to improve patient care and overall health.

REFERENCES

1. Dunsford, I., et al., *A human blood-group chimera.* Br Med J, 1953. **2**(4827): p. 81.
2. Lillie, F.R., *The Theory of the Free-Martin.* Science, 1916. **43**(1113): p. 611-3.
3. Owen, R.D., *Immunogenetic Consequences of Vascular Anastomoses between Bovine Twins.* Science, 1945. **102**(2651): p. 400-1.
4. Vigier, B., et al., *Origin of anti-Mullerian hormone in bovine freemartin fetuses.* J Reprod Fertil, 1984. **70**(2): p. 473-9.
5. Padula, A.M., *The freemartin syndrome: an update.* Anim Reprod Sci, 2005. **87**(1-2): p. 93-109.
6. Scaletti, C., et al., *Microchimerism and systemic sclerosis.* Int Arch Allergy Immunol, 2001. **125**(3): p. 196-202.
7. Kuroki, M., et al., *Detection of maternal-fetal microchimerism in the inflammatory lesions of patients with Sjogren's syndrome.* Ann Rheum Dis, 2002. **61**(12): p. 1041-6.
8. Endo, Y., I. Negishi, and O. Ishikawa, *Possible contribution of microchimerism to the pathogenesis of Sjogren's syndrome.* Rheumatology (Oxford), 2002. **41**(5): p. 490-5.
9. Cirello, V., et al., *Fetal cell microchimerism in papillary thyroid cancer: A role in the outcome of the disease.* Int J Cancer, 2015. **137**(12): p. 2989-93.
10. Gadi, V.K., *Fetal microchimerism in breast from women with and without breast cancer.* Breast Cancer Res Treat, 2010. **121**(1): p. 241-4.
11. Herlin, M., et al., *Prevalence and patient characteristics of Mayer-Rokitansky-Kuster-Hauser syndrome: a nationwide registry-based study.* Hum Reprod, 2016. **31**(10): p. 2384-90.
12. Oppelt, P., et al., *Clinical aspects of Mayer-Rokitansky-Kuester-Hauser syndrome: recommendations for clinical diagnosis and staging.* Hum Reprod, 2006. **21**(3): p. 792-7.
13. Peters, H.E., et al., *Unusual Twinning Resulting in Chimerism: A Systematic Review on Monochorionic Dizygotic Twins.* Twin Res Hum Genet, 2017. **20**(2): p. 161-168.
14. van Dijk, B.A., D.I. Boomsma, and A.J. de Man, *Blood group chimerism in human multiple births is not rare.* Am J Med Genet, 1996. **61**(3): p. 264-8.
15. Bogdanova, N., et al., *Blood chimerism in a girl with Down syndrome and possible freemartin effect leading to aplasia of the Mullerian derivatives.* Hum Reprod, 2010. **25**(5): p. 1339-43.
16. Muller, A.C., et al., *Microchimerism of male origin in a cohort of Danish girls.* Chimerism, 2015. **6**(4): p. 65-71.
17. Yan, Z., et al., *Male microchimerism in women without sons: quantitative assessment and correlation with pregnancy history.* Am J Med, 2005. **118**(8): p. 899-906.
18. Dierselhuis, M.P., et al., *Transmaternal cell flow leads to antigen-experienced cord blood.* Blood, 2012. **120**(3): p. 505-10.
19. Utter, G.H., et al., *Transfusion-associated microchimerism.* Vox Sang, 2007. **93**(3): p. 188-95.
20. Dierselhuis, M.P. and E. Goulmy, *We are all born as microchimera.* Chimerism, 2013. **4**(1): p. 18-9.
21. Bianchi, D.W., et al., *Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum.* Proc Natl Acad Sci U S A, 1996. **93**(2): p. 705-8.

22. Gammill, H.S. and W.E. Harrington, *Microchimerism: Defining and redefining the prepregnancy context - A review.* Placenta, 2017. **60**: p. 130-133.

23. Lo, Y.M., et al., *Two-way cell traffic between mother and fetus: biologic and clinical implications.* Blood, 1996. **88**(11): p. 4390-5.

24. Mosca, M., et al., *Correlations of Y chromosome microchimerism with disease activity in patients with SLE: analysis of preliminary data.* Ann Rheum Dis, 2003. **62**(7): p. 651-4.

25. Sheets, K.M., et al., *A case of chimerism-induced paternity confusion: what ART practitioners can do to prevent future calamity for families.* J Assist Reprod Genet, 2018. **35**(2): p. 345-352.

26. Yu, N., et al., *Disputed maternity leading to identification of tetragametic chimerism.* N Engl J Med, 2002. **346**(20): p. 1545-52.

27. *She's Her Own Twin.* 2006 August 15 [cited 2022 June 27]; Available from: https://abcnews.go.com/Primetime/shes-twin/story?id=2315693.

28. Kamper-Jorgensen, M., et al., *Male microchimerism and survival among women.* Int J Epidemiol, 2014. **43**(1): p. 168-73.

29. Athanasiadis, G., et al., *Nationwide Genomic Study in Denmark Reveals Remarkable Population Homogeneity.* Genetics, 2016. **204**(2): p. 711-722.

30. Bates, T.C., et al., *The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families.* Twin Res Hum Genet, 2018. **21**(2): p. 73-83.

31. Yengo, L., et al., *A Saturated Map of Common Genetic Variants Associated with Human Height from 5.4 Million Individuals of Diverse Ancestries.* bioRxiv, 2022: p. 2022.01.07.475305.

32. Kanaan, S.B., et al., *Immunogenicity of a rheumatoid arthritis protective sequence when acquired through microchimerism.* Proc Natl Acad Sci U S A, 2019.

33. van Dongen, J., et al., *The continuing value of twin studies in the omics era.* Nat Rev Genet, 2012. **13**(9): p. 640-53.

34. van Rood, J.J., et al., *Effect of tolerance to noninherited maternal antigens on the occurrence of graft-versus-host disease after bone marrow transplantation from a parent or an HLA-haploidentical sibling.* Blood, 2002. **99**(5): p. 1572-7.

35. Burlingham, W.J., et al., *The effect of tolerance to noninherited maternal HLA antigens on the survival of renal transplants from sibling donors.* N Engl J Med, 1998. **339**(23): p. 1657-64.

36. Schlich, T., *The origins of organ transplantation.* Lancet, 2011. **378**(9800): p. 1372-3.

37. Khosrotehrani, K. and S. Aractingi, *Fetal cell microchimerism in cancer: a meaningful event?* Future Oncol, 2009. **5**(9): p. 1441-8.

38. Eun, J.K., et al., *In situ breast cancer and microchimerism.* Sci Rep, 2013. **3**: p. 2192.

39. Irgens, H.U., et al., *Long term mortality of mothers and fathers after pre-eclampsia: population based cohort study.* BMJ, 2001. **323**(7323): p. 1213-7.

40. Murphy, M.F., et al., *Breast cancer risk in mothers of twins.* Br J Cancer, 1997. **75**(7): p. 1066-8.

41. Ji, J., et al., *Risks of breast, endometrial, and ovarian cancers after twin births.* Endocr Relat Cancer, 2007. **14**(3): p. 703-11.

42. Wang, W., et al., *Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein.* Science, 1995. **269**(5230): p. 1588-90.

43. Biomarkers Definitions Working, G., *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.

44. Biswas, B., *Clinical Performance Evaluation of Molecular Diagnostic Tests.* J Mol Diagn, 2016. **18**(6): p. 803-812.

45. Wilkinson, E., et al., *A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa.* Science, 2021. **374**(6566): p. 423-431.

46. Faico-Filho, K.S., V.C. Passarelli, and N. Bellei, *Is Higher Viral Load in SARS-CoV-2 Associated with Death?* Am J Trop Med Hyg, 2020. **103**(5): p. 2019-2021.

47. Fajnzylber, J., et al., *SARS-CoV-2 viral load is associated with increased disease severity and mortality.* Nat Commun, 2020. **11**(1): p. 5493.

48. Lee, N., et al., *Viral loads and duration of viral shedding in adult patients hospitalized with influenza.* J Infect Dis, 2009. **200**(4): p. 492-500.

49. Granados, A., et al., *Influenza and rhinovirus viral load and disease severity in upper respiratory tract infections.* J Clin Virol, 2017. **86**: p. 14-19.

50. Pronier, C., et al., *Respiratory Influenza viral load as a marker of poor prognosis in patients with severe symptoms.* J Clin Virol, 2021. **136**: p. 104761.

51. Saag, M.S., et al., *HIV viral load markers in clinical practice.* Nat Med, 1996. **2**(6): p. 625-9.

52. Drain, P.K., et al., *Point-of-Care HIV Viral Load Testing: an Essential Tool for a Sustainable Global HIV/AIDS Response.* Clin Microbiol Rev, 2019. **32**(3).

53. Kidd, M., et al., *S-variant SARS-CoV-2 is associated with significantly higher viral loads in samples tested by ThermoFisher TaqPath RT-QPCR.* medRxiv, 2020: p. 2020.12.24.20248834.

54. Davies, N.G., et al., *Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England.* Science, 2021. **372**(6538).

55. Fall, A., et al., *The displacement of the SARS-CoV-2 variant Delta with Omicron: An investigation of hospital admissions and upper respiratory viral loads.* EBioMedicine, 2022. **79**: p. 104008.

56. Grebely, J., et al., *Time to Detection of Hepatitis C Virus Infection With the Xpert HCV Viral Load Fingerstick Point-of-Care Assay: Facilitating a More Rapid Time to Diagnosis.* J Infect Dis, 2020. **221**(12): p. 2043-2049.

57. Poiteau, L., et al., *Evaluation of the Xpert HBV Viral Load for hepatitis B virus molecular testing.* J Clin Virol, 2020. **129**: p. 104481.

58. Zimmerman, P.A., et al., *Molecular Diagnosis of SARS-CoV-2: Assessing and Interpreting Nucleic Acid and Antigen Tests.* Pathog Immun, 2021. **6**(1): p. 135-156.

59. Cuadros, D.F., et al., *Dynamics of the COVID-19 epidemic in urban and rural areas in the United States.* Ann Epidemiol, 2021. **59**: p. 16-20.

60. Roden, D.M., et al., *Pharmacogenomics.* Lancet, 2019. **394**(10197): p. 521-532.

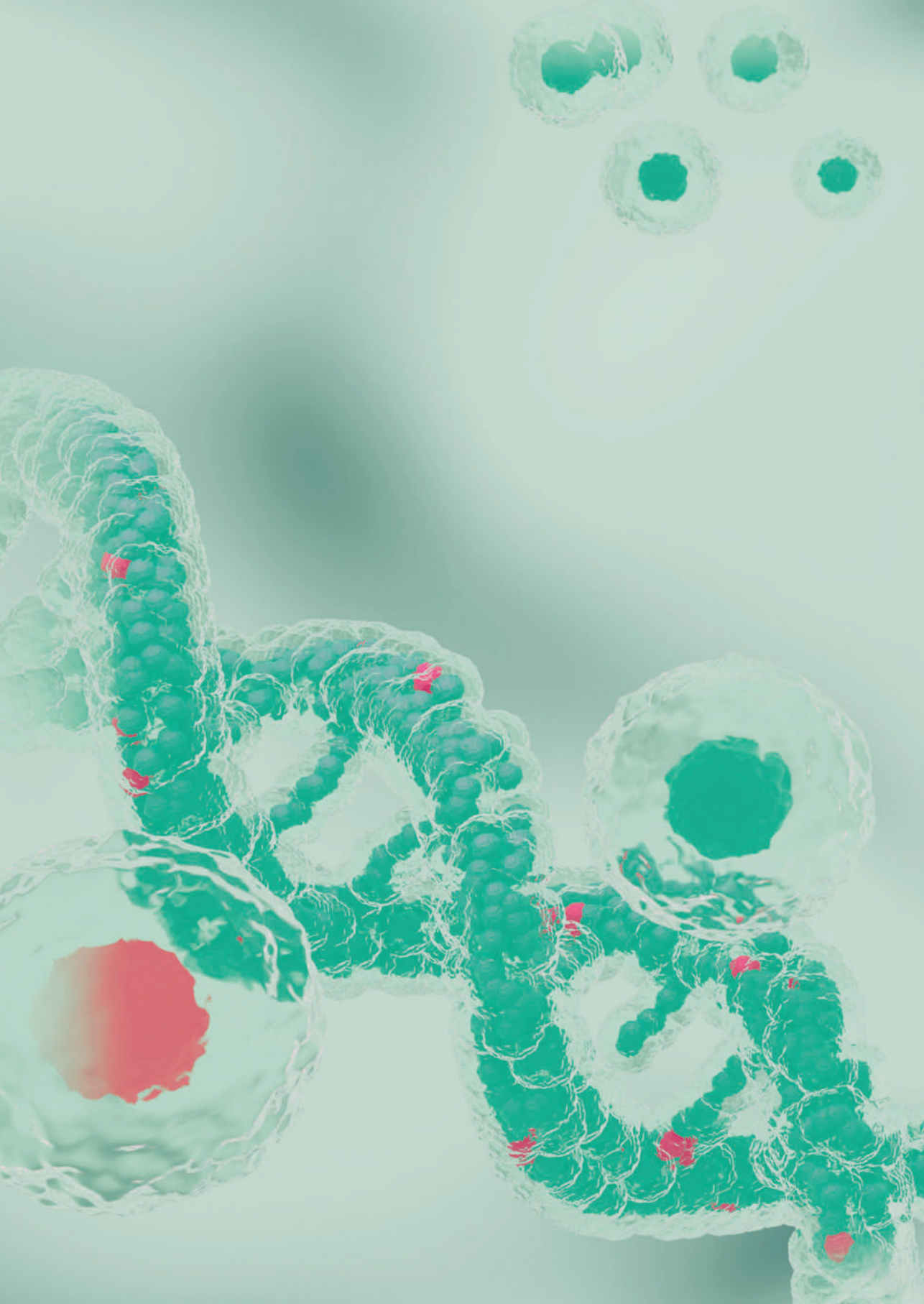
61. Khera, A.V., et al., *Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.* Nat Genet, 2018. **50**(9): p. 1219-1224.

62. Slunecka, J.L., et al., *Implementation and implications for polygenic risk scores in healthcare.* Hum Genomics, 2021. **15**(1): p. 46.

63. Boomsma, D.I. and B. Johnson, *Peer Review Report For: The Genetic Sphygmomanometer: an argument for routine genome-wide genotyping in the population and a new view on its use to inform clinical practice [version 1; peer review: 3 approved, 1 approved with reservations].* Wellcome Open Res, 2018. **3**(138).

64. Kumar, A.J., et al., *Graft-versus-host disease in recipients of male unrelated donor compared with parous female sibling donor transplants.* Blood Adv, 2018. **2**(9): p. 1022-1031.

65. van Dongen, J., et al., *Identical twins carry a persistent epigenetic signature of early genome programming.* Nat Commun, 2021. **12**(1): p. 5618.



7

GENERAL SUMMARY

GENERAL SUMMARY

In this thesis, I outline my work describing exposures to low concentrations of genetically foreign material via measures in biological samples. I begin by presenting a systematic review of the literature as it relates to investigation of naturally occurring chimerism including methods used to study the condition in health and disease of both humans and animals. Next, I conduct two studies to examine natural chimerism in human biology. The first study examines if chimerism as a biomarker of prior exposure to male blood in utero can provide insight into the etiology of Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome. The second study expands on prior research on microchimerism in dizygotic twins and aims at elucidating the prevalence and patterns of microchimerism acquisition in female family members of twin pedigrees. This is followed by a study examining how quantitative measures of a virus in samples collected for diagnostic testing can provide insight into the epidemiology of Coronavirus Disease 2019 (COVID-19) in the population.

The following provides succinct summaries of the principal chapters of this thesis.

Chapter 2 presents a systematic review of literature related to investigation of natural chimerism in both human and animal conditions. This chapter begins with examination of several of the most prominent methods applied in the detection of microchimerism, including blood antigen typing, karyotyping, polymerase chain reaction (PCR) and fluorescent in situ hybridization (FISH). Each method presents advantages and disadvantages due to the characteristics of the technique, although the availability and sensitivity of PCR has made it one of the most prominent choices for chimerism studies. Research of chimerism in the etiology of health and disease has included studies in both humans and animals. Developmental disorders have been implied to be influenced by chimerism exposure due to in utero transplacental exchange. In cattle and other mammals, freemartin syndrome is a developmental disorder that affects the female co-twin of opposite-sex dizygotic twins, and XX/XY chimerism (two cell populations with different sex chromosomes) has been observed as a result of the associated exposure to the blood of the male co-twin. Studies in humans have been primarily focused on elucidating the role of chimerism in common complex diseases such as autoimmune disease and cancer. Several studies of autoimmune diseases which have a female predilection have discovered an increased prevalence among cases, including rheumatic and thyroid autoimmune

diseases. In contrast, studies in cancer have indicated chimerism to be reduced in some types of cancer cases with chimerism associated with generally better prognoses. Such contradictory findings for chimerism in disease suggest a complex relationship between chimerism and disease pathology. Additional studies have implicated chimerism in female-specific diseases including pregnancy related complications. Further research into chimerism will necessitate identifying the origin of donor cells in chimerism, examining the prevalence of chimerism in a population and monitoring the longitudinal changes to chimerism burden to establish patterns in health and disease. Addressing these outstanding questions will provide insight into the understanding of chimerism risk in the population and intricacies of complex diseases.

Chapter 3 focuses on a study of male microchimerism in relation to MRKH syndrome. The MRKH syndrome phenotype has been known to share similarity to freemartinisim in other mammals. The principal factor contributing to the etiology of freemartinism in cattle has been determined to be exposure to hormones of a male co-twin via the placental vasculature, also inadvertently producing XX/XY chimerism in the female. This study investigates if a similar pathophysiological mechanism would explain the pathology observed in women with MRKH. To investigate this, a case-control study was conducted to examine the prevalence of male microchimerism as a persistent biomarker in women with MRKH and a control population of women with no history of pregnancy or cancer. A high sensitivity quantitative PCR assay was developed for relative quantification of male genome mass by Y chromosome gene *DYS14* compared to total genome mass through measure of *β-globin*. The samples tested revealed a greater prevalence of male microchimerism in the control group compared to MRKH cases (17.2% vs. 5.3%) as well as a larger mean concentration of male microchimerism in the control group (71.8 vs. 2.3 male genome equivalents per one million cells). While the results do not support the hypothesis of exposure to male blood in the pathogenesis of MRKH syndrome, these results derived from adult participants do not definitively exclude this explanation. In addition, the findings of this study present an interesting observation of a large proportion of women in the control group with detectable male microchimerism suggesting the need for further research into the source and mechanisms of persistent microchimerism.

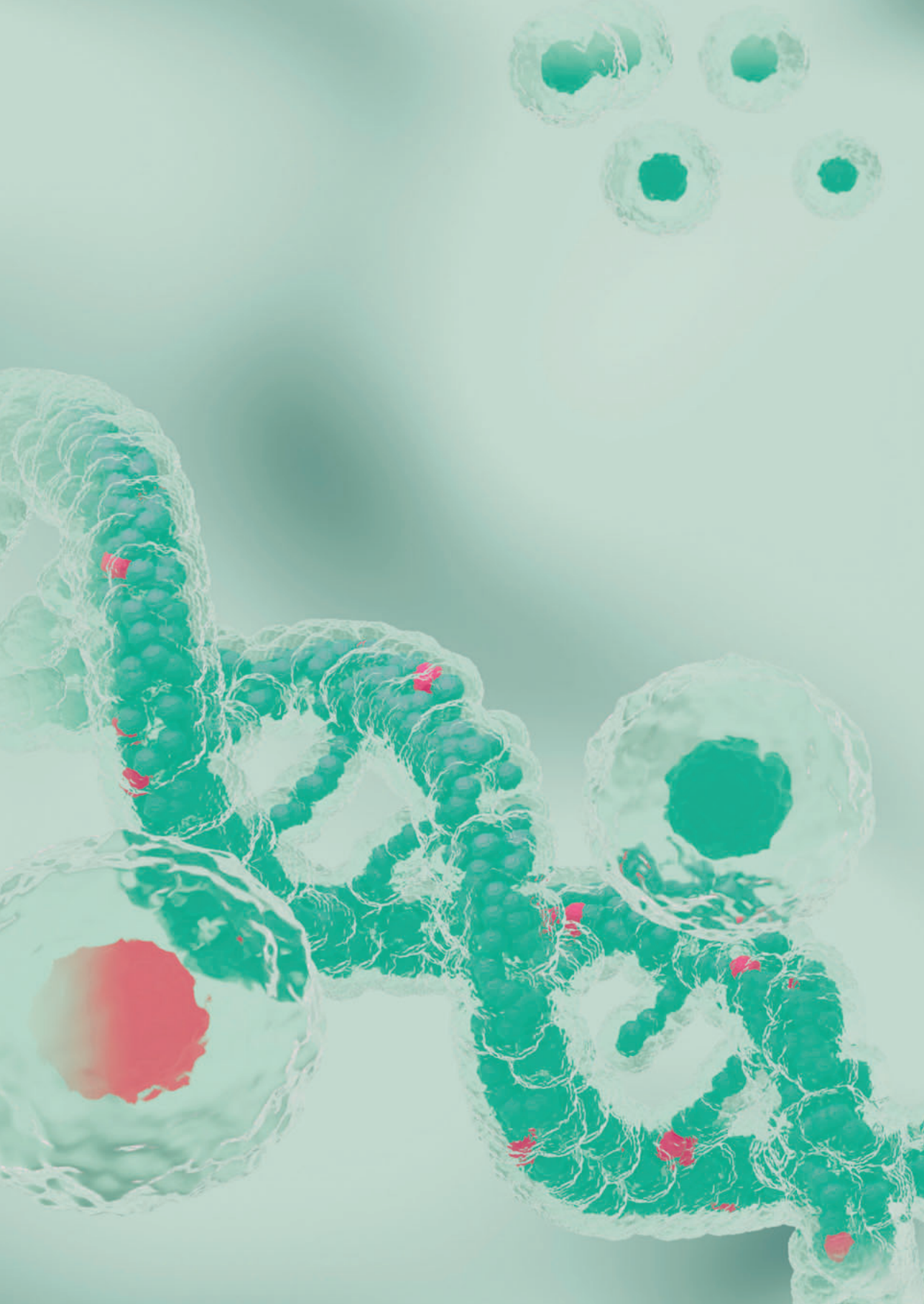
Chapter 4 details a study to investigate how male microchimerism permeates within two generations of twin pedigrees. As observed in the previous chapter,

women with no history of pregnancy still present with a notable prevalence of male microchimerism. This study examined if characteristics within a family increase the prevalence of male microchimerism, including having a male co-twin, older brothers, or sons. Participants from the Netherlands Twin Register (NTR) were selected if peripheral blood samples were collected for female twins, a non-twin sister and their mother for each pedigree. The extracted DNA from 446 females was tested using the male microchimerism assay for relative quantification of male genome described in Chapter 3. In total, 26.9% of women in the study presented with detectable male microchimerism with the greatest prevalence in the older generation, that is the mothers of twins (38.7%). This finding also coincided with a positive relationship between age and detectable male microchimerism. Additionally, monozygotic twins and same sex dizygotic twins had similar concordance in detectable male microchimerism. There was no significant difference in the concentration of chimerism for each type of family member relative to the twins. However, outlier samples with >80 male genome equivalents per one million cells were related to patients with a significantly lower age compared to others with detectable male microchimerism. There was no significant difference in male microchimerism prevalence related to individuals having male family members in the pedigree including male co-twins, sons or an older brother. These results contradict our primary hypothesis but emphasize additional underlying complexities in variance in the chimerism phenomenon. Overall, the findings of this study illustrate that microchimerism is a common phenomenon with substantial complexity in variation between individuals. Understanding these mechanisms may provide valuable insight into the contribution of microchimerism to other common complex diseases.

Chapter 5 evaluates the application of quantitative measures of minor genome concentration in examining the epidemiology of a respiratory virus during the COVID-19 pandemic. The microchimerism assay applied in the preceding chapters provided valuable insight for the development of a diagnostic assay for the detection of low concentration of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) viral genome against the background of human genetic material. The assay developed and implemented in clinical diagnostic testing at the Avera Institute for Human Genetics was a real-time reverse transcription PCR (RT-PCR) assay including detection of two viral genome targets in addition to a human control target. This diagnostic assay received Emergency Use Authorization from the United States

Food and Drug Administration early in the pandemic response and results from 74479 tests conducted between March and December 2020 were included in our study of epidemiological characteristics. Weekly mean Cycle threshold (Ct) values strongly correlated with weekly hospitalizations in South Dakota (SD) with a two-week delay and with deaths four weeks after testing. Further examination of patient symptom data for 101 patients showed symptoms could be reduced to three factors which described groups of symptoms including respiratory virus symptoms, nervous symptoms, and gastrointestinal symptoms. Patients with documented symptoms related to the three symptom groups also revealed variation in additional characteristics such as lower age and lower Ct values in patients with nervous system symptoms (ageusia and anosmia). With the emergence of SARS-CoV-2 variants of concern, I worked with the team at the Avera Institute for Human Genetics to develop and integrate a genetic surveillance program for detecting newly emerging SARS-CoV-2 mutations and variants. Investigation of diagnostic testing quantitative data as Ct values with SARS-CoV-2 variants presented no significant difference between non-N501Y (ancesteral SARS-CoV-2) and N501Y (presumably Alpha) but Ct values did differ significantly between Delta and Omicron. These findings provide real-world data to examine how population measures in a clinical laboratory may change as a result of changing viral dynamics within the community which may be able to inform public health in preparing for changing demands in health care services.

Chapter 6 provides a comprehensive summary of the preceding chapters, followed by a discussion of the central themes and overarching topics described throughout. I present perspectives on chimerism as a condition, the factors contributing to variance in chimerism presentation in the population, and a discussion of the ways in which mechanisms of host and donor immunotolerance may inform our understanding of disease pathophysiology. Following this, I examine how research applications of molecular testing methods are able to inform clinical applications for molecular diagnostics and provide insights for monitoring pathogen dynamics through quantitative measures across a testing population. Finally, I present my thoughts on the future landscape of genetics and associated quantitative measures for improving overall health. Here I highlight emerging applications of genetics in clinical care, opportunities presented by study of host-donor cellular interactions and discuss the need for both longitudinal and multi-omics studies to evaluate the broad implications of biological exposures.



APPENDIX

LIST OF PUBLICATIONS

1. Peters, H. E., **Johnson, B. N.**, Ehli, E. A., Micha, D., Verhoeven, M. O., Davies, G. E., Dekker, J., Overbeek, A., Berg, M., Dulmen-den Broeder, E. V., Leeuwen, F. E. V., Mijatovic, V., Boomsma, D. I., & Lambalk, C. B. (2019). Low prevalence of male microchimerism in women with Mayer-Rokitansky-Kuster-Hauser syndrome. *Hum Reprod*, 34(6), 1117-1125. https://doi.org/10.1093/humrep/dez044

2. **Johnson, B. N.**, Ehli, E. A., Davies, G. E., & Boomsma, D. I. (2020). Chimerism in health and potential implications on behavior: A systematic review. *Am J Med Genet A*, 182(6), 1513-1529. https://doi.org/10.1002/ajmg.a.61565

3. Slunecka, J. L., van der Zee, M. D., Beck, J. J., **Johnson, B. N.**, Finnicum, C. T., Pool, R., Hottenga, J. J., de Geus, E. J. C., & Ehli, E. A. (2021). Implementation and implications for polygenic risk scores in healthcare. *Hum Genomics*, 15(1), 46. https://doi.org/10.1186/s40246-021-00339-y

4. **Johnson, B. N.**, Peters, H. E., Lambalk, C. B., Dolan, C. V., Willemsen, G., Ligthart, L., Mijatovic, V., Hottenga, J. J., Ehli, E. A., & Boomsma, D. I. (2021). Male microchimerism in females: a quantitative study of twin pedigrees to investigate mechanisms. *Hum Reprod*, 36(9), 2529-2537. https://doi.org/10.1093/humrep/deab170

5. **Johnson, B. N.**, Cooper, L., Beck, J. J., Finnicum, C. T., Davis, C. M., Van Asselt, A. J., Kallsen, N., Silvernail, C., Viet, S., Long, K., Huizenga, P., Vaselaar, E., Pfeifle, M., Nelson, L., Streier, Z., Katz, C., Leonard, K. B., Matthiesen, N., Buschette, N., Weisser, L., Jahnke, M., Bohlen, K., Young, S. L., Sulaiman, R. A., Prouse, B. R., Basel, D. A., Black, M., Hansen, R., Petersen, J. L., Willemsen, G., Boomsma, D. I., Bleile, M. J., & Ehli, E. A. (2022). More than results: the clinical and research relationship in the evolving detection and surveillance of SARS-CoV-2. *S D Med*. Accepted.