UNIVERSIDADE FEDERAL DO PARANÁ

LEONARDO MALDANER AMORIM

ANÁLISE DE KIR E HLA EM ALTA RESOLUÇÃO EM UMA POPULAÇÃO EURO-DESCENDENTE E EM PACIENTES COM CARCINOMA MAMÁRIO ESPORÁDICO

CURITIBA
2022

LEONARDO MALDANER AMORIM

ANÁLISE DE KIR E HLA EM ALTA RESOLUÇÃO EM UMA POPULAÇÃO EURO-DESCENDENTE E EM PACIENTES COM CARCINOMA MAMÁRIO ESPORÁDICO

Tese apresentada ao Programa de Pós-Graduação em Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito à obtenção do título de Doutor em Genética.

Orientador: Prof. Dr. Danillo Gardenal Augusto
Coorientadora: Profª Drª. Marcia H. Beltrame

CURITIBA
2022

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS BIOLÓGICAS

Amorim, Leonardo Maldaner
Análise de KIR e HLA em alta resolução em uma população euro-descendente e em pacientes com carcinoma mamário esporádico / Leonardo Maldaner Amorim. – Curitiba, 2022.
1 recurso on-line : PDF.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências Biológicas, Programa de Pós-graduação em Genética.
Orientador: Prof. Dr. Danillo Gardenal Augusto.
Coorientadora: Profª Drª. Marcia H. Beltrame.

1. Receptores KIR. 2. Antígenos HLA. 3. Mamas - Câncer. 4. Genética da população humana. I. Augusto, Danillo Gardenal. II. Beltrame, Márcia Holsbach, 1984. III. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-graduação em Genética. IV. Título.

Bibliotecária: Giana Mara Seniski Silva CRB-9/1406

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA -
40001016006P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **LEONARDO MALDANER AMORIM** intitulada: **ANÁLISE DE KIR E HLA EM ALTA RESOLUÇÃO DE UMA POPULAÇÃO EURO-DESCENDENTE E EM PACIENTES COM CARCINOMA MAMÁRIO ESPORÁDICO**, sob orientação do Prof. Dr. DANILLO GARDENAL AUGUSTO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 19 de Setembro de 2022.

Assinatura Eletrônica
10/10/2022 16:22:49.0
DANILLO GARDENAL AUGUSTO
Presidente da Banca Examinadora

Assinatura Eletrônica
27/09/2022 21:17:08.0
MARCIA REGINA PINCERATI
Avaliador Externo (UNIVERSIDADE POSITIVO)

Assinatura Eletrônica
28/09/2022 08:11:48.0
JAQUELINE CARVALHO DE OLIVEIRA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
27/09/2022 15:55:53.0
LIANA ALVES DE OLIVEIRA
Avaliador Externo (FACULDADES INTEGRADAS DO BRASIL)

Assinatura Eletrônica
28/09/2022 11:59:17.0
MARCIA HOLSBACH BELTRAME
Coorientador(a) (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Primeiramente gostaria de agradecer aos meus pais, que por toda vida me ensinaram o valor que tem uma boa educação, e que se sacrificaram muito para que eu pudesse chegar até aqui. Sem eles eu não teria chegado aonde cheguei.

Gostaria de agradecer também à CAPES, que financiou minha bolsa de mestrado e doutorado. Apesar dos tempos sombrios que vivemos com relação ao investimento em pesquisa, foi graças ao investimento público que hoje entrego essa tese. É um privilégio estar chegando ao fim de um doutorado, e que muitos poucos têm acesso, e por esse motivo eu sou grato a todo investimento feito.

Queria fazer um agradecimento especial ao meu orientador Danillo Augusto, que muito me ensinou, e me ajudou a crescer como profissional. Uma das pessoas mais inteligentes que já conheci, ele foi e continua sendo um modelo de profissional para mim. Não somente um orientador, mas também um amigo, que me recebeu tão bem em sua casa nos Estados Unidos, e não mediu esforços para que nossa pesquisa saísse da melhor forma possível. Agradeço imensamente por todos os momentos dentro e fora da universidade, pela amizade e pela mentoria.

Gostaria também de agradecer aos mestres que me ensinaram tanto ao longo do caminho, em especial as professoras Maria Luiza Petzl-Erler e Ana Cláudia Bonatto, por terem me inspirado a seguir a carreira na genética. Também gostaria de agradecer a professora Márcia Beltrame, por ter aceitado ser minha coorientadora e ter me ensinado tanto ao permitir que eu a acompanhasse durante a prática de docência.

Não poderia esquecer também das amigas que esses 6 anos e meio de pós-graduação me trouxeram. Hellen Issler, Luciana Vargas, Luana Oliveira, Tamyres Carvalho e Verónica Calonga, durante essa jornada nós dividimos muitos momentos, dentro e fora da universidade. Sem a amizade, apoio e suporte de vocês eu não teria chegado aonde cheguei.

À minha namorada Barbara, que me ensinou tanto em tão pouco tempo, e que me ensinou tantas coisas, que apesar de não serem diretamente

relacionadas a minha pesquisa acadêmica, me fizeram uma pessoa muito melhor. Sem você eu definitivamente não estaria aqui tão bem quanto eu estou hoje.

Por fim, gostaria de agradecer à Dra. Jill Hollenbach por ter aberto as portas de seu laboratório para que eu pudesse realizar a minha pesquisa, e à Fulbright por ter acreditado no meu projeto e financiado meu sonho de morar nos Estados Unidos.

Fica aqui essa breve mensagem de agradecimento, que definitivamente não representa todo o carinho e gratidão que eu tenho pelas pessoas citadas, mas para deixar registrado todos que me ajudaram tanto nessa jornada. Deixo aqui meu muito obrigado!

“Nada é tão grande que não possa ser alcançado e nada é tão pequeno que não seja importante.”

Allan Kardec

RESUMO

As células assassinas naturais (ou NK, do inglês *natural killer*) são importantes efetores da imunidade inata e adaptativa, atuando no combate a células infectadas e células tumorais. Os genes *KIR* (*killer-cell immunoglobulin-like receptors*) codificam um grupo de receptores que modulam as respostas dessas células através do reconhecimento de antígenos leucocitários humanos (HLA, do inglês, *human leukocyte antigens*). Combinações de variantes *KIR-HLA* já foram associadas com diversas doenças, incluindo vários tipos de câncer, doenças autoimunes e infecciosas. No entanto, devido à complexidade estrutural da família de genes *KIR*, poucos estudos atingiram um nível de profundidade de análise que permitiu explorar a variação alélica e de número de cópias desses genes em larga escala. Com o objetivo de preencher essa lacuna, nosso primeiro objetivo foi descrever a variação de *KIR* em alta resolução em uma grande amostra populacional, aplicando sequenciamento de nova geração em 2130 Euro-descendentes dos Estados Unidos. A determinação precisa do número de cópias de *KIR* nos permitiu identificar um conjunto de haplótipos *KIR* incomuns, responsáveis por 5,2% da variação estrutural total. Observamos que *KIR2DL4* foi o gene moldura que mais variou em número de cópias (6,5% de todos os indivíduos). Além disso, identificamos 250 alelos com resolução de 5 dígitos, dos quais 90 têm frequências ≥1%. Encontramos padrões de sequência que eram consistentes com a presença de novos alelos em 398 (18,7%) indivíduos e contextualizamos vários polimorfismos de nucleotídeo únicos (SNP) órfãos dentro do complexo. Também identificamos uma nova variante de KIR2DL1 (Pro151Arg), a qual demonstramos por dinâmica molecular que esta substituição possivelmente afeta a interação com HLA-C. No contexto de KIR e doenças, nós escolhemos o câncer de mama como foco do nosso estuado devido à falta de estudos que analisam KIR e HLA de classe I e classe II nessa doença. O câncer de mama (BC) é a neoplasia mais comum entre as mulheres, com tumores que variam em agressividade, tipo histológico, subtipo molecular e desenvolvimento de metástases entre outros. Neste trabalho, analisamos o DNA de 550 mulheres com câncer de mama esporádico e 747 controles. Nós identificamos que a presença de três cópias de *KIR2DL5* estava associada à proteção (OR = 0,21, p^{corr} = 0,027), enquanto a presença de três cópias de *KIR3DL1S1* foi associada ao risco aumentado para BC (OR = 2,44, p^{corr} = 0,009). Também identificamos que o par KIR2DL3*001+HLA-C1 estava significativamente reduzido em pacientes com o subtipo HER2+ não luminal (ER e PR negativo) em comparação com todos os outros subtipos (OR = 0,23, p^{corr} = 0,016), além de uma forte associação de KIR2DS1*002+HLA-C2 em pacientes com tumores triplo negativos em comparação com todos os outros subtipos (OR = 3,34, $p^{corr} < 1,4 \times 10^{-4}$). Por fim, observamos um forte efeito protetor do par KIR2DL1*001+HLA-C2, significativamente aumentado em pacientes com BC de grau I em comparação com os graus mais agressivos II+III (OR = 0,25, p^{corr} = 0,038). Nós também exploramos a diversidade de HLA classe I e classe II em BC e seus subtipos. *HLA-C*12:03* estava significativamente mais frequente em pacientes com câncer de mama triplo negativo (OR = 2.7, p = 0.0025, p^{corr} = 0.033), bem como variações em amino ácidos nas posições 11 e 12 de HLA-B. *HLA-DPB1*17:01* foi associado à proteção ao desenvolvimento de metástases em linfonodos (OR = 0.21, p = 0.0019, p^{corr} = 0.037). E por fim, a presença do amino ácido arginina na posição 151 de HLA-A estava associada ao desenvolvimento de câncer de mama grau 3 (OR = 3.36, p = 0.0016, p^{corr} = 0.016). Já os amino

ácidos lisina e ácido aspártico nas posições 9 e 11 de HLA-DRB1 estavam associadas à proteção contra o câncer de mama de grau 3 (OR = 0.13, p = 0.0033, p^{corr} = 0.033). Nenhum estudo anterior explorou a variação estrutural e de sequência de *KIR* em nenhuma população ou doença com a profundidade que apresentamos neste trabalho. Demonstramos que a junção do sequenciamento de alto rendimento com ferramentas computacionais de última geração permite a exploração de todos os aspectos da variação de *KIR,* incluindo a determinação da diversidade de haplótipos em nível populacional, melhorando a compreensão dessa complexa família gênica e fornecendo uma referência importante para estudos futuros.

Palavras-chave: KIR; HLA; Câncer de mama esporádico; Estudo de associação; Estudo populacional.

ABSTRACT

Natural killer cells (or NK) are important effectors of innate and adaptive immunity, acting in the fight against infected and tumor cells. The *KIR* genes (killer-cell immunoglobulin-like receptors) encode a group of receptors that modulate the responses of these cells through the recognition of human leukocyte antigens (HLA). Combinations of KIR-HLA variants have been associated with several diseases, including many types of cancer, autoimmune and infectious diseases. However, due to the structural complexity of the *KIR* gene family, few studies have reached a level of analysis depth that allowed exploring the allelic and copy number variation of these genes on a large scale. In order to fill this gap, our first objective was to describe the *KIR* variation in high resolution in a large population sample, applying next-generation sequencing in 2130 Euro-descendants from the United States. Precise determination of the *KIR* copy number allowed us to identify a set of unusual *KIR* haplotypes, responsible for 5.2% of the total structural variation. We observed that *KIR2DL4* was the framework gene that most varied in number of copies (6.5% of all individuals). In addition, we identified 250 alleles with 5-digit resolution, of which 90 have frequencies ≥1%. We also found sequence patterns that were consistent with the presence of new alleles in 398 (18.7%) individuals and contextualized several orphan single nucleotide polymorphisms (SNP) within the complex. We also identified a new variant of KIR2DL1 (Pro151Arg), which we demonstrated by molecular dynamics that this substitution possibly affects the interaction with HLA-C. In the context of KIR and disease, we chose breast cancer as the focus of our study due to the lack of studies that analyze *KIR* and *HLA* class I and class II in this disease. Breast cancer (BC) is the most common malignancy among women, with tumors that vary in aggressiveness, histological type, molecular subtype, and development of metastases, among others. In this work, we analyzed the DNA of 550 women with sporadic breast cancer and 747 controls. We found that the presence of three copies of *KIR2DL5* was associated with protection (OR = 0.21, p^{corr} = 0.027), while the presence of three copies of *KIR3DL1S1* was associated with increased risk for BC (OR = 2.44, p^{corr} = 0.009). We also identified that the KIR2DL3*001+HLA-C1 pair was significantly reduced in patients with the non-luminal HER2+ subtype (ER and PR negative) compared to all other subtypes (OR = 0.23, p^{corr} = 0.016), in addition to a strong association of KIR2DS1*002+HLA-C2 in patients with triple negative tumors compared to all other subtypes (OR = 3.34, $p^{corr} < 1.4 \times 10^{-4}$). Finally, we observed a strong protective effect of the KIR2DL1*001+HLA-C2 pair, significantly increased in patients with grade I BC compared to the more aggressive grades II+III (OR = 0.25, p^{corr} = 0.038). We also explored the diversity of HLA class I and class II in BC and their subtypes. *HLA-C*12:03* was observed significantly more frequent in patients with triple negative breast cancer (OR = 2.7, p = 0.0025, p^{corr} = 0.033), as well as variations in amino acids at positions 11 and 12 of HLA-B. *HLA-DPB1*17:01* was associated with protection against the development of lymph node metastases (OR = 0.21, p = 0.0019, p^{corr} = 0.037). Finally, the presence of the amino acid arginine at position 151 of HLA-A was associated with the development of grade III breast cancer (OR = 3.36, p = 0.0016, p^{corr} = 0.016). The amino acids lysine and aspartic acid at positions 9 and 11 of HLA-DRB1 were associated with protection against grade III breast cancer (OR = 0.13, p = 0.0033, p^{corr} = 0.033). No previous study has explored the structural and

sequence variation of *KIR* in any population or disease to the depth that we have presented in this work. We demonstrate that the combination of high-throughput sequencing with state-of-the-art computational tools allows for the exploration of all aspects of KIR variation, including the determination of haplotype diversity at the population level, improving the understanding of this complex gene family, and providing an important reference for future studies.

Keywords: KIR; HLA; Sporadic breast cancer; Association study; Population study.

LISTA DE FIGURAS

FIGURA 1. Organização dos genes localizados no complexo NKC (*Natural Killer Complex*), no cromossomo 12, e do complexo LRC (*Leukocyte Receptor Complex*), no cromossomo 19. 27
FIGURA 2. Fenômeno de ausência do próprio ("*Missing-self*") e fenômeno de reconhecimento induzido ("*Induced-self*"). 28
FIGURA 3. O Complexo Receptor de Leucocitário (LRC) estendido (19q13.4). 30
FIGURA 4. Organização gênica dos diferentes tipos de KIR. 31
Figura 5. Estrutura das proteínas KIR 33
FIGURA 6. Nomenclatura dos genes KIR 35
FIGURA 7. Conteúdo gênico do haplogrupo A e uma representação de genes que podem estar presentes no haplogrupo B. 37
FIGURA 8. Conteúdo gênico dos haplótipos centroméricos e teloméricos de KIR em uma população euro descendente dos Estados Unidos. 40
FIGURA 9. A região do MHC. 46
FIGURA 10. Estrutura das moléculas HLA de classe I e classe II 48
FIGURA 11. Receptores KIR e seus respectivos ligantes HLA conhecidos. 51
FIGURA 12. Distribuição de ligantes de KIR em diferentes populações. 55

LISTA DE TABELAS

TABELA 1. EXEMPLOS DE DOENÇAS ASSOCIADAS A COMBINAÇÕES DE KIR-HLA. 53
TABELA 2. ASSOCIAÇÕES DE *KIR* COM DIVERSOS TIPOS DE CÂNCER 57
TABELA 3. ASSOCIAÇÕES DE *KIR* E SEUS LIGANTES COM CÂNCER DE MAMA. 58

LISTA DE QUADROS

QUADRO 1. PARÂMETROS PARA CLASSIFICAÇÃO DO GRAU HISTOLÓGICO DO CÂNCER 20
QUADRO 2. SUBTIPOS IMUNO-HISTOQUÍMICOS 22
QUADRO 3. FREQUÊNCIAS DE *KIR* EM DIFERENTES POPULAÇÕES 39
QUADRO 4. CARACTERIZAÇÃO DAS AMOSTRAS DE CÂNCER DE MAMA 63

LISTA DE SIGLAS

CTL - Linfócito T citotóxico

DAP-12 -

ER - Receptores de estrogênio
HER2 - *human epidermal growth factor receptor-type 2*
HLA - Antígenos leucocitários humanos

IC – Intervalo de confiança
IL - Interleucina

INF - *Interferon*

IPD *- Immuno Polymorphism Database*

ITAM - Motivo ativador baseado em tirosina

ITIM - Motivo inibidor baseado em tirosina

KIR - Receptores das células NK semelhantes a imunoglobulinas

LRC - Complexo dos receptores leucocitários

MHC - Complexo principal de histocompatibilidade

MicA - *MHC class I polypeptide-related sequence A*
MicB - *MHC class I polypeptide-related sequence B*
NGS – Sequenciamento de nova geração
NK - *Natural killer*

NKC - Complexo *Natural Killer*

OR - *Odds ratio*

PB – pares de base

PCR - Reação em cadeia da polimerase

PCR-SSP - Reação em cadeia da polimerase com iniciadores de sequência específica

PCR-SSOP – Reação em cadeia da polimerase com sondas oligonucleotídicas

PING – *Pushing Immunogenetics to Next Generation*

PR - Receptores de progesterona

TNBC – *Triple negative breast cancer*

USD – Dólares americanos

SUMÁRIO

1.Introdução **16**

2. Revisão Bibliográfica **17**

2.1 CÂNCER DE MAMA 17

2.2 CLASSIFICAÇÕES DOS TUMORES MAMÁRIOS 18

2.3 CÉLULAS ASSASSINAS NATURAIS (*NATURAL KILLER*) 24

2.4 RECEPTORES DE CÉLULAS NK 25

2.5 RECEPTORES DE CELULAS NK SEMELHANTES À IMUNOGLOBULINA 29

2.6 A NOMENCLATURA DE *KIR* 34

2.7 O POLIMORFISMO DE *KIR* 36

2.8 A COMPLEXIDADE DOS GENES *KIR* E OS DESAFIOS PARA GENOTIPAGEM 41

2.9 O reconhecimendo dos ANTÍGENOS LEUCOCITÁRIOS HUMANOS (HLA) por KIR 45

2.10 INTERAÇÕES KIR-HLA E SEU IMPACTO FUNCIONAL E EVOLUTIVO 49

2.11 INTERAÇÕES ENTRE *KIR* E *HLA* E O DESENVOLVIMENTO DE TUMORES 56

3. JUSTIFICATIVA **60**

4. OBJETIVOS **61**

4.1 OBJETIVOS GERAIS 61

4.2 OBJETIVOS ESPECÍFICOS 61

5. MATERIAL E MÉTODOS **62**

5.1 AMOSTRAS BIOLÓGICAS 62

5.1.1 Amostras de euro descendentes dos Estados Unidos 62

5.1.2 Amostras de pacientes e controles de câncer de mama 62

5.3 EXTRAÇÃO DE DNA DAS AMOSTRAS DE PACIENTES DE CÂNCER DE MAMA 64

5.4 SEQUENCIAMENTO DE *KIR* E *HLA* 64

5.5 ANÁLISE DE DADOS DA POPULACAO EURO DESCENDENTE 65

5.5.1 Identificação de haplótipos de conteúdo gênico de *KIR* 65

5.5.2 Análise de desequilíbrio de ligação 65

5.5.3 Identificação de novos alelos 66

5.6 ANÁLISE DE DADOS DAS AMOSTRAS DE CANCER DE MAMA E CONTROLES.. 66
6. RESULTADOS.. 67
6.1 CAPÍTULO I: High-Resolution Characterization of *KIR* Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity 68
6.2 CAPÍTULO II: Natural Killer Cell Receptor Variation is Associated with More Aggressive Subtypes of Breast Cancer. 81
6.3 CAPÍTULO III: *HLA* Alleles Are Associated with Aggressive Tumors in Brazilian Women Diagnosed with Sporadic Breast Cancer......................... 108
7.Conclusões 126
Referências 128

1.INTRODUÇÃO

O tumor de mama é o tipo mais comum de neoplasia entre as mulheres. Seu desenvolvimento está associado a diversos fatores, como estilo de vida, exposição à radiação e predisposição genética. A classificação dos tumores mamários é importante para o prognóstico e direcionamento do tratamento, sendo eles classificados em relação ao grau de agressividade, tipo histológico, subtipo imuno-histoquimico, presença de metástase em linfonodos, entre outros (1).

No contexto imunológico, as células *natural killer* (NK) são linfócitos que atuam tanto na resposta imune inata quanto adaptativa e são importantes no reconhecimento e no combate a tumores e células neoplásicas. Elas são também responsáveis por combater infecções e têm importante papel na reprodução. A capacidade dessas células de reconhecer as células alvo, possivelmente cancerosa ou infectada, se dá através de diversos receptores de superfície celular. Os receptores semelhantes a imunoglobulinas (ou KIR, do inglês, *killer cell immunologlobulin-like receptor*) reconhecem moléculas HLA (do inglês, *human leukocyte antigen*) de classe I. Essas moléculas HLA, por sua vez, são expressas na superfície de todas as células nucleadas saudáveis. Em situações de anormalidade, como infecção ou neoplasia, há uma expressão anormal de moléculas HLA de classe I na superfície celular, o que serve de sinal para o reconhecimento e extermínio pelas células NK (2).

Os genes *KIR* apresentam um polimorfismo incomum de presença ou ausência de genes, possibilitando a existência de diferentes haplótipos de conteúdo gênico. Cada haplótipo, portanto, apresenta uma composição diferente de genes (3). Além da variação de presença ou ausência, os genes *KIR* também apresentam grande variabilidade alélica em cada gene (4). Essa complexa e extensa variação torna praticamente impossível dois indivíduos não aparentados apresentarem as mesmas variantes *KIR*.

Diversas doenças já foram associadas com a variação de genes *KIR* e *HLA*, incluindo doenças autoimunes, infecciosas e câncer (5,6). Apesar da importância das células NK e dos receptores KIR no combate a tumores, os estudos que avaliam o polimorfismo de *KIR* na suscetibilidade ao câncer de mama são escassos e conflitantes, sendo todos compostos por análise em baixa

resolução e a maioria limitada por um pequeno tamanho amostral. (7–11). Já no aspecto da diversidade normal, poucos trabalhos exploraram a diversidade alélica de *KIR* em alta resolução, devido à dificuldade técnica imposta por essa complexa família de genes. Portanto o objetivo deste trabalho foi preencher essas lacunas, analisando o polimorfismo de *KIR* tanto em uma população euro-descendente dos Estados Unidos quanto em pacientes com carcinoma mamário esporádico.

2. REVISÃO BIBLIOGRÁFICA

2.1 CÂNCER DE MAMA

O câncer de mama é uma doença heterogênea em diferentes aspectos biológicos, incluindo diferentes prognósticos e respostas a tratamento (12). O tumor mamário é o tipo mais comum de malignidade (13–16) e, entre as mulheres, é a principal causa de morte por câncer no mundo (15–18). Genes como o *BRCA1, BRCA2, CHEK2, PTEN e TP53*, entre outros, foram identificados como causadores de câncer mamário hereditário em famílias nas quais o câncer de mama apresentava alta frequência (1,19,20).

Diferentemente do câncer de mama hereditário, em que normalmente uma variante genética tem papel determinante para causar a doença, o câncer de mama esporádico não apresenta um padrão de herança clássico, tratando-se de uma doença complexa em que múltiplos fatores genéticos e ambientais contribuem para a sua etiologia (19). Os fatores de risco para o desenvolvimento dessa doença incluem o sexo (mulheres tem maior probabilidade de desenvolver um tumor mamário do que homens), obesidade pós menopausa, tratamento hormonal durante a menopausa, radiação ionizante, idade, histórico familiar, entre outros fatores (21–27).

O desenvolvimento do câncer de mama está relacionado a uma variedade de alterações genéticas, que irão fazer com que células normais adquiram um conjunto de capacidades funcionais que as levam para estados de crescimento neoplásicos. Dentre essas alterações estão as alterações de ganho de função em proto-oncogenes e de perda de função em genes supressores de tumor (28). O ganho de função de proto-oncogenes resulta na ativação de

oncogenes. Essa ativação dos oncogenes pode se dar através de expressão aumentada do gene devido a amplificações gênicas, translocações cromossômicas ou mutações de ponto. Por exemplo, em alguns tipos de câncer de mama ocorre o aumento de cópias nas células tumorais do gene *human epidermal growth factor receptor-type 2* (*ERBB2* ou *HER2)*, o que vem a causar um aumento da expressão desse receptor nas células tumorais (29–33). Outros oncogenes que aparecem desregulados no câncer de mama são *PI3KCA* (34), *MYC* (35,36) e *CCND1* (37).

Genes supressores de tumor atuam como reguladores negativos do crescimento celular, o que inibe a progressão do ciclo celular ou ativa o sistema de reparo quando há danos no DNA. Defeitos nesses genes podem desencadear um potencial efeito invasivo ou metastático na célula (38). Vários genes supressores de tumor já foram relacionados ao câncer de mama, como por exemplo *BRCA1* (39–41), *BRCA2* (41,42), *PTEN* (43) e o *TP53* (44) .

2.2 CLASSIFICAÇÕES DOS TUMORES MAMÁRIOS

Devido ao fato do câncer de mama ser uma doença complexa e heterogênea, a estratificação dos tumores é essencial para o prognóstico da doença e definição do melhor tipo de tratamento (45,46). Aqui, abordaremos apenas dois tipos de classificação, que serão elucidados ao longo deste trabalho. A primeira, leva em consideração o potencial de agressividade do tumor. Um dos sistemas de classificação por potencial de agressividade, conhecido *como Nottingham Histologic Score System* (modificação do método *Scarff-Bloom-Richardson grading system*, por Elston e Elvis em 1991) (47,48), divide os tumores em graus, I, II e III, sendo o grau I o de menor agressividade e o grau III o de pior prognóstico (48–50). Para essa classificação, é levada em consideração a similaridade microscópica das células tumorais em relação às células normais do tecido mamário. Os parâmetros para classificação são: a quantidade de formação de estruturas tubulares (proporção de células tumorais que apresentam estruturas tubulares), as características dos núcleos (se eles são uniformes como em células epiteliais normais, ou tem formato irregular) e por último sua atividade mitótica (identificação de figuras mitóticas) (48). Cada um desses fatores recebe um valor de 1-3, e ao final da classificação os valores

são somados e usados para determinar o grau do tumor. O grau I tem uma contagem de 3-5 pontos, enquanto grau II são de 6-7 pontos, e grau III apresenta de 8-9 pontos (Quadro 1) (48,49). Tumores de grau I se multiplicam com menor velocidade e tem um prognóstico melhor do que tumores de graus maiores (50). Quanto maior o grau de agressividade, maior sua taxa de multiplicação e, consequentemente, pior o seu prognóstico (12).

QUADRO 1. PARÂMETROS PARA CLASSIFICAÇÃO DO GRAU HISTOLÓGICO DO CÂNCER.

Pontuação	Estruturas tubulares	Características do núcleo	Contagem mitótica
1	>75% da área do tumor formando estruturas tubulares	Pequena variação de tamanho	Menor ou igual a 7 por 10 campos de visão
2	10% a 75% da área do tumor formando estruturas tubulares	Variação de tamanho moderada	8-14 mitoses por 10 campos de visão
3	<10% da área do tumor formando estruturas tubulares	Grande variação de forma e tamanho	15 ou mais mitoses por 10 campos de visão

A soma das pontuações determina o grau histológico do câncer: grau **I** em tumores com contagem de 3-5 pontos; grau **II** com contagem de 6-7 pontos; e grau **III** com contagem de 8-9 pontos. Fonte: (ARGANI;CIMINO-MATHEWS, 2015) (49).

No passado, a escolha do tratamento para o tumor era essencialmente baseada no grau do câncer e em seu estágio. Porém, nas últimas décadas muito se tem estudado para melhorar a classificação do câncer de mama, utilizando uma classificação molecular para tentar levar em consideração outros aspectos da complexidade e heterogeneidade da doença. Perou e colaboradores (2000) classificaram os tumores em quatro subgrupos baseados na expressão de mais de 8000 genes (51). Os subgrupos são: luminal, "*normal like*", basal e *HER2* enriquecido. Posteriormente, o subgrupo Luminal foi subdividido em A e B (52).

Devido às dificuldades metodológicas e o alto custo, para que esta classificação se tornasse útil na prática clínica, criou-se uma classificação imuno-histoquímica. Essa classificação reflete de maneira aproximada, e com o uso de menos marcadores, os subgrupos mencionados anteriormente, com a inclusão do subgrupo triplo negativo (TNBC) (29,53,54). Nesta classificação é analisada a presença ou ausência dos receptores de estrogênio e progesterona na célula tumoral, assim como a amplificação do gene *HER2* e o índice de marcação Ki-67, um marcador de proliferação celular (29,31,55). A amplificação de *HER2,* que é um proto-oncogene responsável pelo crescimento da célula, acontece quando há um funcionamento anormal, onde há um excesso de cópias do próprio gene (processo conhecido como amplificação gênica), o que causa uma expressão muito aumentada de receptores HER2 na célula (nesse caso, diz-se que o tumor é enriquecido para HER2) (30). Uma explicação mais detalhada de cada subtipo será apresentada a seguir, e está sumarizada no quadro 2.

De acordo com o encontro de St. Gallen, decidiu-se por adotar cinco subtipos distintos para classificação imuno-histoquímica (31). O subtipo Luminal A representa 50% dos tumores, sendo assim o subtipo imunoistoquímico mais comum, e é caracterizado por ser positivo para os receptores de estrogênio (ER) e progesterona (PR), HER2 negativo, e ter uma expressão de KI67 abaixo de 14% (31,32). O ponto de corte entre os valores 'alto' e 'baixo' para Ki-67 varia entre laboratórios. Um nível menor que 14% é melhor correlacionado com a definição de expressão gênica de Luminal A, com base nos resultados de um único laboratório de referência (29). Os pacientes do subtipo luminal A tem um melhor prognóstico e são tipicamente de grau I (56). Normalmente utiliza-se a terapia hormonal como tratamento, sendo a quimioterapia restrita a poucos casos (32,57).

QUADRO 2. SUBTIPOS IMUNO-HISTOQUÍMICOS DO CÂNCER DE MAMA.

SUBTIPO	ER	PR	*HER2*	KI67
Luminal A	+	+	-	<14%
Luminal B/ *HER2* negativo	+	+/-	-	>14%
Luminal B/ *HER2* positivo	+	+/-	+	>14%
***HER2* positivo não luminal**	-	-	+	>14%
Triplo negativo	-	-	-	>14%

Legenda: ER = Presença de receptores de estrogênio; PR = Presença de receptores de progesterona. Fonte: o autor, 2022.

O subtipo Luminal B representa 20% dos tumores. Ele pode ser do tipo Luminal B HER2 negativo, que é positivo para ER e pode ser positivo também para PR, negativo para amplificação de HER2, apresentando uma expressão alta de KI67 (14% ou mais) (31,32). Ele também pode ser do tipo Luminal B HER2 positivo, sendo ER positivo e tendo a amplificação de HER2. Geralmente possui um grau de agressividade maior que Luminal A, e um prognóstico piorado em relação ao anterior. Como tratamento utiliza-se a terapia hormonal e a quimioterapia (29,31,32,57,58). No caso de existir a amplificação de HER2, ainda pode ser usado como tratamento a terapia de anticorpos monoclonais (31,32).

O próximo subtipo é o HER2 enriquecido não luminal, que representa de 15 a 20% dos tumores mamários, e em qual tanto ER quanto PR são negativos e HER2 é positivo (31,32). A expressão de Ki67 também é alta, sendo esses tumores geralmente de grau III e apresentando um prognóstico desfavorável. Esse subtipo está associado com uma doença mais agressiva, alta taxa de recorrência e uma menor taxa de sobrevivência (33). O tratamento é feito utilizando anticorpos monoclonais direcionados ao receptor HER2, além da quimioterapia (59–63).

O subtipo triplo negativo, apresenta ER, PR e HER2 negativos (64), com um nível de expressão de Ki67 elevado (31,32). Esse subtipo é o mais agressivo de todos e apresenta o pior prognóstico. Por não apresentar nenhum receptor ER ou PR, o tratamento por hormônios não é efetivo, e a ausência de super expressão de HER2 contraindica o uso de anticorpo monoclonal específico (65). Esse subtipo costuma ser um carcinoma ductal invasivo com alto índice mitótico. O tratamento indicado é a quimioterapia, porém a reincidência desse subtipo costuma ser agressiva, com metástases atingindo diferentes órgãos (32).

O câncer pode ainda invadir outros tecidos do corpo, o que caracteriza a metástase. O tumor pode se propagar através do próprio tecido para células próximas, ou pode se espalhar através dos sistemas linfático, através dos vasos linfáticos ou sanguíneos (66). É importante ressaltar que o tumor metastático é do mesmo tipo que o tumor original. Por exemplo, se o tumor mamário se espalhar para o ovário, este tumor ainda será composto por células mamárias e não células do órgão onde este se inseriu.

2.3 CÉLULAS ASSASSINAS NATURAIS (*NATURAL KILLER*)

As células *natural killer* (NK) são linfócitos citotóxicos que estão espalhadas por tecidos linfoides e não linfoides, e representam de 2% a 18% dos linfócitos presentes no sangue periférico humano (67). Elas têm papel crítico em respostas imunológicas, especialmente contra vírus e bactérias intracelulares (68). As células NK são tradicionalmente consideradas partes da imunidade inata, mas também apresentam funções na imunidade adaptativa (69–71). Elas são efetoras da resposta inicial contra infecções e do controle antitumoral. Ainda, as células NK apresentam função relevante na reprodução, estando envolvidas no processo de implantação do embrião e estabelecimento da placenta durante a fase formativa da reprodução (72,73).

As células NK foram descobertas na década de 70, em um estudo no qual mediu-se a atividade citotóxica específica de linfócitos de camundongos contra tumores (74). Naquele estudo, foi demonstrada uma resposta não especifica para matar células tumorais desempenhada por linfócitos grandes e granulares, que mais tarde foram chamados de células *natural killer*. Na mesma década, linfócitos humanos com capacidade semelhante de lisar células tumorais espontaneamente *in vitro* foram identificados (75,76). Atualmente, sabe-se que as células NK atuam na vigilância imunológica não apenas por meio de funções citotóxicas efetoras, mas também como linfócitos reguladores capazes de secretar citocinas e interagir com células imunes inatas e adaptativas, como monócitos/macrófagos, células dendríticas e linfócitos T (77–80).

As células NK têm a capacidade de reconhecer células infectadas sem a necessidade de anticorpos e apresentação antigênica pelo complexo principal de histocompatibilidade (MHC), garantindo respostas mais rápidas (81). As NK sanguíneas se diferem dos linfócitos T citotóxicos (CTLs) por expressarem a glicoproteína CD56, e ter a ausência do complexo proteico CD3. A molécula CD56 é uma molécula glicoproteica heterodimérica da superfície celular que é uma isoforma da molécula de adesão das células neurais (NCAM) (82), enquanto CD3 é um complexo proteico e correceptor de células T que está envolvido na ativação de células T citotóxicas (células T virgens CD8+) e células T auxiliares

(células T virgens CD4+) (83). Por outro lado, ambos CD56 e CD3 estão presentes na membrana celular de CTLs (68).

As células NK podem ser divididas em dois tipos: $CD56^{dim}$ (baixa expressão de CD56 na superfície) e $CD56^{bright}$ (alta expressão de CD56 na superfície) (81). Células $CD56^{dim}$ são CD16+, expressam perforina e produzem intérferon gama (IFN-γ). CD16 é um receptor expresso em células NK que facilita a citotoxicidade celular dependente de anticorpos (do inglês *antibody-dependent cellular cytotoxicity* ou ADCC) pela ligação à porção Fc de vários anticorpos (84). A perforina liberada será responsável pela criação de poros na membrana da célula alvo, levando essa célula à apoptose (85). Já células $CD56^{bright}$ são $CD16^{-}$ e não possuem perforina, produzindo IFN-γ a partir do contato com interleucinas (81). Naturalmente, células $CD56^{dim}$ são mais citotóxicas que células $CD56^{bright}$ (81). A atividade citotóxica das NK depende do ambiente de citocinas presente, assim como a interação com outras células do sistema imune (67). Interferons do tipo I, e as interleucinas 12, 15 e 18 são ativadores das células NK (85).

2.4 RECEPTORES DE CÉLULAS NK

As células NK apresentam uma grande variedade de receptores em sua superfície, que podem aumentar ou diminuir sua resposta à célula alvo (86). Através dos sinais gerados por receptores inibidores e ativadores, essas células são capazes de diferenciar células infectadas de células saudáveis (87).

Os genes que codificam esses receptores estão localizados principalmente em duas regiões genômicas. No cromossomo 12 está localizado o *natural killer complex* (NKC), que contém, dentre outros, os genes *NKG2* e *CD94* da família de receptores de lectina do tipo C (88). Por outro lado, no cromossomo 19 está localizado o l*eukocyte receptor complex* (LRC), onde estão a família de receptores semelhantes a imunoglobulina (Figura 1). Juntos, esses dois complexos contêm cerca de 90 genes, dentre os quais muitos tem variação estrutural de presença e ausência e homologia entre si. O que determina a atividade da NK é o balanço entre os sinais gerados pelos receptores inibidores e ativadores dessas duas superfamílias (67,89,90).

A maioria dos receptores ativadores reconhecem ligantes solúveis, como citocinas, ou interagem com moléculas de superfície celular, como no caso do

receptor NKG2D, que reconhece moléculas MHC (do inglês *Major Histocompatibility Complex*) de classe I como MICA e MICB (81). Os receptores inibidores (e alguns receptores ativadores) reconhecem moléculas de classe I, que estão sendo expressas na superfície celular de quase todas as células do organismo (91).

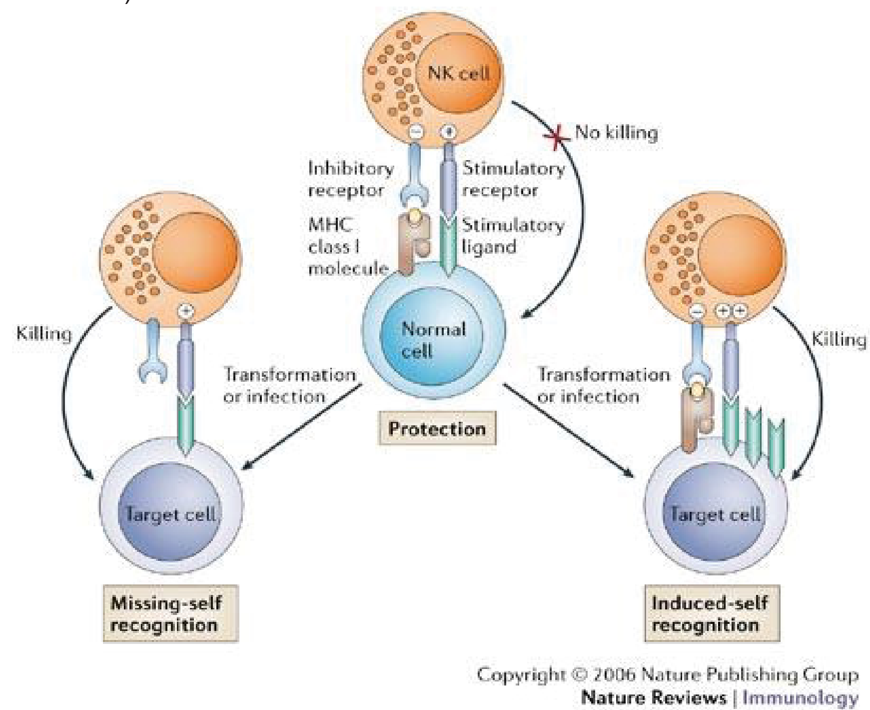
Todas as células nucleadas expressam moléculas HLA de classe I naturalmente em sua superfície celular. Algumas infecções virais, porém, reduzem a expressão de HLA de classe I, como uma adaptação para escapar das respostas imunes via apresentação de antígeno. Portanto, células infectadas e células neoplásicas podem apresentar uma expressão anormal dessas moléculas, o que é reconhecido pelas NK como um sinal para o ataque. Assim quando a expressão de HLA de classe I está diminuída devido a uma infecção ou neoplasia, os receptores inibidores deixam de gerar sinais inibidores que impediriam o ataque das células NK e, portanto, as células NK são ativadas, liberando citocinas e lisando a célula sob estresse (67). Esse fenômeno também é conhecido com ausência de reconhecimento do próprio, do inglês *missing-self* (Figura 2) (91).

Além da ausência de moléculas HLA de classe I e, por conseguinte a ausência de sinais inibidores, para que haja a atividade da NK, ainda se faz necessária a atuação dos receptores ativadores, gerando o balanço entre sinais inibidores/ativadores comentado anteriormente. Caso a presença de sinais ativadores seja mais forte que a de sinais inibidores, haverá o ataque pela célula NK (fenômeno conhecido como *induced-self recognition*) (Figura 2) (87).

FIGURA 1. Organização dos genes localizados no complexo NKC (*Natural Killer Complex*), no cromossomo 12, e do complexo LRC (*Leukocyte Receptor Complex*), no cromossomo 19.

Os complexos NKC e LRC possuem os genes que codificam os receptores das células NK. A região do LRC possui um tamanho de aproximadamente1MB, e contém os genes da família *KIR* (do inglês, *Killer Cell Immunoglobulin-like Receptors*) que serão abordados mais a fundo neste trabalho. Fonte: (Adptado de Kelley et al, 2005) (92).

FIGURA 2. Fenômeno de ausência do próprio (“*Missing-self*”) e fenômeno de reconhecimento induzido (“*Induced-self*”).



A imagem central mostra uma célula saudável, que expressa MHC de classe I, que é reconhecido pelo receptor inibidor, e impede o ataque da célula NK. A imagem à esquerda mostra uma célula que não expressa MHC de classe I, o que faz com que os receptores inibidores de NK não tenham um ligante, e, portanto, parem de inibir a ação da NK, o que ativa o seu ataque (fenômeno do *missing-self*). Por fim a imagem da direita mostra uma célula que expressa MHC de classe I, porém apresenta um conjunto de sinais ativadores mais fortes que o sinal inibitório, ativando o ataque da NK (fenômeno do *induced-self*). Fonte: (Raulet et al, 2006 (93).

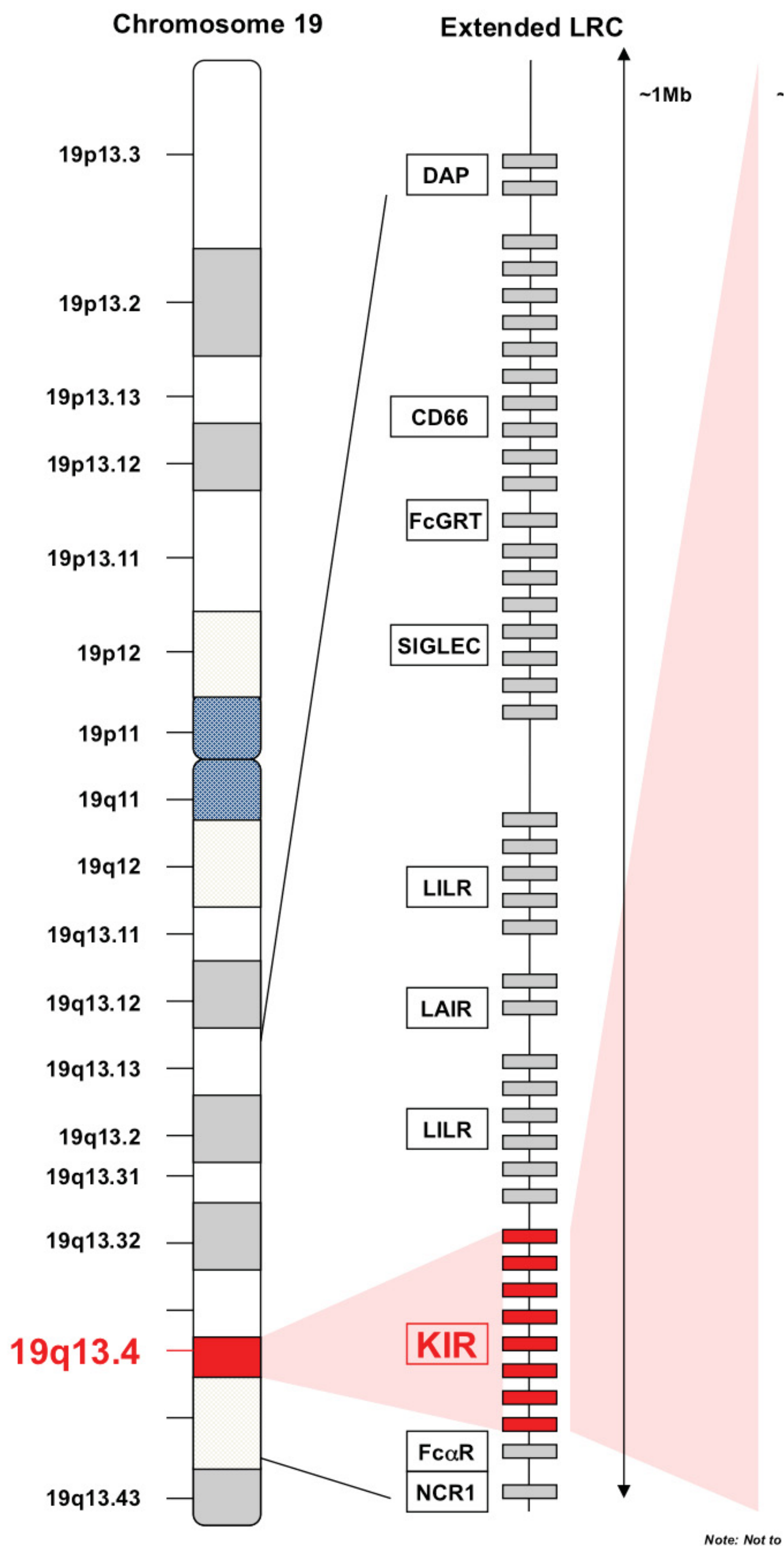
2.5 RECEPTORES DE CELULAS NK SEMELHANTES À IMUNOGLOBULINA

Dentre as famílias de receptores ativadores e inibidores das células NK, estão os receptores semelhantes a imunoglobulinas (KIR). Os receptores KIR exibem dois ou três domínios extracelulares, além de uma cauda citoplasmática, responsável por transduzir sinais ativadores ou inibidores (94). Existem 13 genes *KIR* conhecidos (*KIR3DL3, KIR2DS2, KIR2DS5, KIR2DL23, KIR2DL5A, KIR2DL5B, KIR2DS3, KIR2DL1, KIR2DL4, KIR3DL1S1, KIR2DS1, KIR2DS4* e *KIR3DL2*), além de dois pseudogenes (*KIR2DP1* e *KIR3DP1*), codificados dentro de uma região de 100-200 Kb do complexo receptor leucocitário (LRC) localizado no cromossomo 19 (19q13.4) (86) (Figura 3).

Os genes *KIR* variam em comprimento de 4 a 16 Kb (sequência genômica completa) e podem conter de quatro a nove éxons. Os genes *KIR* são classificados como pertencentes a um dos três grupos de acordo com suas características estruturais: 1) genes *KIR2D* tipo I, que codificam duas proteínas de domínio extracelular com conformação D1 e D2; 2) Os genes *KIR2D* Tipo II, que codificam duas proteínas de domínio extracelular com uma conformação D0 e D2 e finalmente; 3) genes *KIR3D* que codificam proteínas com três domínios tipo imunoglobulina extracelulares (D0, D1 e D2) (94) (Figura 4).

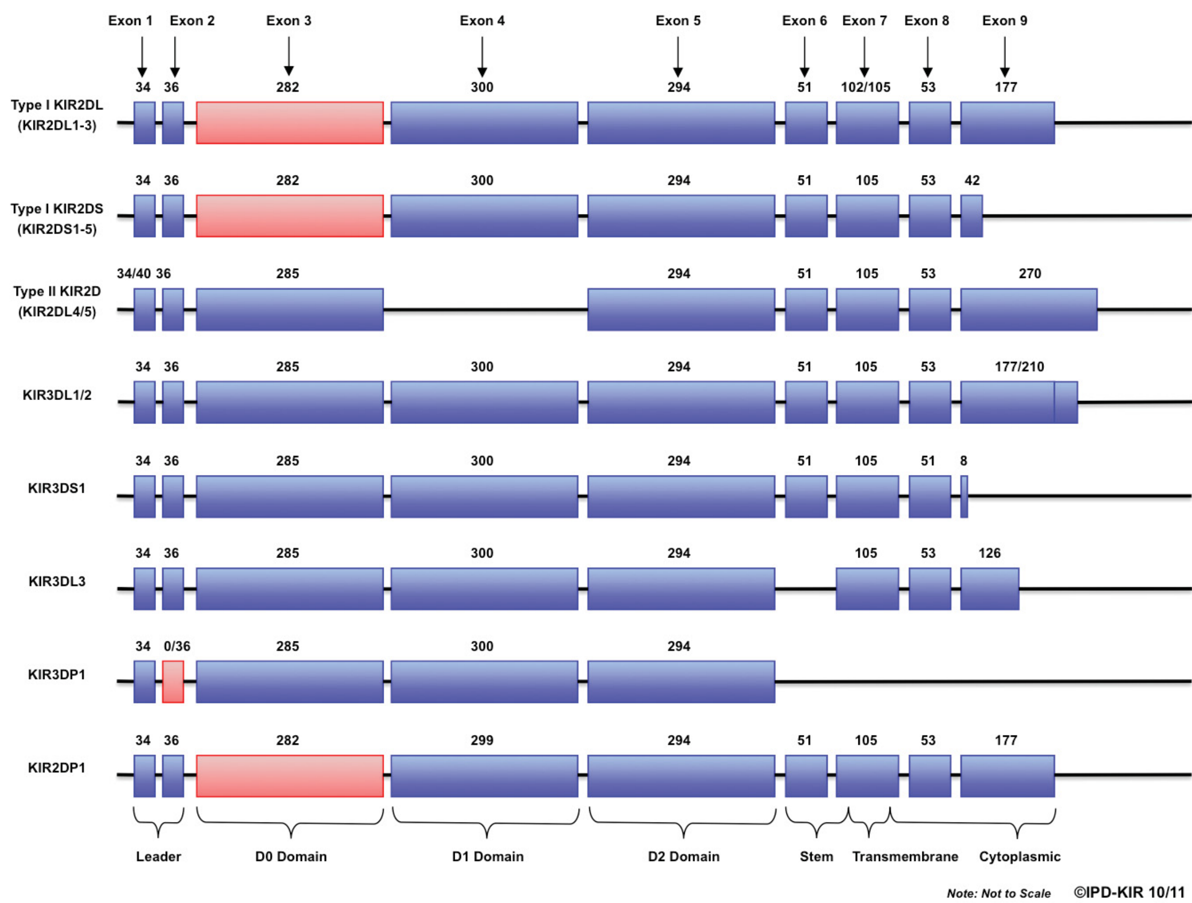
O tipo I contém o pseudogene *KIR2DP1*, assim como *KIR2DL1-3* e *KIR2DS1-5*. Eles possuem 8 éxons, além do pseudo éxon 3 (95–97). Esse pseudo éxon não é funcional, devido em alguns casos a uma substituição nucleotídica ou em outros casos a um códon de parada prematuro (98). Os *KIR* de tipo II incluem os genes *KIR2DL4* e *KIR2DL5*. Nesses genes, a região correspondente ao éxon 4 é ausente, e o éxon 3 é transcrito, ao contrário do que ocorre nos *KIR* de tipo I. Os *KIR* de tipo III possuem 9 éxons e incluem os genes *KIR3DL1*, *KIR3DS1*, *KIR3DL2* e *KIR3DL3*. Esses 4 genes diferem entre si pelo tamanho da região codificadora da cauda citoplasmática. O pseudogene *KIR3DP1* é o que mais difere em estrutura, tendo ausência das regiões do éxon 6 ao 9 (4) (Figura 4).

FIGURA 3. O Complexo Receptor de Leucocitário (LRC) estendido (19q13.4).



Os genes *KIR* são codificados dentro de um trecho de 150 Kb do LRC. O LRC estendido também contém os genes que codificam proteínas adaptadoras DAP, antígenos CD66, bem como receptores SIGLEC, FcGRT, LILR, LAIR, FcAlphaR e NCR1. Fonte: (Robinson et al, 2020) (4).

FIGURA 4. Organização gênica dos diferentes tipos de *KIR*.



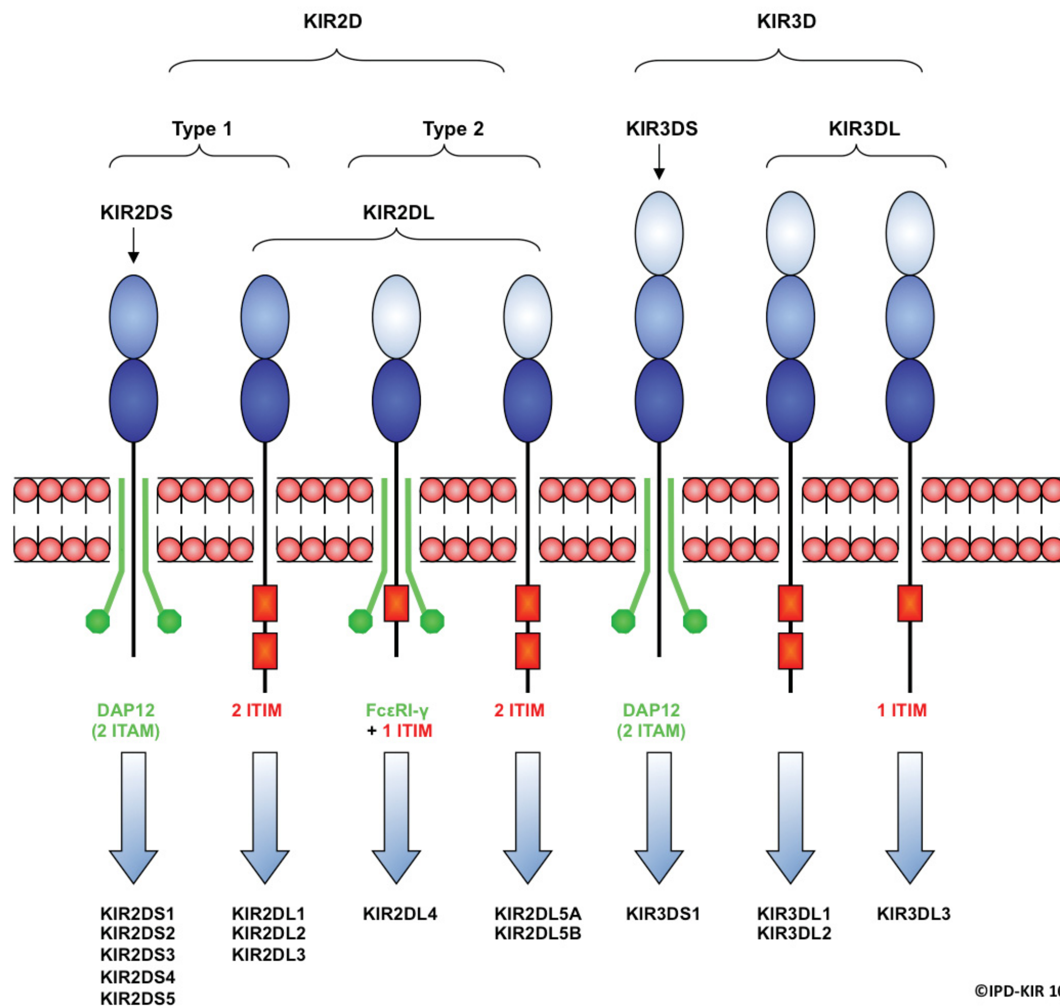
As regiões de codificação dos éxons são representadas por retângulos azuis, e seu tamanho em pares de bases é mostrado acima deles. O pseudoéxon 3 dos *KIR* de tipo I e de *KIR2DP1*, e o éxon 2 de *KIR3DP1* deletado são mostrados em vermelho. Fonte: (Robinson et al, 2020) (4).

Os receptores KIR possuem domínios extracelulares semelhantes a imunoglobulina, que em algumas proteínas estão envolvidas com a função de se ligar as moléculas HLA de classe I. Eles também possuem caudas citoplasmáticas que definem o tipo de sinal transduzido à célula NK. Portanto, os receptores KIR podem conter dois ou três domínios extracelulares, assim como uma cauda citoplasmática curta ou longa. Como regra geral, KIRs de cauda curta transduzem sinais ativadores enquanto os KIR de cauda longa são inibidores. A exceção é KIR2DL4, que apresar de ter cauda longa, possui motivos inibidores a ativadores. Foi demonstrado que KIR2DL4 é capaz de transduzir ambos os sinais, com preponderância de sinais ativadores (99).

Caudas citoplasmáticas longas geralmente contêm dois motivos inibidores baseados em tirosina (ITIM) que transduzem sinais inibidores para a célula NK (94,100). Caudas citoplasmáticas curtas possuem um resíduo de aminoácido carregado positivamente em sua região transmembrana que lhes permite associar-se a uma molécula sinalizadora DAP12 (adaptador transmembrana do tipo tirosina quinase), que contém motivos de ativação baseados em tirosina (ITAM), e é capaz de gerar um sinal de ativação (101,102). DAP-12 também é um membro da superfamília de imunoglobulinas e é codificado na extremidade centromérica do LRC. ITIMs e ITAMs também são característicos de vários receptores imunologicamente importantes, como CD5, CD22 e FcγRII (103) (Figura 5).

A expressão de alguns genes e alelos *KIR* é afetada por variações que podem impedir a transcrição, como é o caso dos pseudogenes *KIR2DP1* e *KIR3DP1* e de alguns alelos de *KIR2DL5* (97). Além disso, alguns genes *KIR* estruturalmente intactos são transcritos, mas apenas expressos em níveis baixos na superfície das células NK por razões que permanecem desconhecidas, como acontece para KIR3DL3 por exemplo (104). A expressão de KIR na superfície celular também não é igual em todas as células. Cada clone de célula NK pode expressar um conjunto diferente de receptores, que podem estar relacionados ao local onde essa célula se encontra, ou a uma resposta a um patógeno (105).

Figura 5. Estrutura das proteínas KIR



Os receptores KIR podem conter 2 ou 3 domínios extracelulares, além de uma cauda citoplasmática curta ou longa. Caudas longas geram sinais inibidores, pois possuem motivos ITIM. Já caudas ativadoras normalmente geram sinais ativadores, devido a associação com moléculas DAP12 que contém motivos ITAM. A exceção a regra são os KIR2DL4, que apesar de possuírem uma cauda longa apresentam motivos inibidores e ativadores. Fonte: (Robinson et al, 2020) (4)

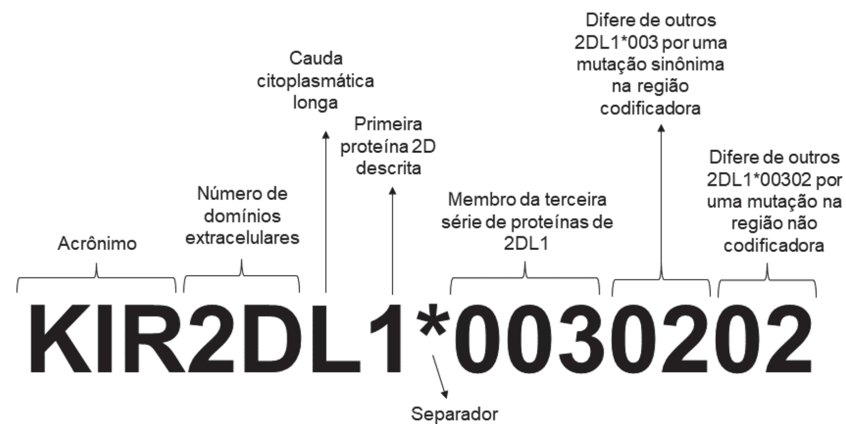
2.6 A NOMENCLATURA DE *KIR*

A nomenclatura dos genes *KIR* é baseada no número de domínios extracelulares presentes (2D ou 3D) e também na natureza da sua cauda citoplasmática (S indicando uma cauda curta, e L indicando uma cauda longa), enquanto a letra P indica a presença de um pseudogene (94,106). Portanto, KIR2DL se refere, por exemplo, a um receptor com dois domínios extracelulares de cauda longa, enquanto KIR3DP se refere a um pseudogene com estrutura similar a um gene codificador de receptor com três domínios extracelulares, por exemplo. Os números 1 a 5 em cada receptor indica a ordem pela qual eles foram descritos (Figura 6).

KIR2DL5A e *KIR2DL5B* excepcionalmente apresentam as letras A e B devido ao fato de que posteriormente à descoberta do receptor KIR2DL5 foi verificado que existem dois genes muito semelhantes, localizados em duas regiões diferentes dentro do haplótipo (um na porção centromérica e outro na telomérica), que posteriormente foram distinguidos em A e B. (107). Além disso, *KIR2DL2* e *KIR2DL3*, assim como *KIR3DL1* e *KIR3DS1*, foram inicialmente descritos como quatro genes independentes, mas posteriormente *KIR2DL23* foi reconhecido como apenas um único loco, da mesma forma que *KIR3DL1S1*.

Para nomenclatura alélica é utilizado um asterisco para separar o sufixo *KIR2D* ou *KIR3D* da identificação específica do alelo. Dependendo da resolução de genotipagem (Ver tópico 2.9) o alelo terá um diferente número de dígitos. Para média resolução são usados 3 dígitos, que representam a série de proteínas do gene. Em alta resolução temos 5 dígitos, onde os dígitos finais representam mudanças sinônimas em regiões codificadoras da proteína. Também podemos ter a genotipagem em alta resolução, na qual temos 7 dígitos, onde os dois dígitos finais vão representar substituições no DNA em uma região não codificante (Figura 6). Portanto, apenas o sequenciamento completo de cada gene, e sem ambiguidades, pode identificar os alelos em máxima resolução. Mais características acerca da diversidade alélica de *KIR* serão exploradas nos próximos tópicos.

FIGURA 6. Nomenclatura dos genes *KIR*



Na figura temos o exemplo de um alelo do gene *KIR2DL1*. A proteína desse gene possui 2 domínios extracelulares (2D), além de uma cauda citoplasmática longa (L), e foi a primeira a ser descoberta dentro dos KIR2DL. O asterisco separa o nome do gene da sua designação alélica. Os 3 primeiros números representam a série de proteínas a qual ele pertence. Os dígitos 4 e 5 representam mutações sinônimas dentro da região codificadora. Os dois últimos dígitos irão representar mutações na região não-codificadora. Fonte: O autor, 2022.

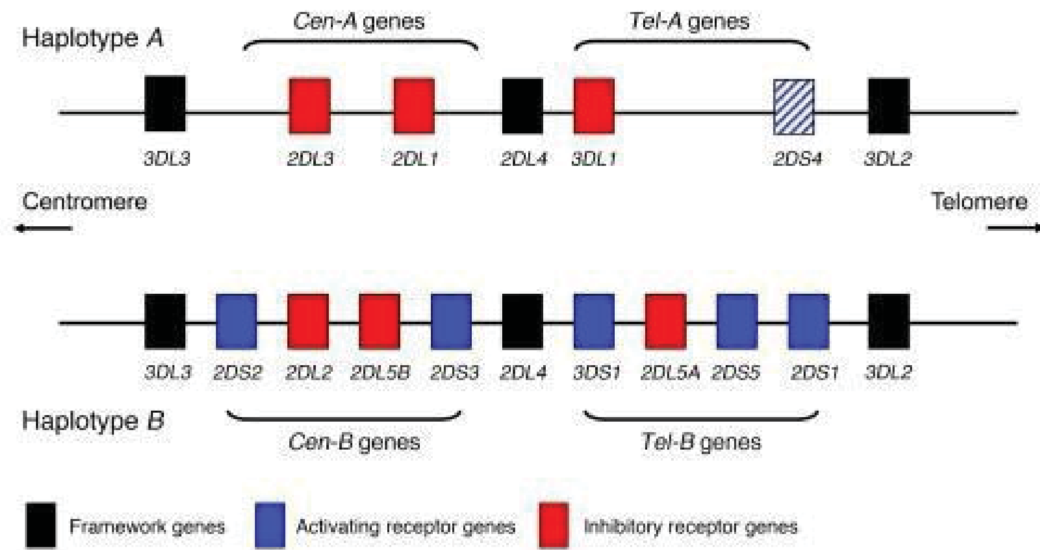
2.7 O POLIMORFISMO DE *KIR*

Os genes *KIR* apresentam um polimorfismo incomum de presença ou ausência, além de um grande número de alelos, e também uma grande diversidade haplotípica, resultando em uma grande diversidade na família desses genes (3,97,108–110). Além disso, *KIR* também possui uma variação de número de cópias, em que cada gente pode conter, normalmente, de 0 a 4 cópias em um mesmo indivíduo.

O polimorfismo de presença e ausência gera haplótipos de conteúdo gênico que podem ser divididos em dois grupos: o haplogrupo A, caracterizado pela presença de apenas um gene ativador de cauda curta (*KIR2DS4)* além dos genes *KIR3DL3, KIR2DL3, KIR2DP1, KIR2DL1, KIR3DP1, KIR2DL4, KIR3DL1,* e *KIR3DL2*; e o haplogrupo B, que contém uma variedade de combinações e genes ativadores e inibidores (3) (Figura 7). Portanto, o haplogrupo B é mais diverso, principalmente por conter mais genes ativadores (111,112).

Um fato curioso é que o único gene ativador de cauda curta presente no haplogrupo A, o *KIR2DS4*, possui um grupo de alelos não funcionais devido a uma deleção de 22pb no éxon 5 do gene, o que leva a produção de uma proteína truncada (113). Esse grupo alélico possui uma frequência alta nas populações mundiais (em Curitiba, por exemplo, a frequência do alelo é de 75% (114). Além disso, *KIR2DL4*, que também tem função ativadora, pode apresentar alelos que codificam receptores não funcionais, neste caso devido a deleção de uma adenina na posição 811, que dá origem a um códon de parada. Essa deleção é comum em diversas populações, como por exemplo na população de Curitiba, na qual a frequência do alelo não funcional é de 57,6% (114). Como haplogrupo A é muito frequente na maioria das populações e esses alelos não funcionais de *KIR2DS4* e *KIR2DL4* são também muito frequentes, muitos indivíduos muitos indivíduos homozigotos para o haplótipos A não possuem nenhum receptor ativador funcional.

FIGURA 7. Conteúdo gênico do haplogrupo A e uma representação de genes que podem estar presentes no haplogrupo B.



O haplogrupo A é caracterizado pela presença de apenas um gene ativador de cauda curta (*KIR2DS4*) além dos genes *KIR3DL3, KIR2DL3, KIR2DP1, KIR2DL1, KIR3DP1, KIR2DL4, KIR3DL1 e KIR3DL2*. O haplogrupo B contém uma variedade de combinações e genes ativadores e inibidores. Os haplótipos de *KIR* são flanqueados pelos genes *KIR3DL3* e *KIR3DL2*, e ao centro são encontrados os genes *KIR2DL4* e *KIR3DP1* (não representado na figura). Por serem encontrados em praticamente todos os haplótipos, esses são chamados de genes moldura. Fonte: (Hiby et al, 2020) (115)

Os haplótipos *KIR* são flanqueados na ponta da região centromérica pelo gene *KIR3DL3* e na região telomérica por *KIR3DL2*. Ao centro, encontram-se *KIR3DP1* e *KIR2DL4* (94,95,116) (Figura 7*)*. Os genes *KIR2DL4*, *KIR3DL2* e *KIR3DL3*, assim como o pseudogene *KIR3DP1* são encontrados praticamente todos os haplótipos e são conhecidos como genes moldura (94). Outros genes como o *KIR2DS3,* por exemplo, são menos frequentes.

As frequências dos genes *KIR* variam em diferentes populações. Em um estudo com uma população de japoneses que residem em Curitiba (PR), por exemplo, foi demonstrado que o haplogrupo mais frequente era o haplogrupo A (60,8%), enquanto genes comumente encontrados em haplogrupos B tiveram frequências baixas, entre 13,7% e 33% (117). Apesar de menos diverso em termos de conteúdo gênico (3,108), o haplogrupo A é o mais frequente em diversas populações como Chineses (118), Sul Coreanos (119), Japoneses (120,121), uma população de Curitiba no Brasil (114), Vietnamitas (122), Tailandeses (123), entre outros. No entanto esse nem sempre é o caso, como mostra Toneva e colaboradores que encontraram frequências de 31,2% em euro-descendentes e de apenas 1,5% em aborígenes australianos, para o mesmo haplogrupo A (124). O quadro 3 exemplifica a variação na frequência de presença e ausência de genes *KIR* em diferentes populações.

Apesar de centenas de genótipos *KIR* terem sido descritos (125), eles são na maioria das vezes variações de sete haplótipos de conteúdo gênico. Esses haplótipos são combinações de quatro motivos centroméricos e dois motivos teloméricos (126), e são responsáveis por 90% da variação haplotípica de *KIR* encontrada na maioria das populações (109,127,128). Apesar de menos frequentes, nossos resultados (109) mostram que em estudos com número amostral suficiente, alguns haplótipos anteriormente considerados raros são observados com frequências entre 0,05% e 1% (109). Esses haplótipos incomuns são geralmente frutos de duplicações, gerando números de cópia de *KIR* que fogem do padrão. Por exemplo, o haplótipo *tB04* possui duas cópias de *KIR2DL4*, fazendo com que o indivíduo que o possua apresente no total três cópias de *KIR2DL4*, um número que foge do padrão de duas cópias geralmente apresentado por esse gene (Figura 8).

QUADRO 3. FREQUÊNCIAS DE *KIR* EM DIFERENTES POPULAÇÕES.

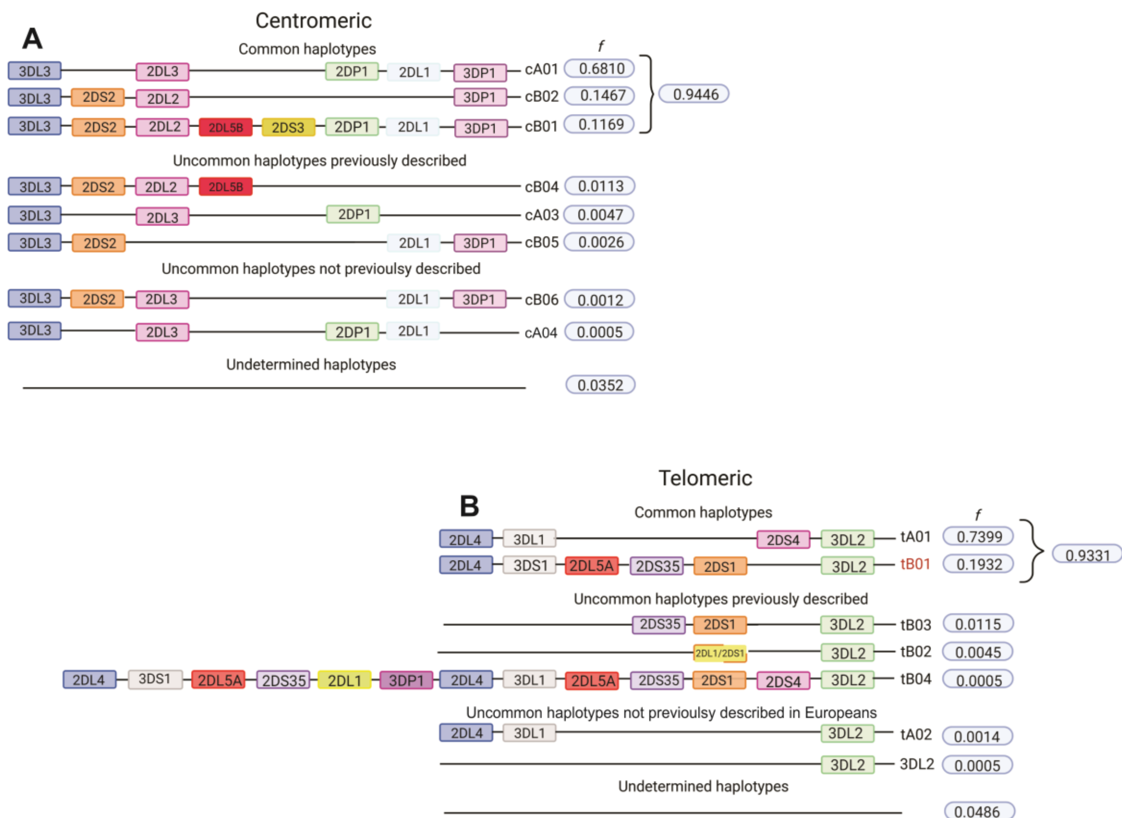
		Curitiba		Czech	BHorizonte	Argentina	France	Ireland	Palestine	India	USAfricans	Senegal	Camoros	Singapore	Hong Kong	Japan	China	S. Korea	Amazon
n		153		125	90	102	108	200	105	145	58	118	54	47	100	132	106	154	40
	F (%)	f_G	f_A	*F* (%)															
2DL1	97.4	0.84	–	95.0	96.7	96.1	97.0	96.0	83.0***	99.3	98.3	100.0	100.0	100.0	99.0	100.0	99.0	99.4	93.0
2DL2	59.4	0.36	0.36	59.0	52.2	62.5	50.0***	51.0	62.0	62.8	56.9	55.0	41.0*	28.2***	28.0***	11.4***	17.3***	14.3***	65.0
2DL3	84.6	0.61	0.55	86.0	94.4*	86.5	91.0***	93.0*	85.0	81.9	82.0	90.0	93.0	100.0**	98.0***	100.0***	99.0***	99.4***	80.0
2DL4	100.0	1.00	–	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2DL5	51.9	0.31	–	52.0	58.9	55.7	47.0	44.0	63.0	71.0***	53.4	52.0	62.7	39.1	45.0	35.6**	68.7*	38.3**	85.0***
2DP1	94.7	0.77	–	94.0	96.7	96.1	97.0	97.0	NT	99.3	98.3	100.0	100.0	100.0	99.0	100.0	99.0	NT	NT
2DS1	42.1	0.24	–	43.0	37.8	45.4	36.0	34.0	44.0	62.8***	22.4	13.0***	17.0***	28.3	40.0	33.3	34.3	37.7	88.0***
2DS2	59.2	0.36	–	57.0	53.3	54.5	51.0	52.0	64.0	62.8	46.6	42.0**	30.0***	28.3***	28.0***	11.4***	17.3***	16.9***	58.0
2DS3	32.5	0.18	–	36.0	38.9	28.6	31.0	28.0	37.0	53.8***	27.6	24.0	20.0	17.4*	25.0	13.6***	12.5***	16.2***	10.0**
2DS4	93.6	0.75	–	92.0	95.6	95.0	96.0	95.0	88.0	86.1***	100.0	100.0**	96.0	97.8	94.0	99.2	94.2	94.2	78.0**
2DS5	31.9	0.17	–	26.0	32.2	35.4	27.0	26.0	27.0	51.0***	37.9	30.0	37.0	21.7	26.0	22.0	23.1	26.6	90.0***
3DL1	92.7	0.73	0.74	94.0	95.6	95.0	96.0	95.0	88.0	87.4	100.0*	99.0*	98.0	97.8	94.0	99.2**	94.2	94.2	65.0***
3DS1	37.3	0.21	0.26	38.0	41.1	42.0	44.0	35.0	39.0	62.2***	13.8***	4.0***	15.0**	30.4	39.0	31.9	32.7	36.4	70.0***
3DL2	100.0	1.00	–	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
3DL3	100.0	1.00	–	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

The asterisks indicate p values for the populations that differed significantly from Curitiba: $^{*}0.01 \leq p < 0.05$; $^{**}0.001 \leq p < 0.01$, and $^{***}p < 0.001$. References are listed in Table 3. The populations are named as in the dendrogram

n number of individuals in the sample, *F* carrier frequency, f_G gene frequency obtained by Bernstein's method, f_A the gene frequency, considering the allelic status of *2DL2* and *2DL3*, and of *3DL1*, *3DS1*, *NT* not tested

A figura exemplifica a frequência de presença e ausência de *KIR* em diversas populações mundiais, mostrando que populações diferentes apresentam diferentes frequências de *KIR*. Fonte: (Augusto et al, 2012) (114)

FIGURA 8. Conteúdo gênico dos haplótipos centroméricos e teloméricos de *KIR* em uma população euro descendente dos Estados Unidos.



Haplótipos de conteúdo gênico encontrados em uma população euro descendente dos Estados Unidos, mostrando que a maior parte da variação haplotípica encontrada pode ser reduzida a poucos haplótipos muito frequentes. No entanto, com um número amostral suficiente podemos observar haplótipos incomuns que antes eram negligenciados nas análises de *KIR*. Fonte: (Amorim et al, 2021) *(109)*

Apesar do polimorfismo de presença e ausência de *KIR* ter sido relativamente bem explorado em diversas populações, o número de estudos explorando o polimorfismo alélico é muito reduzido, caindo drasticamente o número de populações descritas (109,128,129). Apenas dois estudos haviam analisado a diversidade alélica de *KIR* em populações até 2021. Porém, ambos estudos analisaram a variação de *KIR* em resolução média (3 dígitos) para todos os genes. A falta de estudos da variação alélica de *KIR* se deve principalmente à dificuldade técnica de sequenciamento inerente à complexidade da variação desses genes, assunto que será abordado a seguir.

2.8 A COMPLEXIDADE DOS GENES *KIR* E OS DESAFIOS PARA GENOTIPAGEM

Como mencionado anteriormente, os genes *KIR* são uma família de genes localizadas no complexo LRC no cromossomo 19. Com base em elementos *alu*, que podem ser considerados como um relógio molecular, estima-se que a expansão inicial do *cluster* de genes *KIR* remonta a aproximadamente 31 a 44 milhões de anos atrás (130). Os genes da família *KIR* encontram-se justapostos e possuem elevada semelhança em sua sequência gênica e intergênica, o que sugere que esses genes derivaram de um ancestral comum e sofreram subsequentes duplicações (86) e processos recombinatórios (131). A região desse complexo varia em tamanho de 100 a 350kb como resultado de haplótipos estruturalmente diversos, com segmentos duplicados, grandes deleções, e fusões gênicas (132,133). Além disso, estes genes possuem uma grande variedade alélica, indicando uma rápida evolução dessa família de genes (134).

Características cruciais que distinguem os alelos de *KIR* da maioria dos outros genes são a profundidade, amplitude e importância funcional das divergências em suas sequências. Assim, os alelos podem diferir por múltiplas substituições de nucleotídeos, e três ou quatro nucleotídeos alternativos podem estar presentes em posições funcionalmente críticas, o que dificulta a determinação dos alelos durante a genotipagem (135). A alta identidade de sequência causada pela homologia desses genes acrescenta uma barreira extra para o desenho de oligonucleotídeos inciadores (*primers*) e sondas específicas.

Ainda, a extensiva variação estrutural de conteúdo gênico e grandes deleções e inserções nessa região são observadas em níveis incomuns em relação ao restante do genoma. Toda essa complexidade impediu por muito tempo que a diversidade alélica de *KIR* fosse explorada, e, portanto, por muitos anos os estudos dessa família gênica ficaram restritos ao polimorfismo de presença e ausência.

A genotipagem de presença e ausência de *KIR* foi descrita pela primeira vez por Uhrberg em 1997, usando uma abordagem *PCR-SSP* que exigia DNA de alta qualidade devido ao comprimento dos *amplicons* produzidos (~1,5 kb) (108). Em 2001, Norman e colegas desenvolveram um novo *primer* de *KIR2DS3* para superar a amplificação fraca obtida com o *primer* do método anterior (136). Em 2002, Gomez-Lozano e Vilches descreveram um novo método de *PCR-SSP*, parcialmente baseado em Uhrberg e colaboradores (108) que adicionou mais genes *KIR* (*KIR2DL5*, *KIR3DL3*, *KIR3DP1* e *KIR2DP1*)(107). Para superar a limitação de tamanho dos *amplicons*, outros métodos foram desenvolvidos com a amplificação de fragmentos menores. Sun *et al* em 2004 desenvolveram um método multiplex (quatro reações para amplificar todos os genes *KIR*, incluindo o alelo de *KIR2DS4* com deleção de 22pb) com a amplificação de *amplicons* menores (108-565 pb) (137). Em 2007, Vilches e colaboradores atualizaram seu método anterior com novo conjunto de *primers* que produziam fragmentos mais curtos (menos de 200pb) (138).

Os métodos descritos anteriormente dependiam de um único par de *primers* para genotipagem. No entanto, considerando o polimorfismo alélico e a similaridade de sequência para cada *KIR*, usar um único par de *primers* para genotipagem de *KIR* aumenta a chance de resultados imprecisos. Portanto, em 2008, Martin & Carrington descreveram um novo protocolo de genotipagem *KIR* que utilizava dois pares de *primers* para genotipar cada gene, usando *HLA-DRB1* como controle interno para cada reação (139). Aplicando o mesmo conjunto de *primers*, Kulkarni *et al* desenvolveram um método multiplex contendo 12 reações (140). Usando *primers* previamente descritos e seus próprios *primers*, Ashouri *et al* criaram um novo método duplex de *PCR-SSP*, que reduziu o número de reações de 16 para 8 (141). Apesar de ser econômico e rápido, esse método não foi projetado para pequenos conjuntos de amostras, pois utiliza quatro temperaturas de PCR diferentes. Tentando reduzir o número de reações, Abalos

e colegas criaram um novo método *multiplex*, usando *primers* próprios (142). Apesar de reduzir a genotipagem para quatro reações, esse método foi baseado em um único par de *primers* para cada gene.

Alternativamente, os *KIR* também foram genotipados por PCR com sonda oligonucleotídica específica (*PCR-SSOP*). Esse método foi descrito pela primeira vez em 2000 por Crum e colaboradores (143), utilizando 26 sondas marcadas com digoxigenina para identificar 12 genes *KIR*. Em 2007, Nong *et al* publicaram um novo método de *PCR-SSOP* que utilizava 62 sondas para detectar 14 genes *KIR* e 2 pseudogenes (144). Além disso, este método incluía um software de análise para auxiliar na atribuição de tipagem *KIR*. Apesar da sensibilidade da *PCR-SSOP*, as sondas marcadas são relativamente mais caras que os *primers* de *PCR-SSP*, além de exigirem um considerável processamento pós-PCR.

Todas as técnicas de genotipagem *KIR* ora mencionadas envolveram métodos de *PCR-SSP* ou *SSOP* baseados em gel, que são demorados devido ao extenso processamento pós-PCR e propensos a erros quando usados em grandes conjuntos de amostras. Para evitar esses problemas, Koehler e colaboradores descreveram um método de genotipagem de alto rendimento baseado em PCR em tempo real em 2009 (145). No entanto, esse método está disponível apenas para *KIR2DL2/L3 e KIR3DL1/S1*, juntamente com seus respectivos ligantes HLA de classe I. No mesmo ano, Alves e colaboradores desenvolveram um método baseado na análise da curva de *melting* para analisar a presença e ausência de 16 genes *KIR* (146). Eles usaram a tecnologia *Sybr green I* para evitar o processamento pós-PCR das amostras com eletroforese em gel. Embora este método seja mais rápido do que os métodos *SSP* anteriores, ele utiliza apenas um conjunto de *primers* para cada gene, descritos por Vilches *et al* em 2007 (138).

Em 2011, Hong *et al* descreveram um novo método baseado na curva de *melting*, também utilizando a tecnologia *Sybr green* (147). Utilizando *primers* previamente descritos (por Martin *et al*, Vilches *et al* e Alves *et al*) os autores foram capazes de detectar a presença ou ausência de *KIR* de 14 genes e 2 pseudogenes. Este método também incluiu um ensaio para detectar ligantes HLA classe I. Embora esses métodos tenham sido consideravelmente mais rápidos que *SSP* e *SSOP* convencionais, eles ainda apresentavam um grande

número de reações e requerem análise manual, o que não é ideal para estudos com um grande conjunto de amostras.

Para superar esses problemas, Amorim e colaboradores desenvolveram um método *multiplex* baseado em curva de *melting* (148). Este método desenvolvido por nosso grupo de pesquisa possibilita a genotipagem de presença e ausência de 14 genes *KIR* e 2 pseudogenes, juntamente com a detecção da deleção de *KIR2DS4* de 22bp, utilizando uma placa de 384 poços, o que permite a genotipagem simultânea de 16 amostras. Esse método é marcado por uma alta sensibilidade e precisão, além de custos inferiores a 1,65 *USD* por amostra. Para evitar erros inerentes à análise manual de muitos dados, nosso grupo também desenvolveu um programa de análise automatizada, que reduz a probabilidade de erros de genotipagem e o tempo de genotipagem (148).

Métodos alternativos também foram desenvolvidos ao longo do tempo. Em 2007, Houtchens *et al* desenvolveram um método de genotipagem baseado em espectrometria de massa, ou detecção de 17 genes *KIR* e pseudogenes, incluindo também a discriminação de alelos para alguns deles (149). Em 2016, um método de PCR em tempo real de alto rendimento (qKAT) foi descrito para variação do número de cópias e determinação de haplótipos (150). Apesar da melhoria nos métodos de genotipagem de *KIR* com o passar dos anos, nenhum dos métodos citados até agora conseguia explorar a diversidade alélica e de número de cópias de *KIR* com precisão.

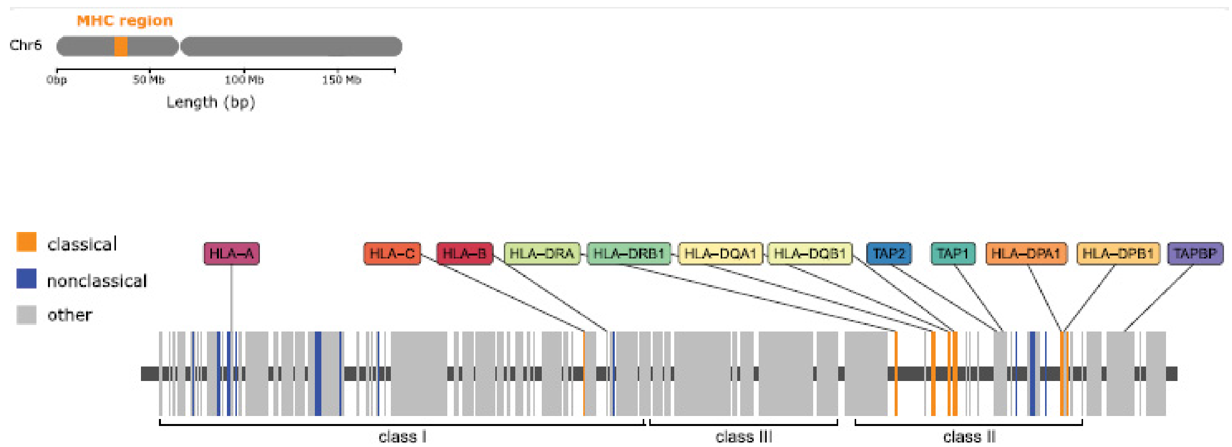
Em 2016 foi desenvolvido o primeiro método de sequenciamento de *KIR* em larga escala baseado em sequenciamento de nova geração (NGS) (135). Foram projetados conjuntos de sondas oligonucleotídicas específicas para capturar a região *KIR* (140-240 kb) de bibliotecas preparadas a partir de DNA genômico. No entanto, a identidade de sequência de *KIR* tornou-se uma barreira significativa para a interpretação de fragmentos curtos gerados pela plataforma Illumina (cerca de 150-250 pb). Mais recentemente, um novo trabalho apresentou uma pipeline de bioinformática (PING, do inglês *Pushing Immunogenetics to the Next Generation*) especificamente para converter dados de sequência curta obtidos de genes *KIR* em genótipos de alta resolução (135,151). Esse método tem a capacidade de, além de identificar a presença e ausência de *KIR*, identificar com precisão o número de cópias de cada gene e determinar com poucas ambiguidades a variação alélica em alta resolução (7

dígitos). Esse método foi utilizado neste trabalho e será detalhado na seção de métodos.

2.9 O RECONHECIMENDO DOS ANTÍGENOS LEUCOCITÁRIOS HUMANOS (HLA) POR KIR

Os genes *HLA* estão localizados em uma região do cromossomo 6 humano chamada MHC (complexo principal de histocompatibilidade, do inglês, *major histocompatibility complex*). No MHC estão localizados cerca de 170 genes, dos quais aproximadamente 50% produzem moléculas com funções relacionadas ao sistema imune (152). O MHC é dividido em três regiões, chamadas de classe I, II e III. A região de classe I se encontra na porção telomérica do MHC, a região de classe II na porção centromérica, e entre as duas regiões encontra-se a região de classe III (153). Os genes *HLA* estão localizados nas regiões de classe I e classe II (Figura 9). No total, existem genes *HLA* de classe I clássicos (*HLA-A, HLA-B* e *HLA-C*) e não clássicos (*HLA-E, HLA-F* e *HLA-G*), além de muitos pseudogenes. Já os *HLA* de classe II são 21 (*HLA-DRA, HLA-DRB1-9, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DPA1, HLA-DPA2, HLA-DPB1, HLA-DPB2, HLA-DMA, HLA-DMB, HLA-DOA* e *HLA-DOB*) (4). A região de classe III, por sua vez, contém genes que codificam para moléculas reguladoras imunes, como por exemplo o fator de necrose tumoral (TNF), fatores C3, C4 e C5 do complemento e proteínas de choque térmico (152).

FIGURA 9. A região do MHC.

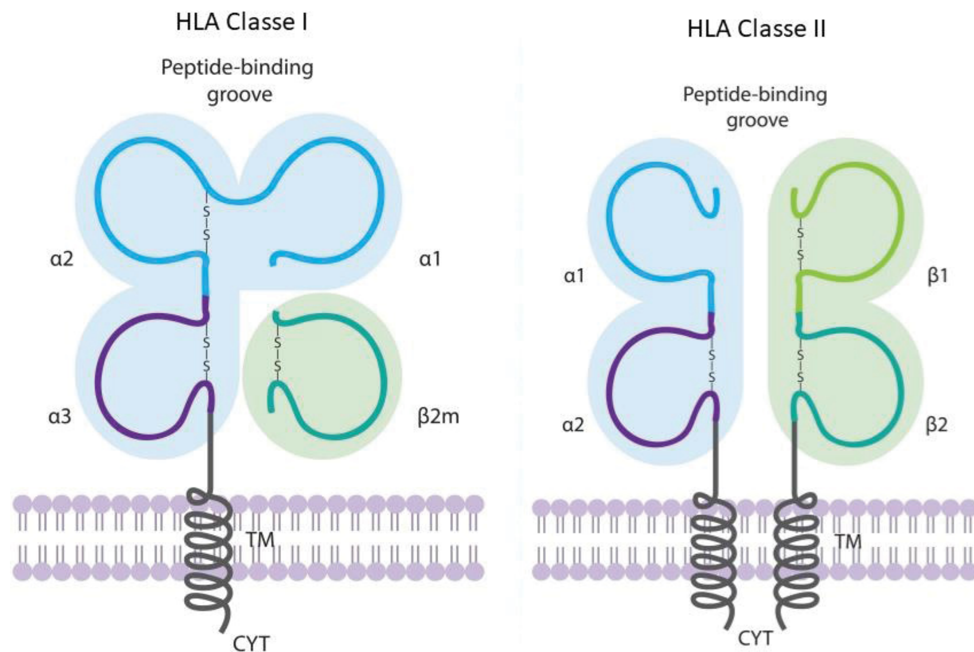


Região do complexo dos genes do MHC, localizado no braço curto do cromossomo 6. Os genes HLA estão localizados apenas nas regiões de Classe I e Classe II, nas porções centromérica e telomérica da região. FONTE: adaptado de (RADWAN et al., 2020) (154).

As moléculas HLA de classe I possuem uma cadeia α ligada de modo não covalente a uma cadeia monomórfica β2-microglobulina, que é codificada por um gene não *HLA* localizado no cromossomo 15 (155). A cadeia α da molécula é dividida em domínios α1, α2 e α3. Os domínios α1 e α2 da proteína interagem entre si formando uma fenda de ligação a peptídeos, e é nessa região que está localizada grande parte da variação da molécula HLA de classe I, pois é nela que se ligam os diferentes peptídeos que serão reconhecidos pelos linfócitos T citotóxicos CD8+. O domínio α3 é o local onde ocorre a interação com a β2-microglobulina, apresentando uma sequência conservada que é responsável pela interação com as moléculas CD8 presentes na membrana de células T citotóxicas. As moléculas HLA de classe I são expressas em todas as células nucleadas e têm como principal função apresentar os peptídeos derivados de antígenos do citosol às células T CD8+(156). A estrutura da molécula HLA de classe I está representada na figura 10.

As moléculas HLA de classe II são compostas por duas cadeias polipeptídicas associadas não covalentemente, uma cadeia α e uma cadeia β (Figura 10) (157). Os genes que codificam ambas as cadeias de moléculas de classe II estão presentes no MHC. Ambas as cadeias têm três regiões, incluindo uma região citoplasmática contendo um sulco de ligação ao peptídeo formado pelos domínios α1 e β1, uma região transmembrana contendo aminoácidos hidrofóbicos pelos quais a molécula é ancorada na membrana celular e os domínios α2 e β2 altamente conservados ao qual o TCR se liga (156). As moléculas de classe II têm expressão restrita às células apresentadoras de antígeno, e processam peptídeos exógenos para apresentação a células T auxiliares CD4+.

FIGURA 10. Estrutura das moléculas HLA de classe I e classe II



Na molécula HLA de classe I, os domínios α1 e α2 da proteína interagem entre si formando uma fenda de ligação a peptídeos, enquanto o domínio α3 é o local onde ocorre a interação com a β2-microglobulina, apresentando uma sequência conservada que é responsável pela interação com as moléculas CD8 presentes na membrana de células T citotóxicas. A região transmembrana (TM) contém aminoácidos hidrofóbicos pelos quais a molécula é ancorada na membrana celular. Na molécula de classe II as cadeias α1 e β1 formam um sulco de ligação ao peptídeo, enquanto os domínios α2 e β2 são responsáveis pela ligação ao TCR. Ambas as moléculas possuem uma cauda citoplasmática representada na figura por CYT. Fonte: Adaptado de (Tapias et al, 2013).

Outra função das moléculas HLA de classe I é a ligação com receptores KIR das células NK. Nos próximos parágrafos serão abordados os *HLA* de classe I dentro do contexto de interação com KIR. Os principais ligantes para os receptores KIR são os HLA de classe I clássicos (HLA-A, HLA-B e HLA-C), mas alguns ligantes podem ser do tipo não clássico (HLA-F e HLA-G) (86,134). Cada loco *HLA* clássico possui milhares de alelos (4) e essa é a família gênica mais polimórfica e polialélica do genoma humano.

HLA-B é o gene que possui a maior quantidade de alelos, passando dos 9000 alelos descritos (4). As moléculas HLA-B podem ser divididas em diferentes grupos, de acordo com epítopos que apresentam. Por exemplo, alótipos HLA-B podem ser caracterizados por polimorfismos nos resíduos 77-83 do domínio α1 no sitio de ligação do peptídeo, que caracterizam os epítopos Bw4 e Bw6 (158,159). Cerca de 40% das moléculas HLA-B apresentam o epítopo Bw4. O epítopo Bw4 pode ainda ser dividido em Bw4 80I, contendo um resíduo de isoleucina na posição 80, ou Bw4 80T, contendo um resíduo de treonina na mesma posição (160).

Os epítopos Bw4 80T são divididos em 80TA (resíduo de arginina na posição 81) e 80TL (resíduo de leucina na posição 81). Assim como HLA-B, as moléculas de HLA-A também podem apresentar o epítopo Bw4 (161). Dentre as moléculas de HLA-A contendo o epítopo Bw4 estão HLA-A*23, HLA-A*24, HLA-A*25 e HLA-A*32, no entanto, a molécula HLA-A*25 não serve como ligante de KIR (162). Todos as moléculas HLA-A que contém o epítopo Bw4 são Bw4 80I.

As moléculas de HLA-C podem ser divididas em dois grupos. O grupo C1 contém um resíduo de asparagina na posição 80, enquanto C2 tem uma lisina nesta mesma posição (163). Esses grupos são conhecidos ligantes de receptores KIR. As interações entre KIR e HLA serão discutidas a seguir.

2.10 INTERAÇÕES KIR-HLA E SEU IMPACTO FUNCIONAL E EVOLUTIVO

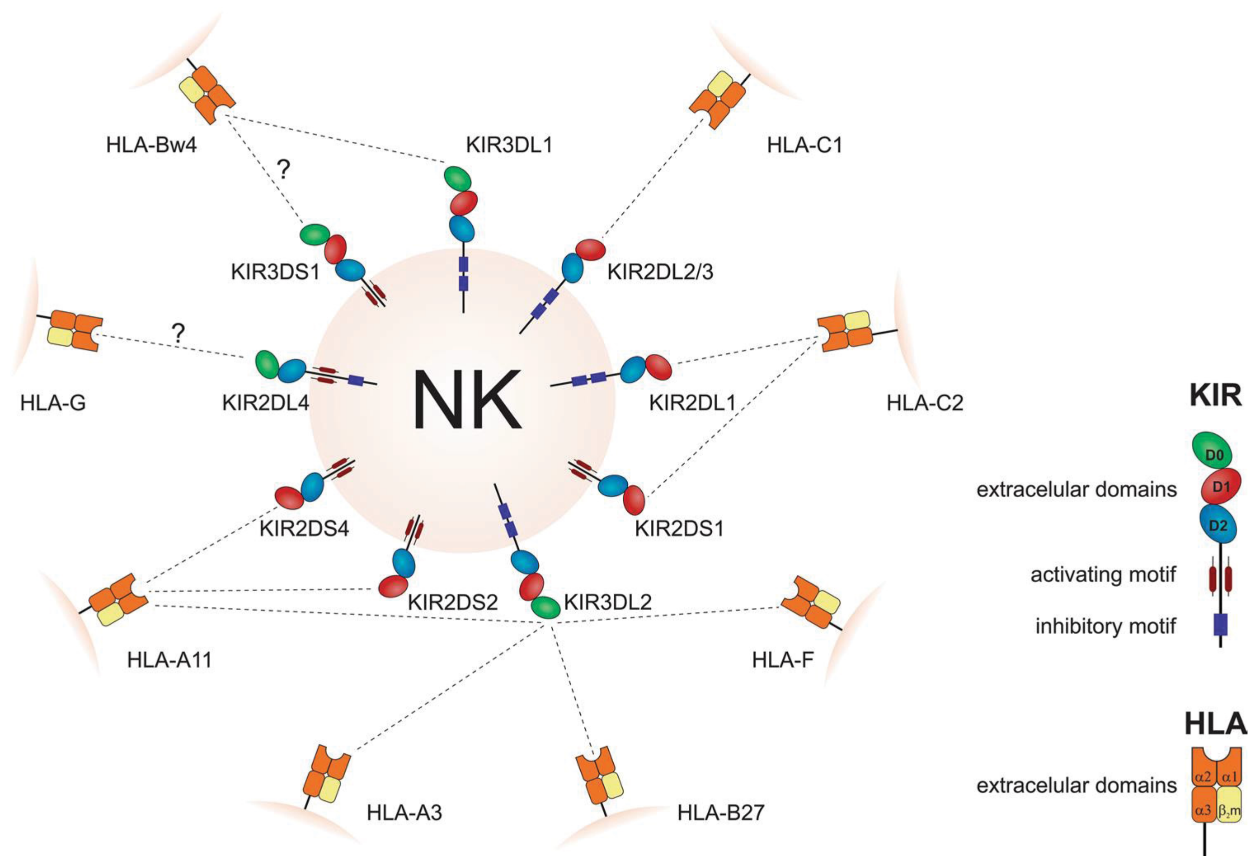
Epítopos Bw4, tanto em moléculas HLA-A quanto HLA-B são reconhecidos por KIR. Porém, sabe-se que moléculas contendo Bw4 80I apresentam interações mais fortes do que Bw4 80T, aumentando o potencial de inibição (164). Algumas das interações entre KIR-HLA são conhecidas e serão brevemente sumarizadas a seguir. O produto do gene *KIR2DS2* se liga a

moléculas de HLA-A*11 (165). Já os receptores KIR2DL1, KIR2DL2 e KIR2DL3 reconhecem moléculas de HLA-C. Porém, enquanto o receptor KIR2DL1 se liga ao grupo C2, KIR2DL2 e KIR2DL3 se ligam ao grupo C1 (163,166).

O receptor KIR3DL2 se liga a diferentes moléculas HLA, como HLA-A*03 e HLA-A*11 (167), HLA-B*27 (168) e HLA-F. KIR3DL1 interage com moléculas HLA-B do grupo Bw4 (mais fortemente com Bw4 80I) (164,169), e alguns HLA-A que contém Bw4, exceto o HLA-A*25 (162). Moléculas Bw4 foram no passado consideradas ligantes de KIR3DS1 devido homologia desse gene com KIR3DL1. No entanto, após diversas tentativas frustradas de evidências diretas dessa interação (170–172), foi demonstrado que somente o alótipo KIR3DS1*014 reconhece HLA-B como ligante *in vivo* (173). KIR3DS1 pode também interagir com HLA-F (174).

KIR2DS1 interage com moléculas HLA-C do grupo C2, porém com baixa afinidade (175). Já KIR2DS4 interage com HLA-A*11 e HLA-F. Essas interações de KIR2DS4 podem vir do fato deste receptor ser derivado de KIR3DL2 (176). A figura 11 mostra os receptores KIR e seus principais ligantes HLA conhecidos.

FIGURA 11. Receptores KIR e seus respectivos ligantes HLA conhecidos.



O símbolo de interrogação mostra ligações que ainda são controversas. A figura também mostra a estrutura das moléculas KIR. Elas podem conter dois ou três domínios extracelulares (D0, D1 e D2). Os receptores inibidores têm uma cauda citoplasmática longa, com motivos inibidores (ITIM), enquanto receptores ativadores possuem motivos ativadores (ITAM). Fonte: (Petzl-Erler & Augusto, 2015) (177).

O estudo dessas interações se faz necessário pois elas podem estar associadas à susceptibilidade ou proteção a doenças (Tabela 1). Por exemplo, a presença de *KIR3DS1* juntamente com a presença de ligantes Bw4, apresentou uma forte associação (no caso proteção) à doença autoimune pênfigo foliáceo (178), assim como a presença do alelo *KIR3DL2*001* e o ligante HLA-A3 ou A11 conferem risco a PF (179). Já a interação entre KIR2DL3+C1 mostrou-se protetora ao desenvolvimento da dengue (180). Geralmente a presença de genes *KIR* ativadores está associada a uma maior suscetibilidade à autoimunidade e a uma maior proteção contra doenças infecciosas, enquanto a presença de genes inibidores confere proteção à autoimunidade e suscetibilidade a doenças infecciosas (177). Além de doenças autoimunes e infecciosas, a interação entre *KIR* e *HLA* já foi demonstrada como fator de risco ou proteção ao desenvolvimento de diversos tipos de tumores, tema que será abordado no tópico 2.11.

A relação de *KIR* e *HLA* com diversas doenças sugere que estes estejam sobre forte pressão seletiva, e existe uma vasta coletânea de evidências de co-evolução destas duas famílias em diferentes populações (177,181–183). Além disso, as combinações KIR-HLA também são importantes para a placentação e suprimento adequado de sangue materno para o feto e, consequentemente, para o sucesso reprodutivo (184). Em outras palavras, as combinações KIR-HLA podem estar sob pressão seletiva não apenas por causa de sua função imunológica, mas também por sua influência direta na reprodução.

A família dos genes *KIR* passou por uma rápida evolução, marcada por eventos de crossing over desiguais e duplicações gênicas (131). Esses eventos evolutivos, tanto em *KIR* quanto em *HLA*, coincidem em sua cronologia, sendo uma das evidências de que essas duas famílias podem estar co-evoluindo (185). Outra evidência é o fato de que HLA-C possivelmente evoluiu mudando sua função primária de apresentação antigênica para células T tornando-se primariamente um ligante de KIR (186). Abi-Rached e colaboradores demonstraram que resíduos chave de KIR2DL23 que interagem com HLA-C1 estão sob seleção positiva (187). Hollenbach e colaboradores também identificaram correlações entre KIR2DL3 e HLA-C1 em 105 diferentes populações (188).

TABELA 1. Exemplos de doenças associadas a combinações de KIR-HLA

Table 1 Examples of associations of KIR–HLA combinations with diseases

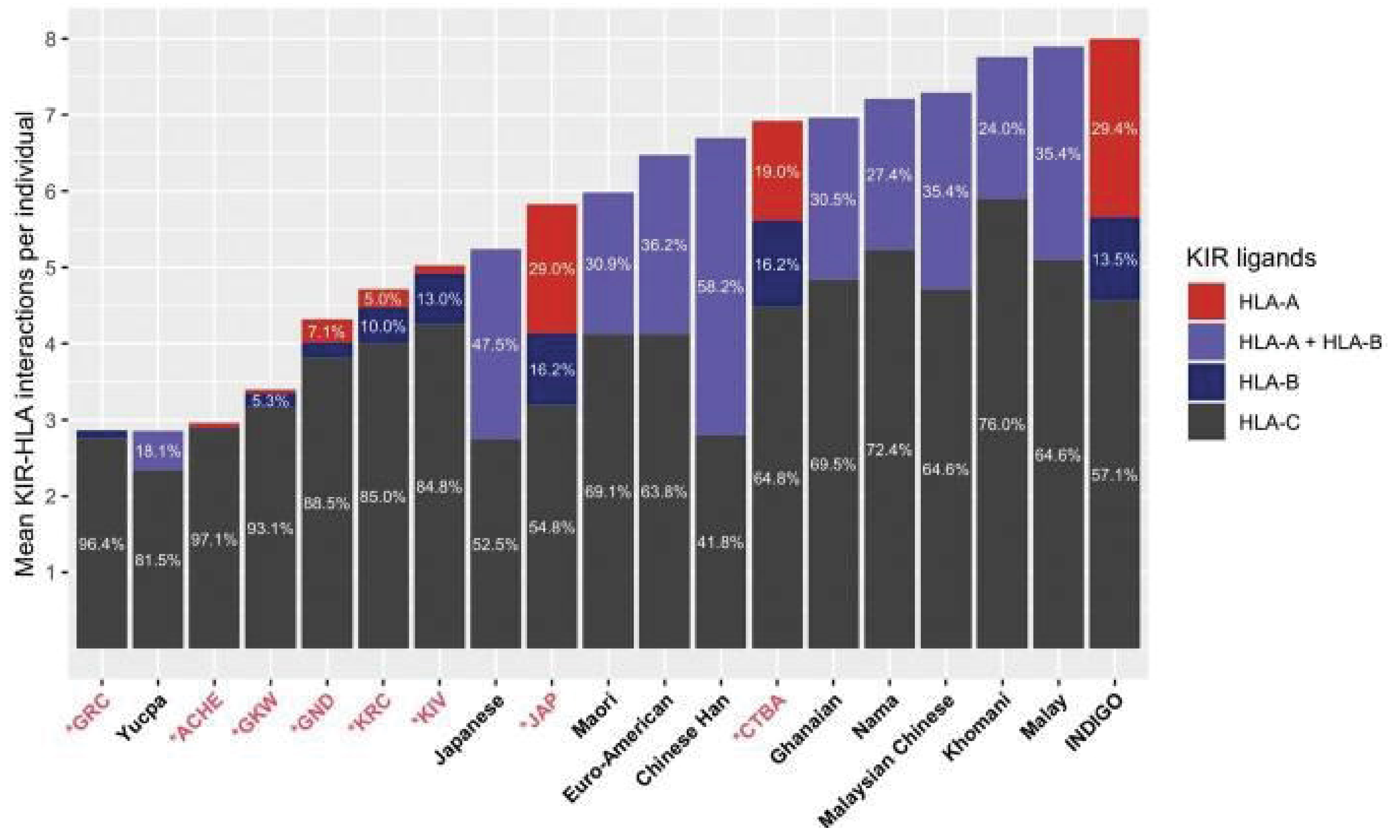
KIR–HLA	Disease	Effect	References
2DL1-C2	Oral squamous cell carcinoma	Increased risk	Dutta et al. (2014)
2DL2/3+C2	Conjunctival scarring in trachoma	Increased risk	Roberts et al. (2014)
2DL2+C1	Leprosy	Increased risk	Jarduli et al. (2014)
	Type 1 diabetes mellitus	Increased risk	Shastry et al. (2008)
2DL3+C1	Co-occurrence of type 1 diabetes and celiac disease	Increased risk	Smigoc Schweiger et al. (2014)
	Dengue	Protection	Beltrame et al. (2013)
	Malaria	Increased risk	Hirayasu et al. (2012)
	Maternal–infant HIV-1 transmission	Increased risk	Paximadis et al. (2011)
	Melanoma	Protection	Campillo et al. (2013)
2DS1+C2	Ankylosing spondylitis	Increased risk	Jiao et al. (2008)
	Autism	Increased risk	Torres et al. (2012)
	Lupus	Increased risk	Hou et al. (2015)
2DS2+C1	Dry eye disease	Increased risk	Ren et al. (2012)
	Type 1 diabetes mellitus	Increased risk	van der Slik et al. (2003)
*3DL1*004*+Bw4	HIV-1	Slow progression	Martin et al. (2007)
*3DL2*001*+A3/A11	Pemphigus foliaceus	Increased risk	Augusto et al. (2015b)
3DS1+Bw4	Autism	Increased risk	Torres et al. (2012)
	Endometriosis	Protection	Kitawaki et al. (2007)
	HCV-related hepatocarcinoma	Increased risk	De Re et al. (2015)
	HIV-1	Reduced replication	Alter et al. (2007)
	HIV-1	Slow progression	Martin et al. (2002)
	Pemphigus foliaceus	Protection	Augusto et al. (2012a)
	Tuberculosis	Protection	Shahsavar et al. (2012)
	Type 1 diabetes mellitus	Increased risk	Jobim et al. (2010)

A tabela informa a associação de *KIR* e HLA com diversas doenças, dentre elas doenças autoimunes, doenças infecciosas e câncer. Fonte: (Petzl-Erler & Augusto, 2015) (177)

Um exemplo da importância de KIR na reprodução foi demonstrado por Hiby e colaboradores(189). Esse trabalho mostrou que a homozigozidade materna de haplótipos A e a presença de HLA-C2 no feto aumentam o risco de pré-eclâmpsia. Devido a seu efeito negativo no sucesso reprodutivo, os autores previram que existiria uma seleção contra essa combinação, e demonstraram que existe de fato uma correlação negativa de frequências do genótipo AA e HLA-C2 (r = -0,82) em oito populações de etnias diferentes (189). Outro exemplo veio de um estudo realizado em uma pequena vila em Gana, onde a malária é endêmica e tem alta incidência (182). Cada indivíduo exibia uma combinação única de receptores KIR e ligantes HLA, uma evidência de seleção balanceadora possivelmente impulsionada pela pressão conferida pela malária.

Mais recentemente, Vargas e colaboradores demonstraram que populações indígenas da América do Sul exibem, até o momento, os menores números de interações KIR-HLA entre todas as populações mundiais descritas. Ainda, 83-97% de suas interações KIR-HLA dependem de algumas poucas moléculas HLA-C (190) (Figura 11). Em 2009, Gendzekhadze e colaboradores sugeriram que essa limitada diversidade observada em populações nativas americanas é provavelmente a diversidade mínima necessária para sobrevivência humana (191). Além disso, Vargas e colaboradores sugerem que o número limitado de interações KIR-HLA, representadas principalmente por poucos ligantes HLA-C, impõe uma intensa pressão estabilizadora específica da população na região centromérica de KIR para manter os sinais inibidores mínimos necessários para a educação de NK e a consequente sobrevivência dessas populações(190).

FIGURA 12. Distribuição de ligantes de KIR em diferentes populações.



Os ligantes HLA-C representam 85-97% do total de ligantes HLA nos ameríndios. As barras verticais representam a média de interações funcionais KIR-HLA por indivíduo em populações mundiais. O código de cores representa a proporção dessas interações que são representadas por diferentes subgrupos HLA. GRC = Guarani de Rio das Cobras (190); Yucpa = Ameríndios da Venezuela (191); ACHE = Ameríndios do Paraguai (190); GKW = Guarani Kaiowa, Brasil (190); GND = Guarani Ñandeva, Brasil (190); KRC = Kaingang de Rio das Cobras, Brasil (190); KIV = Kaingang de Ivaí, Brasil (190); JAP = população descendente de Japoneses do Brasil (190); CTBA = população do sul do Brasil da cidade de Curitiba (190); INDIGO = população Euro descendente do Estados Unidos (109). Fonte: (Vargas et al, 2021) (190)

2.11 INTERAÇÕES ENTRE *KIR* E *HLA* E O DESENVOLVIMENTO DE TUMORES

Estudos *in vitro* com células humanas, assim como estudos *in vivo* com camundongos e ratos, já demonstraram que células neoplásicas são alvos de células NK (67). Dewan e colaboradores mostraram que quando células tumorais da mama foram transplantadas em camundongos sem imunidade adaptativa, formaram-se tumores não invasivos. Porém, quando esses tumores foram inseridos em camundongos sem imunidade adaptativa e sem atividade de células NK, esses tumores passaram por um rápido processo de metástase (192). Isso mostra a importância das células NK no desenvolvimento de tumores, e como elas estão associadas à defesa contra estes tumores (193). Portanto, o progresso do tumor está particularmente ligado à eficiência do sistema imune, especialmente das células *natural killer* (18).

Diversos estudos mostraram associações entre o polimorfismo de ausência e presença de *KIR* e o desenvolvimento de tumores (Tabela 2). Goedete e colaboradores mostraram que a combinação de KIR3DS1 e Bw4 80I confere proteção ao sarcoma de Kaposi (OR = 0,6, p = 0,01) (194). Middleton e colaboradores mostraram que a presença de *KIR3DL1* e *KIR2DS4* estariam associadas a um risco maior ao desenvolvimento de câncer de bexiga (OR=risco, *p*=0,01; OR=risco, *p*=0,01) (195). Kim e colaboradores encontraram uma associação entre a presença de *KIR2DS5* e o desenvolvimento de câncer colorretal em asiáticos, em que a presença do receptor caracterizaria um risco aumentado ao desenvolvimento da doença (OR=1,9 e *p*=0,007) (196). Além desses estudos citados, muitos outros mostraram associações ou falta de associações entre o polimorfismo de KIR e o desenvolvimento de câncer (6).

Com relação ao câncer de mama, poucos estudos foram realizados para tentar identificar a associação entre genes *KIR* e seus ligantes HLA e a suscetibilidade a essa doença (7–11) (Tabela 3). O primeiro estudo utilizou apenas 33 amostras de pacientes com tumor mamário e 77 controles, apresentando fracas associações e resultados não conclusivos. Nesse trabalho, a presença de *KIR2DS1* foi associada a um maior risco (*p*= 0,03), ao passo que

TABELA 2. Associações de *KIR* com diversos tipos de câncer.

		Ethnicity	OR/effect	*P*	*N*	References
Leukemia	2DL2	Euro	Risk	0.007	94	Verheyden et al., 2004
	2DS2		Risk	0.022	94	
	A/A genotype		Protection	0.011	94	
	2DS4	Asian	1.76	0.008	263	Zhang et al., 2010
	2DS4 in CML		3.29	< 0.001	135	
	2DL2	Euro	0.61	0.029	158	Middleton et al., 2009
	2DL2 in CML		0.39	0.004	52	
	Bw4 Ile80		1.72	0.018	158	
	Bw4 Ile80 in AML		3.32	< 0.001	54	
	2DL2+C1		0.55	0.009	158	
	2DL2+C1 in CML		0.28	< 0.001	52	
	2DS2+C1		0.58	0.018	158	
	2DS2+C1 in CML		0.33	0.002	52	
	2DS1	Euro	0.55	0.020	100	Almalte et al., 2011
	2DS2		0.19	< 0.001	100	
	2DS3		0.32	< 0.001	100	
	2DS4		0.49	0.004	100	
	2DS5		0.32	< 0.001	100	
	3DS1		0.27	< 0.001	100	
	>4 Activating KIR		0.06	< 0.001	100	
	Lack of association with KIR	Euro	NA		220	Babor et al., 2012
	A/A	Hispanic	1.86	0.03	114	de Smith et al., 2014
	A/A	Euro	NA	0.37	76	
	Bw4/Bw4	Euro	3.93	0.01	76	
	Lack of association for KIR2DL1/S1 alleles	Euro	NA		320	Babor et al., 2014
	C1/C1 in ALL		0.69	0.005	320	
	Lack of association with KIR in B-CLL	Euro	NA		197	Karabon et al., 2011
	Bw4		0.56	0.005	197	
Lymphoma	3DS1 (familial study) in HL	Euro	0.44	0.006	345*	Besson et al., 2007
	2DL5 (familial study) in HL		0.56	0.02	345*	
	2DS1 (familial study) in HL		0.42	0.01	345*	
	2DS4full (familial study) HL		2.22	0.03	345*	
	Lack of association with KIR (case-control)		NA		68	
	Lack of association with KIR in HL	Arab	NA		41	Hoteit et al., 2015
	Lack of association with KIR in FL	Arab	NA		20	Khalaf et al., 2013
	Lack of association with KIR in DLBCL	Asian	NA		60	Vejbaesya et al., 2014
	Bw4 in DLBCL		0.39	0.003	60	
	3DL1+Bw4 in DLBCL		0.34	0.001	60	
Multiple myeloma	2DS4*001/002	Arab	Risk	0.04	34	Hoteit et al., 2014
	2DS5		Risk	0.007	34	
Nasopharyngeal	>5 Activating KIR	Asian	3.40	0.07	378	Butsch Kovacic et al., 2005
Breast cancer	2DL1	Euro	Risk	0.03	34	Ozturk et al., 2012
	2DS1		Risk	0.03	34	
	2DS4*003/4/6/7		Protection	0.03	34	
	2DL2	Euro	2.18	< 0.001	230	Jobim et al., 2013
	C1		2.71	< 0.001	230	
	2DL2+C1 in absence of 2DS2		9.95	< 0.001	230	
Colorectal cancer	Lack of associations	Euro	NA		128	Al Omar et al., 2010
	2DS5	Asian	1.9	0.007	241	Kim et al., 2014
	Lack of association with KIR	Euro	NA		75	Al Omar et al., 2010
	Bw4 Ile80 in 3DL1+Bw4+ individuals		3.10	< 0.001	75	
	Bw4 Thr80 in 3DL1+Bw4+ individuals		0.30	< 0.001	75	
	Lack of associations	Euro	NA		90	Middleton et al., 2007
Melanoma	Lack of association with KIR	Euro	NA		50	Naumova et al., 2005
	C2		0.27	0.017	50	
Ovarian cancer	Lack of association with KIR	Euro	NA		142	Giebel et al., 2014
	C1		3.07	0.002	103	
Kidney cancer	2DL3+C1	Euro	5.90	0.009	40	Al Omar et al., 2010
	2DL2+C1-		0.08	0.002	40	
Kaposi's sarcoma	2DS1	Euro	3.82	0.008	32	Guerini et al., 2012
	3DS1		4.00	0.006	32	
	2DS1+C2		4.24	0.01	32	
	KIR3DS1+Bw4 Ile80	Euro	0.60	0.01	250	Goedert et al., 2016

Fonte: (Augusto, 2016) (6)

TABELA 3. Associações de *KIR* e seus ligantes com câncer de mama.

KIR/HLA	**Estudo**	**OR / efeito**	**Valor de *p***	**Referência**
KIR2DL1	BC x controles	Proteção	0,02	Ozturk *et al* (7)
KIR2DS1	BC x controles	Risco	0,03	Ozturk *et al* (7)
KIR2DL2	BC x controles	2,18	<0,001	Jobim *et al* (8)
KIR2DL2 + C1	BC x controles	2,32	<0,001	Jobim *et al* (8)
KIR2DS2	BC x controles	0,25	<0,01	Alomar *et al* (9)
KIR2DS3	BC x controles	0,21	<0,01	Alomar *et al* (9)
KIR2DL5A	BC x controles	0,27	<0,01	Alomar *et al* (9)
KIR2DL2 sem C1	BC x controles	0,03	0,001	Alomar *et al* (9)
KIR2DL3 sem C1	BC x controles	0,03	0,001	Alomar *et al* (9)
KIR2DL1	BC x controles	0,193	0,04	Larki *et al* (10)
KIR2DS4 del	BC x controles	0,39	0,0001	Larki *et al* (10)
KIR2DS4	BC x controles	1,8	0,0041	Larki *et al* (10)
KIR2DS1	ER x controles	1,89	0,036	Larki *et al* (10)
KIR3DS1	ER x controles	1,55	0,046	Larki *et al* (10)
KIR2DL5	PR x controles	1,64	0,036	Larki *et al* (10)
KIR2DL1	HER2+ x controles	0,15	0,038	Larki *et al* (10)
KIR2DL5	TN x outros grupos	0,14	0,015	Larki *et al* (10)
KIR2DL3	BC x controles	0,29	0,000003	Ashouri *et al* (11)
KIR3DL1	BC x controles	0,41	0,02	Ashouri *et al* (11)
KIR2DS4	BC x controles	0,44	0,04	Ashouri *et al* (11)
KIR2DL2	BC x controles	1,56	0,035	Ashouri *et al* (11)
KIR2DL5	BC x controles	1,54	0,039	Ashouri *et al* (11)
KIR3DS1	BC x controles	1,66	0,01	Ashouri *et al* (11)
KIR2DS1	BC x controles	2,19	0,0001	Ashouri *et al* (11)
KIR2DS5	BC x controles	2,22	0,0001	Ashouri *et al* (11)

BC = Câncer de mama; ER = receptor de estrogênio presente; PR = receptor de progesterona presente; HER2+ = presença de HER2 enriquecido; TN = câncer de mama triplo negativo; *KIR2DS4 del* = Variante de KIR2DS4 com a deleção de 22pb. Fonte: O autor, 2022.

a presença de *KIR2DL1* foi associado com proteção (*p*= 0,02) (7). Jobim *et al* utilizaram 230 amostras de pacientes e 278 controles, e encontraram uma forte associação entre a presença simultânea de *KIR2DL2* e HLA-C1 com maior risco de desenvolvimento de câncer mamário (OR: 2,32; p = 0,001) (8). O último estudo realizado, também com um número amostral pequeno, utilizou 50 amostras de pacientes sauditas e 65 controles. Os resultados deste estudo sugeriram que a presença dos genes *KIR2DS2*, *KIR2DS3* e *KIR2DL5A* protegem contra a doença (OR = 0,25, 0,21 e 0,27 respectivamente; $p < 0,01$), ao passo que, a presença de HLA-C1 e C2 em heterozigose foi associada a um risco aumentado (OR = 2,33; $p = 0,03$). Por fim a presença de *KIR2DL2* e *KIR2DL3* na ausência do ligante HLA-C1 mostrou-se protetora contra a doença (OR = 0,03; $p = 0,001$) (9). Mais recentemente, um novo estudo analisou o polimorfismo de *KIR* em subtipos agressivos de câncer de mama (10). Com o avanço de técnicas de sequenciamento, a variação alélica de *KIR* passou a ser o próximo objetivo de análise em estudos de caso controle, em detrimento das análises de presença e ausência realizadas no passado. Com isso, ainda não existe nenhum trabalho que analisou o polimorfismo alélico de *KIR* em tumores mamários.

Os estudos apresentados anteriormente chegaram a resultados não conclusivos quanto à relação dos genes *KIR* e *HLA* com o câncer mamário, primariamente consequência de limitações do número amostral e amostras não bem caracterizadas. Além disso, até o momento nenhum estudo avaliou o polimorfismo alélico de *KIR* em carcinomas mamários e seus subtipos. O papel do polimorfismo de *KIR-HLA* no carcinoma mamário continua pouco compreendido. Neste trabalho, utilizamos uma amostra de 550 pacientes e 747 controles, sendo as pacientes majoritariamente bem caracterizadas quanto aos dados clínicos e histopatológicos. Desta forma, pretendemos contribuir de forma inédita para uma melhor compreensão sobre a associação dos polimorfismos de *KIR* e seus ligantes à suscetibilidade e aos parâmetros de progressão e agressividade do câncer de mama.

3. JUSTIFICATIVA

Apesar da importância de KIR no contexto de células NK e doenças, poucos estudos conseguiram explorar o polimorfismo de *KIR* em profundidade devido às limitações metodológicas impostas pela complexa e incomum variação dessa família gênica. Até o momento, a maioria dos estudos populacionais e de associação com doenças ainda se limitam a análise do polimorfismo de presença e ausência de *KIR* baseada em métodos obsoletos e que não são capazes de identificar alelos. Neste trabalho, nós focamos em estabelecer uma referência no estudo de KIR, explorando a diversidade de *KIR* em alta resolução e atingido um nível de análise nunca antes atingido em nível populacional. No primeiro capítulo dessa tese, realizamos o primeiro estudo populacional em larga escala em uma população humana, que analisou profundamente a variação do número de cópias e a variação alélica de alta resolução de todos os genes *KIR* em mais de dois mil indivíduos da população dos Estados Unidos. Estudos populacionais são de extrema importância para entendermos melhor a complexidade dessa família gênica. Em conjunto com um grande número amostral, nós focamos na identificação de haplótipos incomuns que normalmente não podem ser identificados em amostras pequenas. Ainda, essa abordagem possibilita a caracterização de variantes com impacto funcional.

No contexto do papel das células NK na vigilância de tumores, o estudo de receptores KIR é especialmente relevante, pois essas moléculas participam diretamente no controle da citotoxicidade, mediando o ataque e lise da célula alvo. Combinações KIR-HLA já foram associados a diversos tipos de cânceres e a muitas outras doenças. Porém, os estudos entre genes *KIR* e o carcinoma mamário esporádico são escassos, realizados com pequeno número amostral e se limitando a estudar a variação de presença e ausência. Não existem, até o momento, trabalhos com populações bem caracterizadas e com o tamanho amostral proposto nesse estudo. Portanto, ainda não há resultados conclusivos sobre a importância da variação de *KIR-HLA* nessa doença. Desse modo, buscamos uma contribuição inédita e significativa para compreender o impacto da variação de KIR-HLA no câncer de mama. compreensão dos mecanismos genéticos envolvidos na etiologia do câncer de mama esporádico Estudos como

esse podem servir de base para trabalhos futuros que buscam marcadores genéticos que podem ser aplicados em imunoterapia.

4. OBJETIVOS

4.1 OBJETIVOS GERAIS

- Caracterizar o polimorfismo alélico em alta resolução de uma população Euro descendente dos Estados Unidos;
- Verificar se polimorfismos de *KIR* e *HLA* estão associados à suscetibilidade ao câncer de mama esporádico e ao agravamento da doença.

4.2 OBJETIVOS ESPECÍFICOS

- Descrever a diversidade de *KIR* em alta resolução em uma amostra populacional de Euro-descendente dos Estados Unidos;
- Definir o número de cópias de *KIR* e determinar seus haplótipos de conteúdo gênico;
- Identificar a fase de haplótipos de conteúdo gênico e de alelos
- Identificar novas variantes alélicas e estruturais nos genes *KIR*
- Verificar se o polimorfismo de ausência e presença de *KIR* altera o risco de desenvolver câncer de mama em uma população brasileira com ascendência europeia;
- Verificar se o variação do número de copias de *KIR* está relacionado com a suscetibilidade ao câncer de mama esporádico;
- Verificar se o polimorfismo alélico de *KIR* está relacionado com a suscetibilidade ao câncer de mama esporádico
- Verificar se combinações KIR-HLA estão relacionados com a suscetibilidade ao câncer de mama esporádico
- Verificar se o polimorfismo de *KIR* e *HLA* estão relacionados aos subtipos imunoistoquímicos, grau histológico e a ocorrência de metástase;

5. MATERIAL E MÉTODOS

5.1 AMOSTRAS BIOLÓGICAS

5.1.1 Amostras de euro descendentes dos Estados Unidos

Neste trabalho, analisamos uma amostra de 2.130 indivíduos adultos não aparentados previamente descritos por Hollenbach e colaboradores em 2019 (197). Todos os indivíduos se auto identificaram como descendentes de Europeus e eram residentes nos Estados Unidos.

5.1.2 Amostras de pacientes e controles de câncer de mama

Foi realizada a análise de amostras de DNA extraídas de 550 pacientes de câncer de mama, extraídas pelo Laboratório de Citogenética Humana e Oncogenética (LABCHO), e provenientes do Hospital das Clínicas (HC-PR) e Hospital Nossa Senhora das Graças (HNSG). Parte das amostras está classificada de acordo com o tipo histológico, grau de agressividade do câncer e subtipo molecular, além de informações sobre presença de metástases em linfonodos (quadro 2). As pacientes, majoritariamente provenientes do Estado do Paraná, são de ancestralidade predominantemente europeia e tem uma média de idade de 56 anos. Como controles, foram utilizadas 747 amostras de indivíduos do sexo feminino, todas acima de 40 anos, que se autodeclararam brancas. Essas amostras foram provenientes dos biorrepositórios do Laboratório de Genética Molecular Humana (LGMH) e do Laboratório de Imunogenética e Histocompatibilidade (LIGH), ambos da Universidade Federal do Paraná (UFPR). Esse projeto foi aprovado pelo comitê de ética da Universidade Estadual de Santa Cruz, Bahia protocolo CAAE 67400917.3.0000.5526, parecer nº 2.175.650.

QUADRO 4. CARACTERIZAÇÃO DAS AMOSTRAS DE CÂNCER DE MAMA

Casos de câncer de mama (n = 550)

Grau de agressividade	*f*	n
Grau I	0,118	44
Grau II	0,589	218
Grau III	0,293	107
NA	-	180
Total	1	370

Subtipo imunoistoquímico	*f*	n
Luminal	0,815	317
HER2 positivo	0,062	24
Triplo Negativo	0,123	48
NA	-	161
Total	1	389

Metástase em linfonodos	*f*	n
Positivo	0,441	195
Negativo	0,559	247
NA	-	108
Total	1	442

Os valores de frequência representam as porcentagens em relação ao total de amostras que foram classificadas, não levando em consideração as amostras sem informação. *f* = frequência, NA = amostras sem informação. Fonte: o autor (2022).

5.3 EXTRAÇÃO DE DNA DAS AMOSTRAS DE PACIENTES DE CÂNCER DE MAMA

A extração de DNA de amostras de tumor foi feita seguindo o protocolo padrão utilizado no Laboratório de Citogenética Humana e Oncogenética. No primeiro dia de extração as amostras de tumor (estocadas em solução RNA later a -80ºC, após processamento manual para retirada de tecido gorduroso e vasos sanguíneos) foram descongeladas, e uma pequena porção foi retirada e adicionada a um tubo de 2ml. Então adicionaram-se 120µl de tampão e 80µl de proteinase K, tendo sido as amostras posteriormente incubadas a 56ºC por cerca de 18 horas. No dia seguinte começaram as etapas de extração. Adicionou-se 1 volume de fenol (aproximadamente 200µl) à reação, que posteriormente foi misturada vigorosamente por 15 segundos. Após, as amostras foram centrifugadas a 23447g por 5 minutos. O sobrenadante foi retirado e colocado em um novo tubo de 2ml. As etapas de extração foram repetidas, e após a centrifugação e a passagem do sobrenadante para um novo tubo, adicionou-se 40µl de acetato de amônio 10M, e 400µl de etanol absoluto gelado para precipitação do DNA. Os tubos foram alocados no freezer até o dia seguinte. No terceiro dia foram realizadas as etapas de lavagem do DNA. As amostras do dia anterior foram centrifugadas a 23447g por 15 minutos. Subsequentemente o conteúdo do tubo foi descartado e adicionou-se 1ml de etanol 70% gelado. Após homogeneização da amostra, efetuou-se novamente a centrifugação a 23447g por 15 minutos. Em seguida descartou-se o conteúdo do tubo e aguardou-se o etanol presente evaporar. Por fim, o DNA que se encontrava no fundo do tubo foi resuspendido em água. O protocolo de extração da maior parte das amostras foi realizado ao longo de anos por alunos do laboratório de Citogenética Humana e Oncogenética, seguindo a rotina do próprio laboratório.

5.4 SEQUENCIAMENTO DE *KIR* E *HLA*

Um total de 100 ng de DNA de alta qualidade é fragmentado usando o kit KAPA HyperPlus (Roche), incubando por 5 minutos a 37°C. Subsequentemente, o DNA fragmentado tem suas extremidades reparadas, a cauda poli-A adicionada, e são ligadas através de PCR a adaptadores de índice duplo compatíveis com a tecnologia da Illumina, com código de barras exclusivo. Após

a ligação, os fragmentos são purificados com esferas magnéticas HyperCap de razão de 0,8X (Roche) seguidas de seleção de tamanho duplo (proporções de 0,42X e 0,15X) para selecionar bibliotecas de aproximadamente 800 pares de base. Finalmente, as bibliotecas são amplificadas e purificadas com esferas magnéticas. Após a quantificação por PCR quantitativo, 30ng de cada amostra são agrupadas usando energia acústica ultrassônica, e a captura direcionada de enriquecimento é realizada com kits de hibridização da Twist Bioscience. Resumidamente, as bibliotecas estão ligadas a 10.456 sondas biotiniladas descritas anteriormente (135). Ao usar esferas magnéticas de estreptavidina, os fragmentos direcionados pelas sondas são capturados e depois amplificados e purificados. As bibliotecas enriquecidas são analisadas no BioAnalyzer (Agilent) e quantificadas por PCR digital (ddPCR). Finalmente, as bibliotecas enriquecidas são sequenciadas usando o aparelho NovaSeq6000 (Illumina) com protocolo de sequenciamento de extremidade pareada de 250 pares de base. Após o sequenciamento, os arquivos *fastq* brutos são analisados usando nosso *pipeline* de bioinformática personalizado (151) para obter o número de cópias *KIR* e dados alélicos de alta resolução, e o software HLA Explore (Omixon) e AlloSeq Assign (CareDx) para obter genótipos *HLA*.

5.5 ANÁLISE DE DADOS DA POPULACAO EURO DESCENDENTE

5.5.1 Identificação de haplótipos de conteúdo gênico de *KIR*

Os haplótipos de conteúdo gênico de *KIR* foram identificados manualmente, com base na determinação precisa do número de cópias, no desequilíbrio de ligação conhecido entre os genes *KIR* e na genotipagem alélica. Os haplótipos incomuns foram identificados com base em observações anteriores (126,127). Os candidatos a novos haplótipos de conteúdo gênico foram identificados quando pareados com haplótipos comuns e observados em dois ou mais indivíduos. Após identificar os haplótipos de conteúdo gênico da população norte americana, inferimos os haplótipos de seus alelos *KIR* usando o algoritmo de maximização de expectativa (EM) e o pacote do R “haplo.stats” (http://CRAN.R-project.org/package=haplo.stats).

5.5.2 Análise de desequilíbrio de ligação

Para a análise de desequilíbrio de ligação (DL) feita no estudo da população euro descendente, usamos os dados de genotipagem alélica e os transformamos em um arquivo de entrada Arlequin (.arp) usando o GenAlEx 6.5 (198). A estimativa de fase gamética e a análise de desequilíbrio de ligação pareada foram realizadas usando o software Arlequin 3.5.2.2 (199). Os dados de DL gerados neste trabalho (109) serviram de base para a resolução de ambiguidades na genotipagem alélica de *KIR* nos dados de câncer de mama e controles.

5.5.3 Identificação de novos alelos

Procuramos na amostra populacional Euro descendente variantes de nucleotídeo único (SNV) presentes em *KIR2DL1* e *KIR3DL1S1,* que foram identificadas por nosso software, mas não estavam presentes em nenhum alelo listado no Immuno Polymorphism Database (IPD)-KIR versão 2.9.0 (4). Para indivíduos portadores de um SNV candidato em *KIR2DL1* ou *KIR3DL1S1*, estes foram resequenciados usando o método de Sanger (200) usando primers previamente descritos (177,201).

5.6 ANÁLISE DE DADOS DAS AMOSTRAS DE CANCER DE MAMA E CONTROLES

Haplótipos de conteúdo gênico foram identificados manualmente, com base na determinação precisa do número de cópias, o conhecido desequilíbrio de ligação entre os genes *KIR* e informações alélicas, conforme descrito anteriormente (109).

HLA-A, HLA-B e HLA-C foram classificados de acordo com seus epítopos relevantes conhecidos para o reconhecimento de KIR. Foram considerados os seguintes pares KIR-HLA: HLA-A*11 e KIR2DS2 (165); HLA-A*03/HLA-A*11 e KIR3DL2 (167); HLA-C2 e KIR2DL1 (163); HLA-C1 e KIR2DL2 (163); HLA-C1 e KIR2DL3 (163); HLA-Bw4 e KIR3DL1 (164,169); HLA-C2 e KIR2DS1 (175).

A análise de associação genética foi realizada usando o pacote BIGDAWG (202), que lida com dados multialélicos para testar a associação no haplótipo, locus e alelo. As associações de presença e ausência e números de

cópias de *KIR* foram testadas com o software de estatística IBM SPSS (203), usando o método Mantel-Haenszel (204). O limiar de significância foi estabelecido em $p^{corr} < 0,05$ após correção para múltiplos testes pelo método de Bonferroni (205).

6. RESULTADOS

A seção resultados será apresentada em três capítulos. Primeiramente, apresentaremos um manuscrito publicado na *Frontiers in Immunology* com a descrição detalhada do polimorfismo de *KIR* em alta resolução de uma população euro descendente dos Estados Unidos. No capítulo dois apresentamos um manuscrito com a análise do polimorfismo de *KIR* e seus ligantes HLA em uma amostra bem caracterizada de pacientes com câncer de mama esporádico e seus subtipos. Por fim, no capítulo 3 apresentamos os resultados da análise de polimorfismos dos genes *HLA* de classe I e classe II no carcinoma mamário esporádico e seus subtipos, em formato de *Brief Communication*.

frontiers in Immunology

ORIGINAL RESEARCH
published: 07 May 2021
doi: 10.3389/fimmu.2021.674778

High-Resolution Characterization of *KIR* Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity

Leonardo M. Amorim[1†], Danillo G. Augusto[2†], Neda Nemat-Gorgani[3], Gonzalo Montero-Martin[4], Wesley M. Marin[2], Hengameh Shams[2], Ravi Dandekar[2], Stacy Caillier[2], Peter Parham[3], Marcelo A. Fernández-Viña[4], Jorge R. Oksenberg[2], Paul J. Norman[3,5] and Jill A. Hollenbach[2]*

[1] Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil, [2] Department of Neurology, University of California, San Francisco, CA, United States, [3] Department of Structural Biology, Stanford University, Palo Alto, CA, United States, [4] Histocompatibility & Immunogenetics Laboratory, Stanford Blood Center, Palo Alto, CA, United States, [5] Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Denver, CO, United States

OPEN ACCESS

Edited by:
Jeane E. L. Visentainer, State University of Maringá, Brazil

Reviewed by:
Christiane Ayo, Faculty of Medicine of São José do Rio Preto, Brazil
Luís Cristóvão Porto, Rio de Janeiro State University, Brazil

***Correspondence:**
Jill A. Hollenbach
jill.hollenbach@ucsf.edu

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *01 March 2021*
Accepted: *20 April 2021*
Published: *07 May 2021*

Citation:
Amorim LM, Augusto DG, Nemat-Gorgani N, Montero-Martin G, Marin WM, Shams H, Dandekar R, Caillier S, Parham P, Fernández-Viña MA, Oksenberg JR, Norman PJ and Hollenbach JA (2021) High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity. Front. Immunol. 12:674778. doi: 10.3389/fimmu.2021.674778

The *KIR* (*killer-cell immunoglobulin-like receptor*) region is characterized by structural variation and high sequence similarity among genes, imposing technical difficulties for analysis. We undertook the most comprehensive study to date of *KIR* genetic diversity in a large population sample, applying next-generation sequencing in 2,130 United States European-descendant individuals. Data were analyzed using our custom bioinformatics pipeline specifically designed to address technical obstacles in determining *KIR* genotypes. Precise gene copy number determination allowed us to identify a set of uncommon gene-content *KIR* haplotypes accounting for 5.2% of structural variation. In this cohort, *KIR2DL4* is the framework gene that most varies in copy number (6.5% of all individuals). We identified phased high-resolution alleles in large multi-locus insertions and also likely founder haplotypes from which they were deleted. Additionally, we observed 250 alleles at 5-digit resolution, of which 90 have frequencies ≥1%. We found sequence patterns that were consistent with the presence of novel alleles in 398 (18.7%) individuals and contextualized multiple orphan dbSNPs within the *KIR* complex. We also identified a novel KIR2DL1 variant, Pro151Arg, and demonstrated by molecular dynamics that this substitution is predicted to affect interaction with HLA-C. No previous studies have fully explored the full range of structural and sequence variation of *KIR* as we present here. We demonstrate that pairing high-throughput sequencing with state-of-art computational tools in a large cohort permits exploration of all aspects of *KIR* variation including determination of population-level haplotype diversity, improving understanding of the *KIR* system, and providing an important reference for future studies.

Keywords: NK receptors, alleles, high resolution, NGS, sequencing, population

INTRODUCTION

A variety of membrane-bound receptors control the response of natural killer cells (NK) to infected or malignant cells (1, 2). The killer-cell immunoglobulin-like receptors (KIR) are the most polymorphic family of NK receptors and encoded by a gene family located at chromosomal region 19q13.4 (3, 4). The *KIR* genes exhibit extraordinary variation, both within populations and between them (5–7). Although *KIR* gene-content has been extensively studied in numerous populations worldwide (8–11), less is known about *KIR* allele diversity. The unusual structural variation of the *KIR* region, coupled with the numerous alleles for each *KIR* gene (12) and extensive sequence similarity within the *KIR* gene family are distinguishing characteristics of *KIR* variation. Further characterizing the *KIR* region are frequent duplications, large deletions, hybrid genes and recombinant alleles (13–17). Together, these obstacles have impeded high-resolution allelic characterization of all *KIR* genes in population studies, which has been accomplished in only a few studies (18–21).

Transduction of NK cell activating and inhibitory signals is achieved by a subset of human leukocyte antigen (HLA) class I molecules, which serve as KIR ligands (22–24). These two interacting molecule families evolve as a unique and integrated system (25–28), and combinations of KIR and HLA have been associated with numerous diseases (29–32), including autoimmune disorders (33–36), malignancies (37–39) and infections (40–43). Combinations of KIR and HLA class I also affect placentation and the success of reproduction (44–47). Therefore, high-resolution allelic analysis of *KIR* and *HLA* class I diversity across populations will be necessary to understand their evolution and lay a foundation for functional studies to determine disease mechanisms. To facilitate this progression, we have used our custom *KIR* genotyping and bioinformatics pipeline to interrogate *KIR* diversity in a sample of 2,130 US residents. These methods were explicitly designed to cope with the complexities of *KIR* alleles, gene and haplotypes. We describe the most comprehensive analysis to date of the *KIR* genes, exploring copy number variation, haplotype patterns, and novel variation not previously reported.

MATERIAL AND METHODS

Study Population

We analyzed a cohort of 2,130 unrelated healthy adult individuals previously described by Hollenbach et al., 2019 (48). All individuals self-identified as being of European descent and were resident of the United States.

KIR Genotyping

DNA samples were sequenced for all *KIR* genes, according to Norman and collaborators (49). After sequencing, raw fastq files were analyzed using our custom bioinformatics pipeline PING (Pushing Immunogenetics into the Next Generation) to obtain *KIR* gene content and allelic genotypes from next-generation sequencing (NGS) data (49). We applied an updated version of the pipeline that precisely determines the copy number of each locus through multiple alignment and filtration steps, also accurately identifying *KIR* genotypes. The updated pipeline increased the accuracy of *KIR* genotype determination and is publicly available (50).

Haplotype Estimation

Gene-content haplotypes were identified manually, based on the precise copy number determination, the known linkage disequilibrium among *KIR* genes, and allelic information. Uncommon haplotypes were identified based on previous observations (13, 51, 52). Candidates for novel gene-content haplotypes were identified when paired with common haplotypes and observed in two or more individuals. After identifying gene-content haplotypes, we inferred the haplotypes of their *KIR* alleles using the expectation-maximization (EM) algorithm and the R package haplo.stats (http://CRAN.R-project.org/package=haplo.stats).

Linkage Disequilibrium (LD) Analysis

Allelic genotyping data were transformed into an Arlequin entry file (.arp) using GenAlEx 6.5 (53). Gametic phase estimation using an EM algorithm and further pairwise linkage disequilibrium analysis were performed using Arlequin 3.5.2.2 (54).

Identification of Novel Alleles

We searched for the single nucleotide variants (SNV) in *KIR2DL1* and *KIR3DL1S1* that were identified by our software but not present in any allele listed at the Immuno Polymorphism Database (IPD)-KIR release 2.9.0 (12). For individuals carrying a candidate novel SNV in *KIR2DL1* or *KIR3DL1S1*, the respective genes were re-sequenced using the Sanger method (55) using previously described primers (25, 56).

Simulations of Molecular Dynamics

In silico, the KIR2DL1 chain was isolated from the KIR-HLA complex (PDB ID: 1IM9) (57). To map the *rs200879366* variation on the KIR2DL1 structure, the proline at position 151 was replaced by arginine using the Mutate plugin in Visual Molecular Dynamics (VMD) package (58). Both allotypes were solvated in separate simulation boxes using TIP3P solvent, and the ion concentration was adjusted to 150 KCl. Energy minimization was then performed on both systems for 150,000 steps. To mimic the anchorage of KIR2DL1 to the lipid membrane, the atom of residue 200 was fixed in space. Conformational transitions of KIR2DL1 allotypes were modeled using the NAMD software package (59) and the CHARMM36 forcefield (60) in NPT ensembles. Temperature and pressure were maintained at 310 K and 1 bar using the Langevin thermostat and Langevin piston Nose-Hoover, respectively. Periodic boundary condition in all directions and a timestep of 2fs were used. Simulations on both systems ran for 100ns. The angle between the D1 and D2 domains is the leading indicator of KIR2DL1 conformational transition. It was obtained by aligning the corresponding atom selections and calculating

rotation and displacement at every timeframe using in-house tcl scripts. Structure visualizations were performed using VMD.

RESULTS

KIR Gene Copy Number Analysis Identified Numerous Deletions and Duplications Involving *KIR3DP1*, *KIR2DL4*, and *KIR3DL1S1*

We determined *KIR* gene copy number for all 13 *KIR* genes (*KIR2DL1*, *KIR2DL23*, *KIR2DL4*, *KIR2DL5A*, *KIR2DL5B*, *KIR2DS1*, *KIR2DS2*, *KIR2DS3*, *KIR2DS4*, *KIR2DS5*, *KIR3DL1S1*, *KIR3DL2*, *KIR3DL3*) and the two pseudogenes (*KIR2DP1*, and *KIR3DP1*). Carrier frequencies and gene frequencies, based on the direct counting of all copies for each *KIR*, are given in **Supplementary Table 1**.

The *KIR A* haplotype is defined by the presence of only one gene encoding a short-tailed activating receptor (KIR2DS4) and a fixed number of genes encoding inhibitory KIR. In contrast, *KIR B* haplotypes encode various combinations of activating and inhibiting receptors (61). A total of 31.4% of individuals in the cohort of 2,130 Europeans are homozygous for the full-length *KIR A* haplotype (*cA01~tA01/cA01~tA01*, $f = 0.58$). The centromeric and telomeric portions of the *KIR* haplotype are flanked by the framework genes *KIR3DL3-KIR3DP1* and *KIR2DL4-KIR3DL2*, respectively (62, 63). On analyzing these two regions separately, we observed that 94% of the centromeric diversity is explained by just three gene-content haplotypes (*cA01*, *cB02*, and *cB01*; **Figure 1A**), whereas *tA01* and *tB01* correspond to 93% of the telomeric haplotypes (**Figure 1B**). We also indentified the presence of two novel centromeric haplotypes, and two telomeric haplotypes not previoulsy described in Europeans. The novel *cB06* haplotype differs from *cB01* by lacking *KIR2DP1*, and *cA03* differs from the more common *cA01* by lacking *KIR3DP1*. Present in the telomeric region is *tA02*, which only differs from *tA01* by lacking *KIR2DS4*. Of particular interest is a haplotype observed in two individuals that has only the *KIR3DL2* framework gene in the telomeric region. The gene content and organization of 3.5% of the centromeric and 4.9% of the telomeric haplotypes could not be determined.

Of the framework genes, *KIR2DL4* has the most copy number variation in our study. We observed 72 individuals (3.4%) carrying one copy of *KIR2DL4* and 67 (3.1%) carrying three

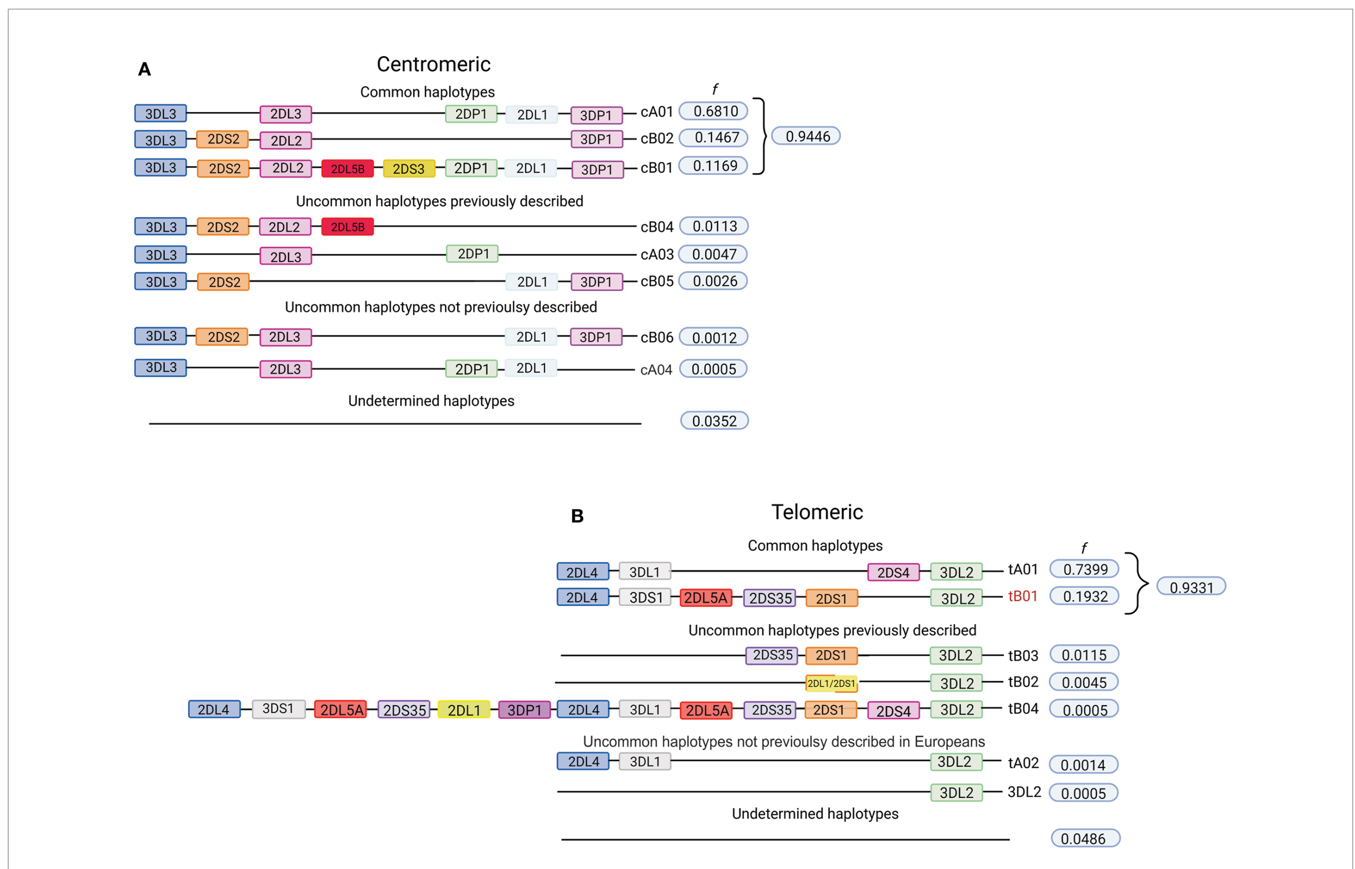


FIGURE 1 | Telomeric and centromeric gene-content haplotypes in European-Americans. Although multiple variations of the *KIR* full-length haplotypes have been described, most are multiple variations of a few centromeric and telomeric haplotypes. The centromeric and telomeric regions of *KIR* haplotypes are flanked by the genes *KIR3DL3-KIR3DP1* and *KIR2DL4-KIR3DL2*, respectively, which are referred to as framework genes (62, 63). **(A)** Frequencies of centromeric *KIR* gene-content haplotypes in the study population. **(B)** Frequencies of telomeric *KIR* gene-content haplotypes in the study population. All listed haplotypes that were not previously described, we have observed in multiple individuals and in combination with a high frequency haplotype, allowing their inference with confidence. All uncommon haplotypes for which the phase could not be determined are grouped as "undetermined". Figure created with BioRender.com.

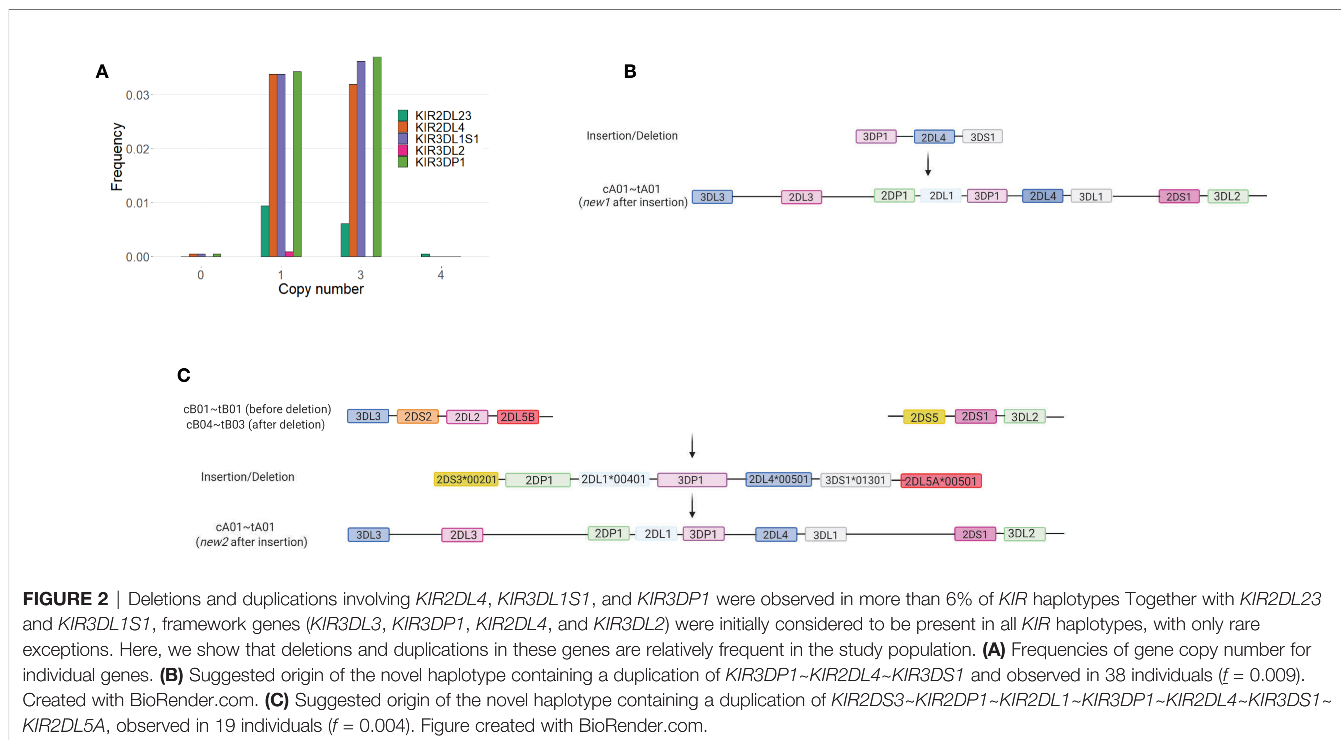


FIGURE 2 | Deletions and duplications involving *KIR2DL4*, *KIR3DL1S1*, and *KIR3DP1* were observed in more than 6% of *KIR* haplotypes Together with *KIR2DL23* and *KIR3DL1S1*, framework genes (*KIR3DL3*, *KIR3DP1*, *KIR2DL4*, and *KIR3DL2*) were initially considered to be present in all *KIR* haplotypes, with only rare exceptions. Here, we show that deletions and duplications in these genes are relatively frequent in the study population. **(A)** Frequencies of gene copy number for individual genes. **(B)** Suggested origin of the novel haplotype containing a duplication of *KIR3DP1~KIR2DL4~KIR3DS1* and observed in 38 individuals (*f* = 0.009). Created with BioRender.com. **(C)** Suggested origin of the novel haplotype containing a duplication of *KIR2DS3~KIR2DP1~KIR2DL1~KIR3DP1~KIR2DL4~KIR3DS1~KIR2DL5A*, observed in 19 individuals (*f* = 0.004). Figure created with BioRender.com.

copies (**Figure 2A**). In this cohort, deletion of *KIR2DL4* is invariably accompanied by the deletion of *KIR3DL1S1*. *KIR3DP1* is deleted in 63 of the 72 (87.5%) haplotypes having the *KIR2DL4~KIR3DL1S1* deletion. *KIR3DL1S1* and *KIR3DP1* are duplicated in 63 out of 68 individuals carrying duplications of *KIR2DL4*. Insertion of *KIR3DP1~KIR2DL4~KIR3DS1* into a *tA01* haplotype created a novel haplotype, carried by 38 individuals (*f* = 0.009; **Figure 2B**). Insertion of the segment *KIR2DS3*00103~KIR2DP1~KIR2DL1*00401~KIR3DP1~KIR2DL4*00501~KIR3DS1*01301~KIR2DL5A*00501* also into *tA01* gave another novel haplotype observed in 19 individuals (*f* = 0.004; **Figure 2C**). In addition, we always observed the centromeric *cB04* haplotype to be in the same gametic phase as *tB03*, and *cA03* always with *tB02*. Duplication of *KIR2DL2* present in *cB02* was observed in 7 individuals (*f* = 0.002)

Several *KIR* Haplotypes Are Marked by Specific *KIR* Alleles

All 13 *KIR* genes were genotyped to five-digit allele resolution in the study sample. We identified 250 *KIR* alleles, of which 90 (37.6%) have frequencies equal or greater than 0.01, and 40 (17%) have frequencies equal or greater than 0.05 (**Figure 3A**). *KIR3DL3* has the highest variety of alleles (n = 83), followed by *KIR3DL2* (n = 48) and *KIR3DL1S1* (n = 39) (**Figure 3B**).

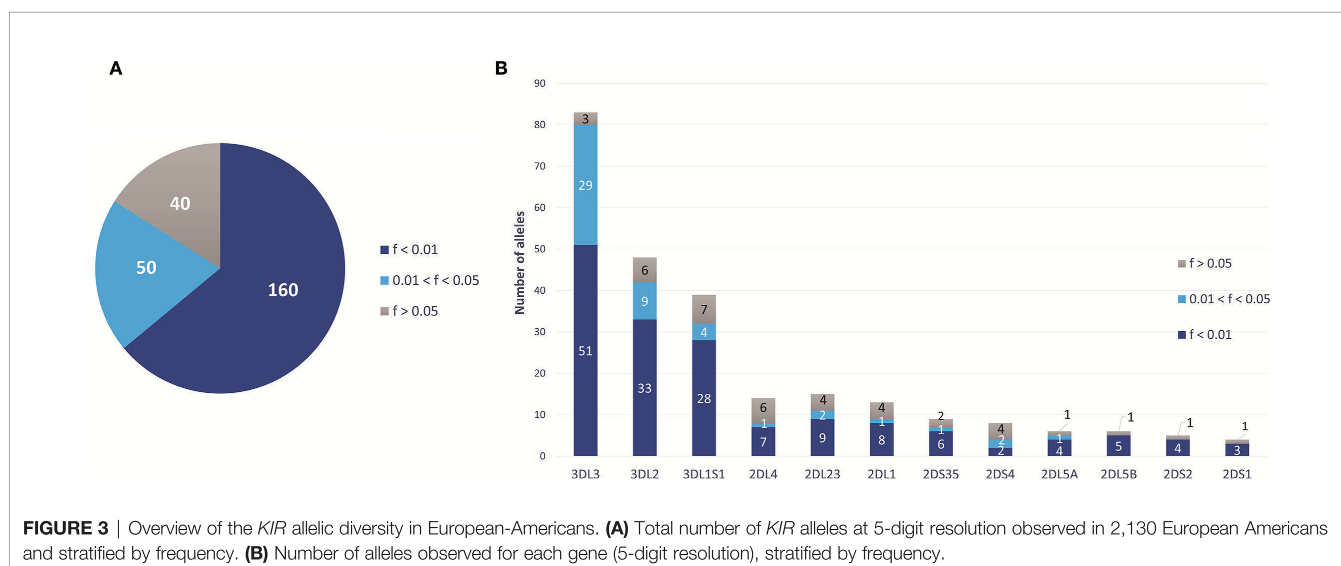


FIGURE 3 | Overview of the *KIR* allelic diversity in European-Americans. **(A)** Total number of *KIR* alleles at 5-digit resolution observed in 2,130 European Americans and stratified by frequency. **(B)** Number of alleles observed for each gene (5-digit resolution), stratified by frequency.

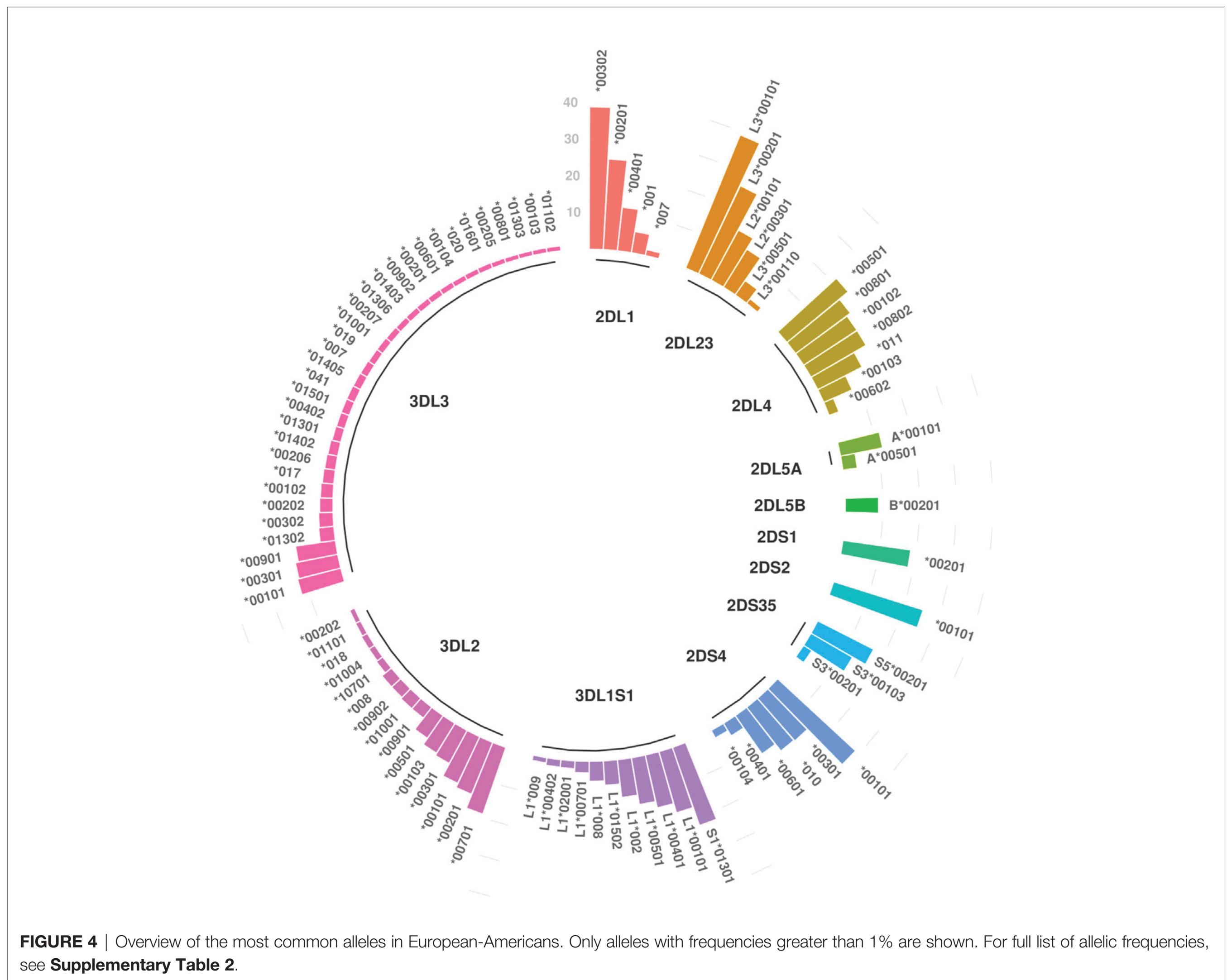


FIGURE 4 | Overview of the most common alleles in European-Americans. Only alleles with frequencies greater than 1% are shown. For full list of allelic frequencies, see **Supplementary Table 2**.

Figure 4 summarizes *KIR* allele diversity of the cohort and complete allele frequencies are given in **Supplementary Table 2**. Among the 20 most common centromeric haplotypes, 16 are *cA01* (**Figure 5A**). Similarly, 17 of the 20 most common telomeric haplotypes were *tA01* (**Figure 5B**).

This is the first study to describe high-resolution (5-digit) *KIR* haplotypes for all functional *KIR* genes in a large population sample. Analyzing this large number of individuals gave us sufficient power to fully explore the patterns of LD and identify alleles that are exclusively or predominately associated with specific haplotypes (**Figure 6**). For example, the alleles *KIR3DL2*00701* and **018* were observed solely in *tB01* haplotypes, whereas *KIR3DL2*00103, *00201, *00501, *00901, *00101*, and **008* were observed only in *tA01*. Similarly, *KIR3DL3*00301* and **00402* are characteristic of *CenB*, whereas **00901* and **00101* are exclusive to *CenA* haplotypes. With few exceptions, *2DL4*00501* is the only *KIR2DL4* allele found in *tB01* haplotypes (99.4%), being in complete LD with *KIR3DS1*013*. Additionally, a few low-frequency alleles are associated with specific, uncommon haplotypes; for example, *KIR3DL2*034* (f = 0.002) is present only in *tB03*, and *KIR2DL5B*00801* (f = 0.004), is present only in *cB04* (f = 0.01). Furthermore, for the first time we describe multiple high-resolution allelic configurations of the full-length *cB04~tB03* haplotype (**Supplementary Table 3**).

As well as the haplotypic associations, we observed many instances of strong LD among specific sets of *KIR* alleles (**Figure 7**). In summary, *KIR2DL4*00501*, *KIR3DS1*01301*, *KIR2DS1*00201*, *KIR2DL5A*00101*, *KIR2DS5*00201*, and *KIR3DL2*00701* are frequently observed together. Many other *KIR2DL4* alleles are in strong LD with specific *KIR3DL1* alleles. Specific examples are *KIR2DL4*00801* with *KIR3DL1*00101* (D' = 0.99, r^2 = 0.85); *KIR2DL4*011* with *KIR3DL1*00501* (D' = 0.95, r^2 = 0.85); *KIR2DL4*00802* with *KIR3DL1*00401* (D'= 0.95, r^2 = 0.73); *KIR2DL4*00602* with *KIR3DL1*00701* (D' = 0.91, r^2 = 0.68); and *KIR2DL4*00103* with *KIR3DL1*008* (D' = 1, r^2 = 0.6). In the centromeric region, *KIR2DL5B*00201* and *KIR2DS3*00103* are always observed together. *KIR2DL1*00401* is in strong LD with *KIR2DS3*00103* (D' = 0.98 and r^2 = 0.83) and *KIR2DL5B*00201* (D' = 0.95 and r^2 = 0.82), whereas

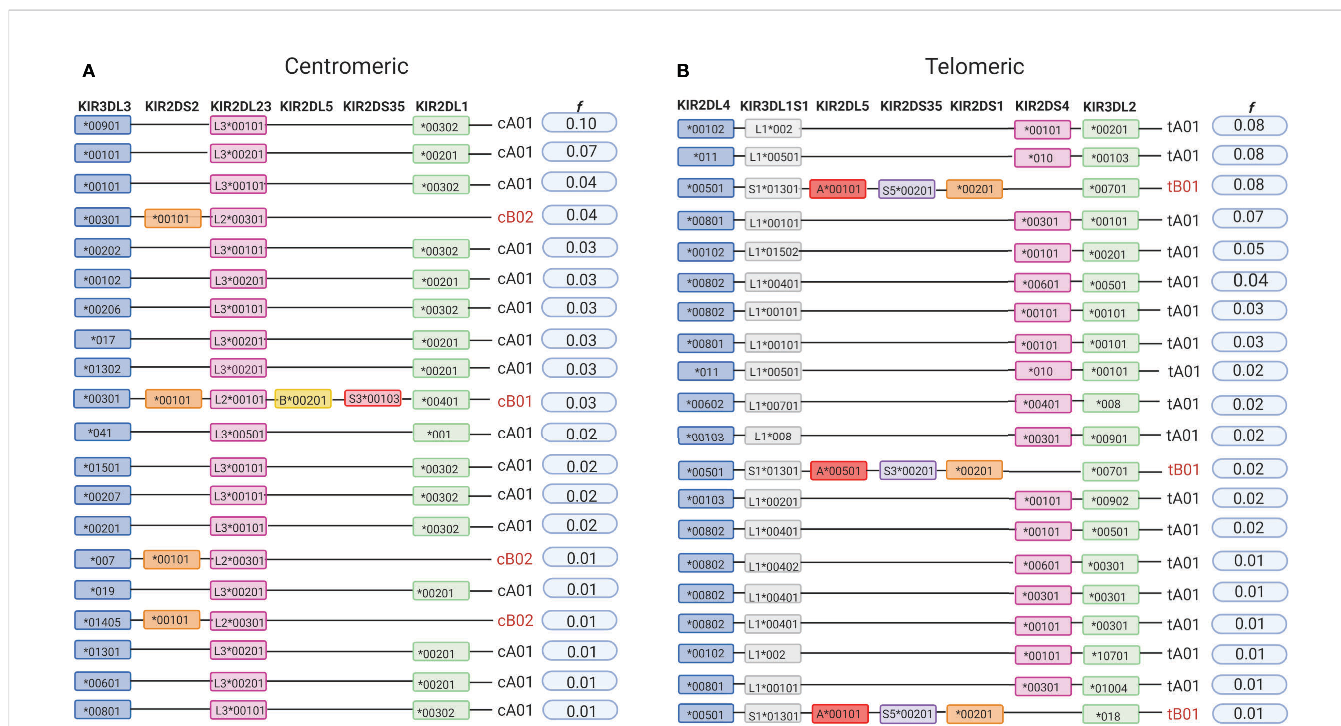


FIGURE 5 | High-resolution allelic haplotypes in European-Americans. **(A)** The 20 most common centromeric *KIR* haplotypes. **(B)** The 20 most common telomeric *KIR* haplotypes observed in European Americans (n = 2,130). Created with BioRender.com.

*KIR2DL3*00201* is associated with *KIR2DL1*00201* (D' = 0.97, r2 = 0.89). A full list of LD values for pairs of *KIR* alleles is given in **Supplementary Table 4**.

Numerous Novel *KIR* Variants Are Present in the Cohort of European Americans

In this analysis of 2,130 individuals, we identified 398 individuals (18.7%) carrying at least one *KIR* recombinant allele that do not correspond to any sequences deposited in the *KIR* database. We define as recombinant allele those that are characterized by different phasing combinations of previously known variable sites. These observations are likely to represent the presence of new alleles that were not present or not detected in previous studies of *KIR* variation. *KIR3DL1S1* accounts for 33% of the observations of candidate new alleles, corresponding to a total allelic frequency of 0.04 at this locus. A large proportion of individuals carrying possible novel alleles were also observed for other KIR2D and KIR3D genes (**Table 1**).

In addition to these candidate novel *KIR* recombinant alleles, we also used our software to identify possible novel SNVs. In some cases, these SNVs may have been reported in the dbSNP database (64), but they were not associated with any *KIR* allele sequence deposited in the IPD-*KIR* database release 2.9.0 (12). Therefore, in the context of *KIR*, these SNVs would be contributing to novel alleles that each differ from a known *KIR* allele by a single nucleotide substitution. To confirm the sequences of novel variants, we used the Sanger method to re-sequence individuals carrying any possible novel SNV in *KIR2DL1* and *KIR3DL1S1*. While confirmation of all novel variants was out of scope for the current project, we selected these loci as exemplars for this work due to substantial previous work examining their structure and function (25, 56, 61, 65–71), including the availability of crystal for molecular modeling structures (57, 72–76). In future work we will continue to explore novel variants that were detected at other loci during this study.

For *KIR2DL1*, 8 of 30 variants were confirmed by Sanger sequencing, while 10 of 32 variants were similarly confirmed for *KIR3DL1S1*. Most of the SNVs in *KIR3DL1S1* were observed in only a single individual, except for two synonymous variants, *rs754894112* and *rs1462310393* (**Table 2**). In contrast, the majority of confirmed novel variants in *KIR2DL1* were observed in several individuals (**Table 3**). Interestingly, 14 out of the 18 confirmed variants in these two genes were non-synonymous substitutions, with functional effects ranging from conservative to radical according to the Grantham scale of physicochemical distances between amino acids (77).

Simulation of Molecular Dynamics Predicts That Dimorphism in Codon 151 of *KIR2DL1* Affects Binding to HLA Class I

To explore the functional differences of *KIR2DL1* alleles that differ by a single nucleotide, we simulated and compared the molecular dynamics simulations of KIR2DL1 allotypes that differ by the non-conservative substitution of proline to arginine at position 151. Underlying this difference is *rs200879366*C>G*. Different conformations were sampled during the simulation trajectory, so that each time step features an individual

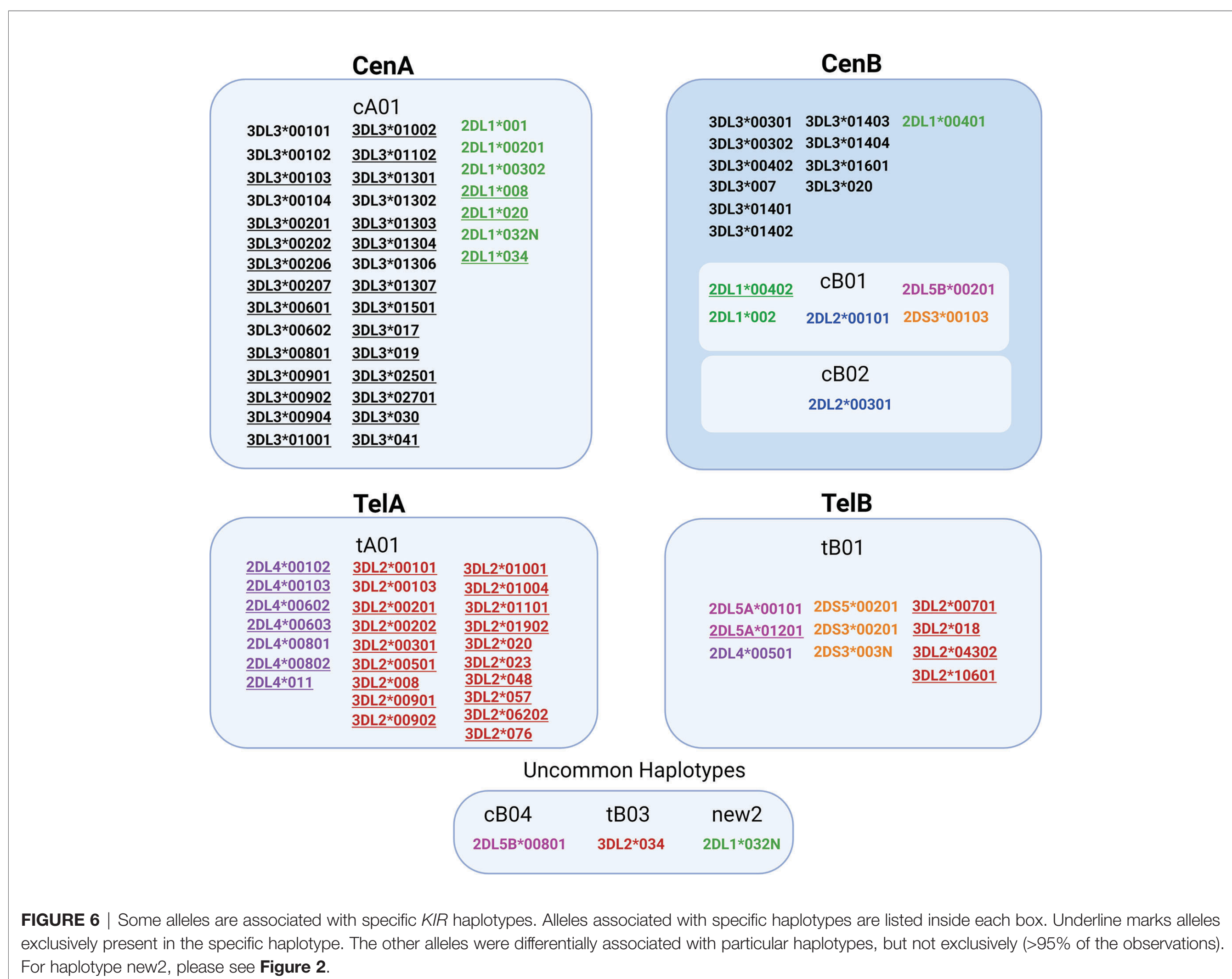


FIGURE 6 | Some alleles are associated with specific *KIR* haplotypes. Alleles associated with specific haplotypes are listed inside each box. Underline marks alleles exclusively present in the specific haplotype. The other alleles were differentially associated with particular haplotypes, but not exclusively (>95% of the observations). For haplotype new2, please see **Figure 2**.

conformer. The angle change mediated the transition between free and HLA-bound states of KIR2DL1 between the Ig domains (D1 and D2) that eventually affected the HLA binding region (**Figure 8A**). Within 100ns of simulation, the angle between the D1 and D2 domains was decreased by 10° in the wildtype (Pro) but was not perturbed in the mutant (Arg) (**Figure 8B**). The conformational transition appears to be mediated by a network of interactions spanning from Met44 to Arg 151. In the wildtype, Met44 is released from Pro185, allowing the angle between D1 and D2 to decrease (**Figure 8C**). By contrast, the angle is increased in the mutant. This enables Arg151 to form a salt bridge with Asp135, leading to an interaction between Met136 and Pro185 (**Figure 8D**) that allows Pro185 to more strongly associate with Met44. This set of localized rearrangements in KIR2DL1 is likely essential for its stable binding to HLA class I.

DISCUSSION

The general configuration of *KIR* gene-content haplotypes was first described two decades ago, when it was observed that four framework genes separate two distinct sub-clusters of genes (7, 24, 62, 78, 79). *KIR3DL3* and *KIR3DP1* were seen to delimit the centromeric region, while *KIR2DL4* and *KIR3DL2* delimit the telomeric region of the *KIR* gene family. The other two genes that are present in most *KIR* haplotypes are *KIR2DL23* and *KIR3DL1S1* (6). Although deletions and duplications of these genes have been previously reported (15–17, 80–82), technical limitations have precluded direct copy number determination of all *KIR* in large-scale population studies. We show here that large structural deletions and duplications involving the framework genes are relatively frequent in European-descendant individuals. For instance, more than 6% of individuals carry a deletion or duplication of *KIR2DL4*. Similar to haplotype variants described in the literature (13, 51, 52), all gene-content *KIR* haplotypes lacking *KIR2DL4* also lacked *KIR3DL1S1* (*tB02* and *tB03*). These haplotypes have been described for other European-Americans (52), while an extensive study of Europeans from Germany did not seek to analyze novel gene-content haplotypes (83). Observation of *KIR2DL4*~*KIR3DL1S1* deletions at high-frequencies in Africans (18, 25) raises the possibility that these variant *KIR* haplotypes originated prior to the modern human

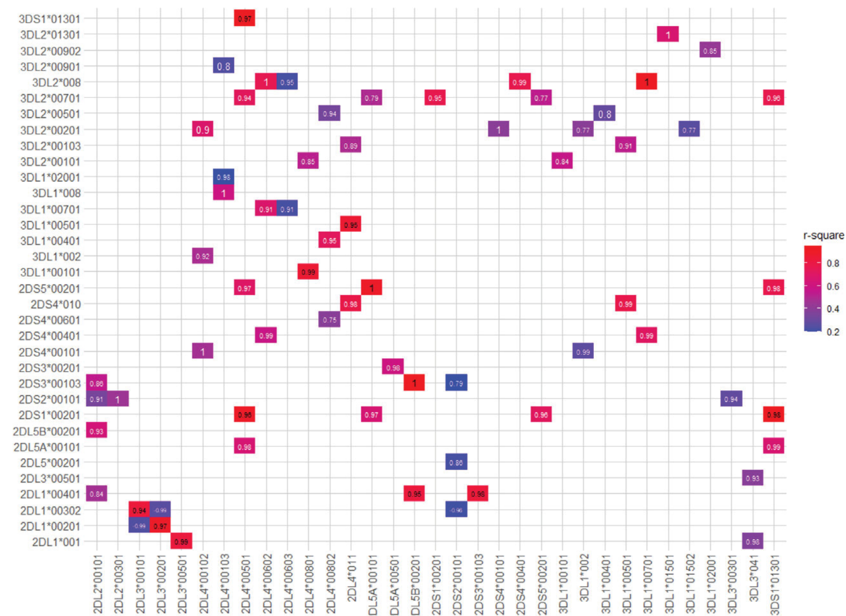


FIGURE 7 | Strong linkage disequilibrium (LD) between *KIR* alleles. Boxes represent the pairs of *KIR* alleles with the strongest LD in the present study. D' values are written inside of each box and color scale represent r^2 values. Only pairs with $r^2 > 0.2$, D' > 0.7 were shown. The *p*-value of all pairs was < 10^{-5}.

migration out of Africa, and that they might also be present in most worldwide populations.

Interestingly, haplotypes carrying *KIR2DL4* duplications also have duplications of *KIR3DL1S1* and *KIR3DP1*. Based on our observations, including the fact that *cB04~tB03* are always in phase, we propose that a single deletion of *KIR2DS3*00103~KIR2DP1~KIR2DL1*00401~KIR3DP1~KIR2DL4*00501~KIR3DS1*01301~KIR2DL5A*00501* from the haplotype *cB01~tB01* originated the *cB04~tB03* haplotype. This seven-locus fragment was possibly inserted into *cA01~tA01*, which would explain the novel full-length haplotype that we identified in multiple individuals. The large cohort that we analyzed allowed us to phase the alleles of the seven-locus indel at high-resolution, providing a unique opportunity to infer the origin of these haplotypic variants.

Previous studies described *KIR* haplotypes at lower genotyping resolution and in smaller sample sizes. For example, Vierra-Green et al. (52) described *KIR* haplotypes at 3-digit resolution in 506 Euro-Americans while Hou et al. (20) analyzed most *KIR* genes at higher resolution but in a small cohort. Here, we present the first study to show 5-digit allelic haplotypes of all *KIR* genes for a large sample of the European-descent U.S population. Notably, the most common centromeric haplotype in our study cohort, *KIR3DL3*00901~ KIR2DL3*00101~KIR2DL1*00302* (*f* = 0.10), is also the most common in four African populations (Datoonga, *f* = 0.11; Baka, *f* = 0.15; Dogon, *f* = 0.18; and Fulani, *f* = 0.14) (18). This haplotype is likely the same reported as the most frequent (*f* = 0.23) in a smaller European American cohort that was not analyzed for all *KIR* genes at high resolution (20). Interestingly, two of the low-frequency telomeric haplotypes present in our sample (*tA02* and *KIR3DL2*) have not been reported in other European populations, but were observed in African populations from Mali (*3DL2*, *f* = 0.01), Democratic Republic of Congo (*tA02*, *f* = 0.08) and Tanzania (*3DL2*, *f* = 0.01; *tA02*, *f* = 0.01) (18). A limitation of determining haplotypes without family segregation studies or confirmation by long-range sequencing is the impossibility to identify unknown haplotype structures or to precisely infer those haplotypes observed in lower frequencies. For this reason, we were not able to confidently identify the less common haplotypes, therefore, presenting data only for the most common ones. However, our large population sample coupled with the curated high-quality data allow us to identify the haplotypic diversity that represent most of the *KIR* diversity in Europeans.

Our well-powered analysis of LD across the *KIR* region shows that some alleles are clearly associated with specific structural haplotypes. Because *KIR2DL4*00501* and *KIR3DL2*00701* are present in *tB01* and are associated with other *tB01*-associated alleles, such as *KIR3DS1*01301*, it was possible to verify that the

TABLE 1 | Large proportion of individuals carrying possible novel alleles.

Locus	n	f
KIR3DL1S1	154	7.23%
KIR3DL2	57	2.68%
KIR2DL5	56	2.63%
KIR2DL4	53	2.49%
KIR3DL3	48	2.25%
KIR2DL1	43	2.02%
KIR2DL23	31	1.46%
KIR2DS3	24	1.13%
KIR2DS5	10	0.47%
KIR2DS4	5	0.23%

n, absolute number individuals carrying potential at least one possible novel allele; f, relative frequency of individuals carrying possible novel alleles.

TABLE 2 | Ten confirmed novel single nucleotide variants in *KIR3DL1S1*.

rsID	n	Exon	Change	Mature protein	Amino acid	Grantham
rs771871523	1	3	G>T	12	Ala>Ser	99
rs1272096635	1	4	C>A	107	Ala>Asp	126
rs769147743	1	4	T>C	127	Ile>Thr	89
rs200893904	1	4	C>T	171	Thr>Ile	89
rs1182774591	1	4	G>C	187	Ala>Pro	27
new SNP	1	5	G>C	249	Arg>Pro	103
rs754894112	12	5	C>T	266	synonymous	
rs1462310393	2	5	C>A	275	synonymous	
new SNP	1	7	C>T	319	His>Tyr	83
rs984592565	1	9	G>A	392	Arg>His	29

These variants have been have been reported and assigned rsID in the dbSNP database (64), but not present in any KIR allele deposited in the IPD-KIR database release 2.9.0 (12), therefore representing novel KIR alleles differing from others by only one nucleotide position. New SNP represent variants not previously assigned rsID in the dbSNP database.

TABLE 3 | Eight confirmed novel single nucleotide variants in *KIR2DL1*.

rsID	n	Exon	Change	Mature protein	Amino acid	Grantham
rs201225013	7	1	A>G	signal peptide	Met>Val	21
rs148427642	4	1	T>G	signal peptide	Cys>Gly	159
rs749640662	13	5	G>A	120	Ala>Thr	58
rs200879366	8	5	C>G	151	Pro>Arg	103
rs749653872	3	5	C>T	157	synonymous	
rs570412759	16	7	G>A	245	Arg>Hist	29
rs201527316	1	9	G>A	296	Arg>Hist	29
rs778821930	2	9	C>A	309	synonymous	

These variants have been have been reported and assigned rsID in the dbSNP database (64), but not present in any KIR allele deposited in the IPD-KIR database release 2.9.0 (12), therefore representing novel KIR alleles differing from others by only one nucleotide position.

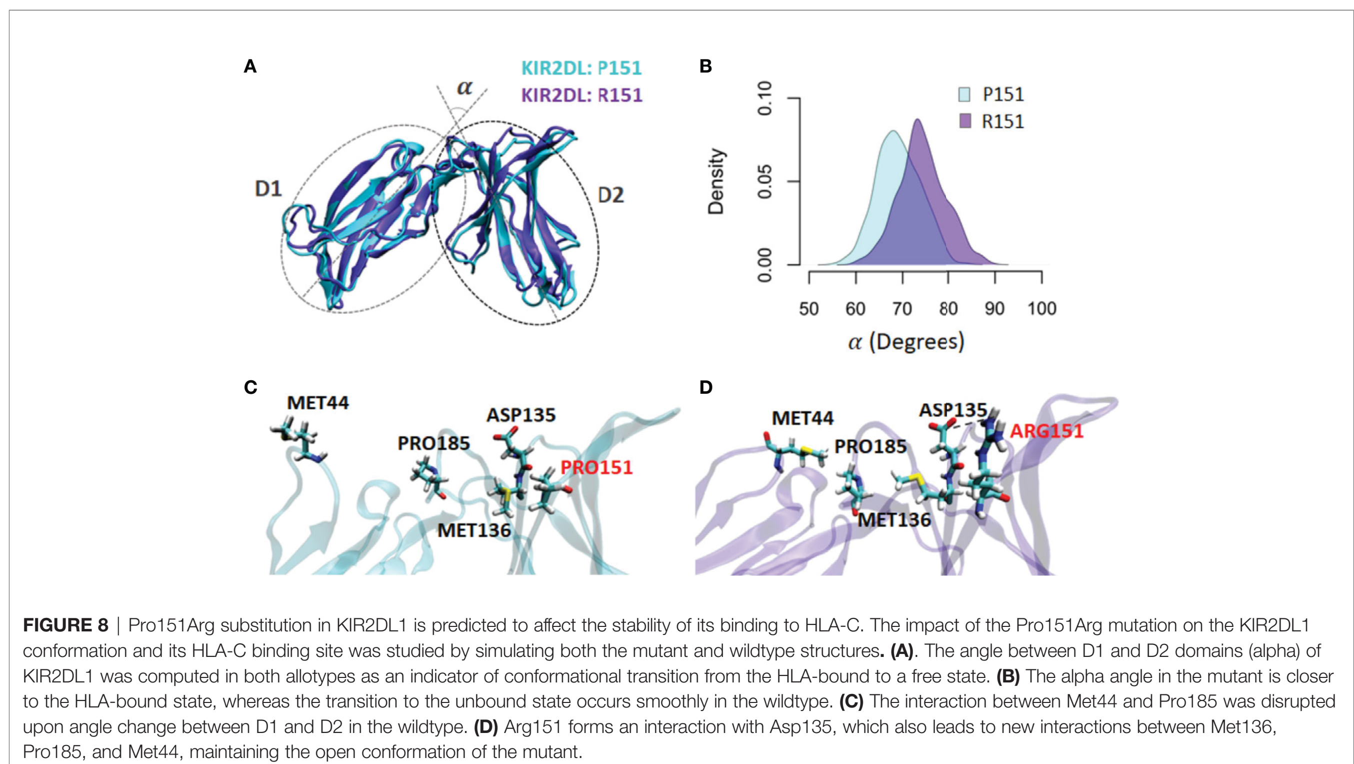


FIGURE 8 | Pro151Arg substitution in KIR2DL1 is predicted to affect the stability of its binding to HLA-C. The impact of the Pro151Arg mutation on the KIR2DL1 conformation and its HLA-C binding site was studied by simulating both the mutant and wildtype structures**. (A)**. The angle between D1 and D2 domains (alpha) of KIR2DL1 was computed in both allotypes as an indicator of conformational transition from the HLA-bound to a free state. **(B)** The alpha angle in the mutant is closer to the HLA-bound state, whereas the transition to the unbound state occurs smoothly in the wildtype. **(C)** The interaction between Met44 and Pro185 was disrupted upon angle change between D1 and D2 in the wildtype. **(D)** Arg151 forms an interaction with Asp135, which also leads to new interactions between Met136, Pro185, and Met44, maintaining the open conformation of the mutant.

insertion of fragments containing *KIR3DS1* occurred on the *cA01~tA01* haplotype. In some cases, specific alleles also associate with uncommon haplotypes, which may be used as markers for these unique haplotypes. These examples highlight how our detailed LD information for high-resolution *KIR* sequencing constitutes a significant resource, yielding valuable information that will facilitate the comprehension and identification of *KIR* haplotypes in future studies.

With the development of robust high-resolution *KIR* typing methods, the discovery of novel *KIR* variants has become more achievable. However, short read misalignment is a major confounding factor in discovering novel variants, particularly those differing from known variants by only one nucleotide. Because the identification of possible novel SNVs is overall not likely in comparison to the possibility of being artifacts due to misalignment of highly similar sequence reads, our pipeline is initially set to exclude them from the primary analysis but flag them for further detailed inspection. This is a limitation intrinsic to short-sequence data analysis in the context of the *KIR* sequence homology, and rare novel variants might be missed. Aiming for an overview of the extent of this methodological limitation, we applied Sanger method to re-sequence all possible novel variants in two highly polymorphic genes, regardless of their frequencies. Although we confirmed approximately one-third of the possible novel SNVs for *KIR3DL1S1* and *KIR2DL1*, most of these variants were observed in only one or a few individuals. In other words, even though our pipeline can possibly miss some of the novel variation, the new SNVs do not represent significant overall distortions in our dataset. In fact, our method has been proven to show an overall high performance for determining accurate genotypes (median 96.5%), and only 1% to 3% of unresolved genotypes (50).

Low frequent variants may, however, be of particular interest especially if they may cause impact on the receptors' function. It is remarkable that most confirmed novel variants cause moderate to radical non-synonymous substitutions, ultimately leading to functional protein variation. To provide new insights to the functional significance of describing novel variants, we focused on variant *rs200879366**G, which was previously reported with a frequency of 0.01 in the Finnish population (84) but had not previously been associated with a specific *KIR* allele. Although residue 151 is in the D2 domain it does not make direct contact with the HLA-C ligand. Nevertheless, polymorphism at the neighboring residue, position 154, has been implicated in differential avidity for the HLA-C ligand (85). Our prediction shows that Asp135, which is directly engaged in HLA binding, forms a bond with Arg151, allowing us to speculate that *rs200879366**G may result in reduced binding to HLA. This example demonstrates the likelihood that many other functional variants will be identified as we interrogate *KIR* allelic diversity in worldwide populations.

According to the recently updated IPD-KIR database (Release 2.10.0, 16 December 2020), the variant *rs200879366*G* marks two unconfirmed alleles, *KIR2DL1*044* and *KIR2DL1*046*, corroborating our findings. These unconfirmed allele sequences were freshly submitted by the same group that genotyped *KIR* in over a million European samples from the DKMS donor registry (86). Although a remarkable effort in its scale and importance to the field, that study targeted specific exons of each *KIR*, resulting in a 3-digit resolution genotyping with substantial ambiguities. In contrast, our study sought to analyze all *KIR* exons and introns of each gene (5-digit resolution). While smaller than the DKMS study by orders of magnitude, our study is nevertheless the largest sample to-date to comprehensively analyze all aspects of *KIR* variation at this resolution, including copy number, allele-haplotype associations, pairwise LD, and functional consequences of novel variation.

For decades, most *KIR* studies in populations were limited to analyzing the presence and absence of genes (5, 7, 9, 62, 87–89). The study of *KIR* gene content laid the basis of the field and suggested that *KIR* diversity and plasticity were and may still be ahead of our technical capabilities. Here, we aimed to set new ground for exploring *KIR* diversity by providing the first large-scale study to deeply analyze copy number variation and high-resolution allelic variation of all genes in a large population sample from the United States. Our results show a large proportion of multi-locus deletions and duplications of genes that were until recently considered rare, in addition to unusual gene-content haplotypes and a high frequency of novel alleles. We argue that as we continue to interrogate *KIR* at high-resolution, we will continue to uncover more layers of this region's complexity, discovering frequent novel variants with functional relevance that have been previously missed due to technical limitations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JH, PN, and DA designed the study. NN-G, GM-M, and DA performed NGS and Sanger sequencing. LA and DA analyzed the data. WM and RD contributed with bioinformatics analysis. HS performed molecular dynamics simulations. SC managed and organized the samples. PP, MF-V, JO, PN, and JH contributed with samples and/or reagents. LA, DA, and JH drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants U19NS095774 and U01 AI090905 from the U.S. National Institutes of Health (NIH). The UCSF DNA biorepository is supported by Si-2001-35701 from the National Multiple Sclerosis Society.

ACKNOWLEDGMENTS

LA acknowledges the scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and the Doctorate Dissertation Research Award scholarship from Fulbright Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.674778/full#supplementary-material

Supplementary Table 1 | *KIR* gene-content frequencies in European-Americans.

Supplementary Table 2 | High-resolution allelic frequencies of the *KIR* genes in European-Americans.

Supplementary Table 3 | Multiple configurations of the haplotype *cB04~tB03* in European-Americans.

Supplementary Table 4 | Pairwise linkage disequilibrium between *KIR* alleles. Only pairs with r2 > 0.2 and D' > 0.7 are shown.

REFERENCES

1. Martinet L, Smyth MJ. Balancing Natural Killer Cell Activation Through Paired Receptors. *Nat Rev Immunol* (2015) 15:243–54. doi: 10.1038/nri3799
2. Lanier LL. NK Cell Receptors. *Annu Rev Immunol* (1998) 16:359–93. doi: 10.1146/annurev.immunol.16.1.359
3. Wende H, Colonna M, Ziegler A, Volz A. Organization of the Leukocyte Receptor Cluster (LRC) on Human Chromosome 19q13.4. *Mamm Genome* (1999) 10:154–60. doi: 10.1007/s003359900961
4. Wilson MJ, Torkar M, Trowsdale J. Genomic Organization of a Human Killer Cell Inhibitory Receptor Gene. *Tissue Antigens* (1997) 49:574–9. doi: 10.1111/j.1399-0039.1997.tb02804.x
5. Norman PJ, Chandanayingyong D, Vaughan RW, Verity DH, Stephens HAF. Distribution of Natural Killer Cell Immunoglobulin-Like Receptor Sequences in Three Ethnic Groups. *Immunogenetics* (2001) 52:195–205. doi: 10.1007/s002510000281
6. Uhrberg M, Valiante NM, Shum BP, Shilling HG, Lienert-Weidenbach K, Corliss B, et al. Human Diversity in Killer Cell Inhibitory Receptor Genes. *Immunity* (1997) 7:753–63. doi: 10.1016/S1074-7613(00)80394-5
7. Hsu KC, Liu X-R, Selvakumar A, Mickelson E, O'Reilly RJ, Dupont B. Killer Ig-Like Receptor Haplotype Analysis by Gene Content: Evidence for Genomic Diversity With a Minimum of Six Basic Framework Haplotypes, Each With Multiple Subsets. *J Immunol* (2002) 169:5118–29. doi: 10.4049/jimmunol.169.9.5118
8. Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, et al. Allele Frequency Net Database (AFND) 2020 Update: Gold-Standard Data Classification, Open Access Genotype Data and New Query Tools. *Nucleic Acids Res* (2020) 48:D783–8. doi: 10.1093/nar/gkz1029
9. Hollenbach JA, Nocedal I, Ladner MB, Single RM, Trachtenberg EA. Killer Cell Immunoglobulin-Like Receptor (KIR) Gene Content Variation in the HGDP-CEPH Populations. *Immunogenetics* (2012) 64:719–37. doi: 10.1007/s00251-012-0629-x
10. Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR, et al. Global Diversity and Evidence for Coevolution of KIR and HLA. *Nat Genet* (2007) 39:1114–9. doi: 10.1038/ng2077
11. Misra MK, Augusto DG, Martin GM, Nemat-Gorgani N, Sauter J, Hofmann JA, et al. Report From the Killer-cell Immunoglobulin-Like Receptors (KIR) Component of the 17th International HLA and Immunogenetics Workshop. *Hum Immunol* (2018) 79:825–33. doi: 10.1016/j.humimm.2018.10.003
12. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. Ipd-IMGT/HLA Database. *Nucleic Acids Res* (2020) 48:D948–55. doi: 10.1093/nar/gkz950
13. Roe D, Vierra-Green C, Pyo C-W, Eng K, Hall R, Kuang R, et al. Revealing Complete Complex KIR Haplotypes Phased by Long-Read Sequencing Technology. *Genes Immun* (2017) 18:127–34. doi: 10.1038/gene.2017.10
14. Traherne JA, Martin M, Ward R, Ohashi M, Pellett F, Gladman D, et al. Mechanisms of Copy Number Variation and Hybrid Gene Formation in the KIR Immune Gene Complex. *Hum Mol Genet* (2010) 19:737–51. doi: 10.1093/hmg/ddp538
15. Martin MP, Bashirova A, Traherne J, Trowsdale J, Carrington M. Cutting Edge: Expansion of the KIR Locus by Unequal Crossing Over. *J Immunol* (2003) 171:2192. doi: 10.4049/jimmunol.171.5.2192
16. Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, et al. Meiotic Recombination Generates Rich Diversity in NK Cell Receptor Genes, Alleles, and Haplotypes. *Genome Res* (2009) 19:757–69. doi: 10.1101/gr.085738.108
17. Pyo C-W, Wang R, Vu Q, Cereb N, Yang SY, Duh F-M, et al. Recombinant Structures Expand and Contract Inter and Intragenic Diversification At the KIR Locus. *BMC Genomics* (2013) 14:89. doi: 10.1186/1471-2164-14-89
18. Nemat-Gorgani N, Guethlein LA, Henn BM, Norberg SJ, Chiaroni J, Sikora M, et al. Diversity of KIR, Hla Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. *JI* (2019) 202:2636–47. doi: 10.4049/jimmunol.1801586
19. Alicata C, Ashouri E, Nemat-Gorgani N, Guethlein LA, Marin WM, Tao S, et al. Kir Variation in Iranians Combines High Haplotype and Allotype Diversity With an Abundance of Functional Inhibitory Receptors. *Front Immunol* (2020) 11:556. doi: 10.3389/fimmu.2020.00556
20. Hou L, Chen M, Ng J, Hurley CK. Conserved KIR Allele-Level Haplotypes are Altered by Microvariation in Individuals With European Ancestry. *Genes Immun* (2012) 13:47–58. doi: 10.1038/gene.2011.52
21. Tao S, Kichula KM, Harrison GF, Farias TDJ, Palmer WH, Leaton LA, et al. The Combinatorial Diversity of KIR and HLA Class I Allotypes in Peninsular Malaysia. *Immunology* (2020) 162:389–404. doi: 10.1111/imm.13289
22. Long EO, Rajagopalan S. HLA Class I Recognition by Killer Cell Ig-like Receptors. *Semin Immunol* (2000) 12:101–8. doi: 10.1006/smim.2000.0212
23. Norman PJ, Parham P. Complex Interactions: The Immunogenetics of Human Leukocyte Antigen and Killer Cell Immunoglobulin-Like Receptors. *Semin Hematol* (2005) 42:65–75. doi: 10.1053/j.seminhematol.2005.01.007
24. Trowsdale J, Barten R, Haude A, Stewart CA, Beck S, Wilson MJ. The Genomic Context of Natural Killer Receptor Extended Gene Families: Genomics of NK Receptor Gene Families. *Immunol Rev* (2001) 181:20–38. doi: 10.1034/j.1600-065X.2001.1810102.x
25. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, et al. Co-Evolution of Human Leukocyte Antigen (Hla) Class I Ligands With Killer-Cell Immunoglobulin-Like Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. *PloS Genet* (2013) 9:e1003938. doi: 10.1371/journal.pgen.1003938
26. Augusto DG, Petzl-Erler ML. KIR and HLA Under Pressure: Evidences of Coevolution Across Worldwide Populations. *Hum Genet* (2015) 134:929–40. doi: 10.1007/s00439-015-1579-9
27. Gendzekhadze K, Norman PJ, Abi-Rached L, Graef T, Moesta AK, Layrisse Z, et al. Co-Evolution of KIR2DL3 With HLA-C in a Human Population Retaining Minimal Essential Diversity of KIR and HLA Class I Ligands. *PNAS* (2009) 106:18692–7. doi: 10.1073/pnas.0906051106
28. Moffett A, Colucci F. Co-Evolution of NK Receptors and HLA Ligands in Humans is Driven by Reproduction. *Immunol Rev* (2015) 267:283–97. doi: 10.1111/imr.12323
29. Kulkarni S, Martin MP, Carrington M. The Yin and Yang of HLA and KIR in Human Disease. *Semin Immunol* (2008) 20:343–52. doi: 10.1016/j.smim.2008.06.003
30. Williams AP, Bateman AR, Khakoo SI. Hanging in the Balance. KIR and Their Role in Disease. *Mol Interv* (2005) 5:226–40. doi: 10.1124/mi.5.4.6
31. Boudreau JE, Hsu KC. Natural Killer Cell Education in Human Health and Disease. *Curr Opin Immunol* (2018) 50:102–11. doi: 10.1016/j.coi.2017.11.003
32. Augusto DG. The Impact of KIR Polymorphism on the Risk of Developing Cancer: Not as Strong as Imagined? *Front Genet* (2016) 7:121. doi: 10.3389/fgene.2016.00121
33. Anderson KM, Augusto DG, Dandekar R, Shams H, Zhao C, Yusufali T, et al. Killer Cell Immunoglobulin-Like Receptor Variants Are Associated With Protection From Symptoms Associated With More Severe Course in Parkinson Disease. *J Immunol* (2020) 205:1323–30. doi: 10.4049/jimmunol.2000144
34. Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF, Petzl-Erler ML. Activating KIR and HLA Bw4 Ligands are Associated to Decreased Susceptibility to Pemphigus Foliaceus, an Autoimmune Blistering Skin Disease. *PloS One* (2012) 7:e39991. doi: 10.1371/journal.pone.0039991
35. Suzuki Y, Hamamoto Y, Ogasawara Y, Ishikawa K, Yoshikawa Y, Sasazuki T, et al. Genetic Polymorphisms of Killer Cell Immunoglobulin-Like Receptors

are Associated With Susceptibility to Psoriasis Vulgaris. *J Invest Dermatol* (2004) 122:1133–6. doi: 10.1111/j.0022-202X.2004.22517.x

36. Hollenbach JA, Pando MJ, Caillier SJ, Gourraud P-A, Oksenberg JR. The Killer Immunoglobulin-Like Receptor KIR3DL1 in Combination With HLA-Bw4 is Protective Against Multiple Sclerosis in African Americans. *Genes Immun* (2016) 17:199–202. doi: 10.1038/gene.2016.5
37. Kim H-J, Choi H-B, Jang J-P, Baek I-C, Choi E-J, Park M, et al. Hla-Cw Polypmorphism and Killer Cell Immunoglobulin-Like Receptor (KIR) Gene Analysis in Korean Colorectal Cancer Patients. *Int J Surg* (2014) 12:815–20. doi: 10.1016/j.ijsu.2014.06.012
38. Middleton D, Vilchez JR, Cabrera T, Meenagh A, Williams F, Halfpenny I, et al. Analysis of KIR Gene Frequencies in HLA Class I Characterised Bladder, Colorectal and Laryngeal Tumours. *Tissue Antigens* (2007) 69:220–6. doi: 10.1111/j.1399-0039.2006.00792.x
39. Al Omar S, Middleton D, Marshall E, Porter D, Xinarianos G, Raji O, et al. Associations Between Genes for Killer Immunoglobulin-Like Receptors and Their Ligands in Patients With Solid Tumors. *Hum Immunol* (2010) 71:976–81. doi: 10.1016/j.humimm.2010.06.019
40. Auer ED, Tong HV, Amorim LM, Malheiros D, Hoan NX, Issler HC, et al. Natural Killer Cell Receptor Variants and Chronic Hepatitis B Virus Infection in the Vietnamese Population. *Int J Infect Dis* (2020) 96:541–7. doi: 10.1016/j.ijid.2020.05.033
41. Martin MP, Qi Y, Gao X, Yamada E, Martin JN, Pereyra F, et al. Innate Partnership of HLA-B and KIR3DL1 Subtypes Against HIV-1. *Nat Genet* (2007) 39:733–40. doi: 10.1038/ng2035
42. Podhorzer A, Dirchwolf M, Machicote A, Belen S, Montal S, Paz S, et al. The Clinical Features of Patients With Chronic Hepatitis C Virus Infections are Associated With Killer Cell Immunoglobulin-Like Receptor Genes and Their Expression on the Surface of Natural Killer Cells. *Front Immunol* (2017) 8:1912. doi: 10.3389/fimmu.2017.01912
43. Alves HV, de Moraes AG, Pepineli AC, Tiyo BT, de Lima Neto QA, Santos T da S, et al. The Impact of KIR/HLA Genes on the Risk of Developing Multibacillary Leprosy. *PloS Negl Trop Dis* (2019) 13:e0007696. doi: 10.1371/journal.pntd.0007696
44. Trowsdale J, Moffett A. NK Receptor Interactions With MHC Class I Molecules in Pregnancy. *Semin Immunol* (2008) 20:317–20. doi: 10.1016/j.smim.2008.06.002
45. Moffett A, Chazara O, Colucci F, Johnson MH. Variation of Maternal KIR and Fetal HLA-C Genes in Reproductive Failure: Too Early for Clinical Intervention. *Reprod BioMed Online* (2016) 33:763–9. doi: 10.1016/j.rbmo.2016.08.019
46. Colucci F. The Role of KIR and HLA Interactions in Pregnancy Complications. *Immunogenetics* (2017) 69:557–65. doi: 10.1007/s00251-017-1003-9
47. Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, et al. Maternal Activating KIRs Protect Against Human Reproductive Failure Mediated by Fetal HLA-C2. *J Clin Invest* (2010) 120:4102–10. doi: 10.1172/JCI43998
48. Hollenbach JA, Norman PJ, Creary LE, Damotte V, Montero-Martin G, Caillier S, et al. A Specific Amino Acid Motif of HLA-DRB1 Mediates Risk and Interacts With Smoking History in Parkinson's Disease. *Proc Natl Acad Sci USA* (2019) 116:7419–24. doi: 10.1073/pnas.1821778116
49. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, et al. Defining KIR and HLA Class I Genotypes At Highest Resolution Via High-Throughput Sequencing. *Am J Hum Genet* (2016) 99:375–91. doi: 10.1016/j.ajhg.2016.06.023
50. Marin W, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, et al. High-Throughput Interpretation of Killer-cell Immunoglobulin-Like Receptor Short-Read Sequencing Data With PING. *bioRxiv [Preprint]* (2021). doi: 10.1101/2021.03.24.436770
51. Pyo C-W, Guethlein LA, Vu Q, Wang R, Abi-Rached L, Norman PJ, et al. Different Patterns of Evolution in the Centromeric and Telomeric Regions of Group A and B Haplotypes of the Human Killer Cell Ig-Like Receptor Locus. *PloS One* (2010) 5:e15115. doi: 10.1371/journal.pone.0015115
52. Vierra-Green C, Roe D, Hou L, Hurley CK, Rajalingam R, Reed E, et al. Allele-Level Haplotype Frequencies and Pairwise Linkage Disequilibrium for 14 KIR Loci in 506 European-American Individuals. *PloS One* (2012) 7:e47491. doi: 10.1371/journal.pone.0047491
53. Peakall R, Smouse PE. GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—an Update. *Bioinformatics* (2012) 28:2537–9. doi: 10.1093/bioinformatics/bts460
54. Excoffier L, Lischer HEL. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows. *Mol Ecol Resour* (2010) 10:564–7. doi: 10.1111/j.1755-0998.2010.02847.x
55. Sanger F, Nicklen S, Coulson AR. DNA Sequencing With Chain-Terminating Inhibitors. *PNAS* (1977) 74:5463–7. doi: 10.1073/pnas.74.12.5463
56. Vargas L de B, Dourado RM, Amorim LM, Ho B, Calonga-Solís V, Issler HC, et al. Single Nucleotide Polymorphism in KIR2DL1 is Associated With HLA-C Expression in Global Populations. *Front Immunol* (2020) 11:1881. doi: 10.3389/fimmu.2020.01881
57. Fan QR, Long EO, Wiley DC. Crystal Structure of the Human Natural Killer Cell Inhibitory Receptor KIR2DL1-HLA-Cw4 Complex. *Nat Immunol* (2001) 2:452–60. doi: 10.1038/87766
58. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* (2006) 22:2695–6. doi: 10.1093/bioinformatics/btl461
59. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, et al. Scalable Molecular Dynamics on CPU and GPU Architectures With NAMD. *J Chem Phys* (2020) 153:044130. doi: 10.1063/5.0014475
60. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* (1983) 79:926–35. doi: 10.1063/1.445869
61. Pugh J, Nemat-Gorgani N, Djaoud Z, Guethlein LA, Norman PJ, Parham P. In Vitro Education of Human Natural Killer Cells by KIR3DL1. *Life Sci Alliance* (2019) 2:e201900434. doi: 10.26508/lsa.201900434
62. Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, et al. Plasticity in the Organization and Sequences of Human KIR/ILT Gene Families. *Proc Natl Acad Sci USA* (2000) 97:4778. doi: 10.1073/pnas.080588597
63. Vilches C, Parham P. K IR : Diverse, Rapidly Evolving Receptors of Innate and Adaptive Immunity. *Annu Rev Immunol* (2002) 20:217–51. doi: 10.1146/annurev.immunol.20.092501.134942
64. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Res* (2001) 29:308–11. doi: 10.1093/nar/29.1.308
65. Stewart CA, Laugier-Anfossi F, Vély F, Saulquin X, Riedmuller J, Tisserant A, et al. Recognition of Peptide–MHC Class I Complexes by Activating Killer Immunoglobulin-Like Receptors. *PNAS* (2005) 102:13224–9. doi: 10.1073/pnas.0503594102
66. Hilton HG, Guethlein LA, Goyos A, Nemat-Gorgani N, Bushnell DA, Norman PJ, et al. Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes. *J Immunol* (2015) 195:3160–70. doi: 10.4049/jimmunol.1501358
67. Bari R, Bell T, Leung W-H, Vong QP, Chan WK, Das Gupta N, et al. Significant Functional Heterogeneity Among KIR2DL1 Alleles and a Pivotal Role of Arginine245. *Blood* (2009) 114:5182–90. doi: 10.1182/blood-2009-07-231977
68. Béziat V, Traherne JA, Liu LL, Jayaraman J, Enqvist M, Larsson S, et al. Influence of KIR Gene Copy Number on Natural Killer Cell Education. *Blood* (2013) 121:4703–7. doi: 10.1182/blood-2012-10-461442
69. Le Luduec J-B, Boudreau JE, Freiberg JC, Hsu KC. Novel Approach to Cell Surface Discrimination Between KIR2DL1 Subtypes and KIR2DS1 Identifies Hierarchies in NK Repertoire, Education, and Tolerance. *Front Immunol* (2019) 10:734. doi: 10.3389/fimmu.2019.00734
70. Norman PJ, Abi-Rached L, Gendzekhadze K, Korbel D, Gleimer M, Rowley D, et al. Unusual Selection on the KIR3DL1/S1 Natural Killer Cell Receptor in Africans. *Nat Genet* (2007) 39:1092–9. doi: 10.1038/ng2111
71. Augusto DG, Norman PJ, Dandekar R, Hollenbach JA. Fluctuating and Geographically Specific Selection Characterize Rapid Evolution of the Human Kir Region. *Front Immunol* (2019) 10:989. doi: 10.3389/fimmu.2019.00989
72. Vivian JP, Duncan RC, Berry R, O'Connor GM, Reid HH, Beddoe T, et al. Killer Cell Immunoglobulin-Like Receptor 3DL1-Mediated Recognition of Human Leukocyte Antigen B. *Nature* (2011) 479:401–5. doi: 10.1038/nature10517
73. Pymm P, Illing PT, Ramarathinam SH, O'Connor GM, Hughes VA, Hitchen C, et al. MHC-I Peptides Get Out of the Groove and Enable a Novel Mechanism of HIV-1 Escape. *Nat Struct Mol Biol* (2017) 24:387–94. doi: 10.1038/nsmb.3381

74. Saunders PM, MacLachlan BJ, Pymm P, Illing PT, Deng Y, Wong SC, et al. The Molecular Basis of How Buried Human Leukocyte Antigen Polymorphism Modulates Natural Killer Cell Function. *PNAS* (2020) 117:11636–47. doi: 10.1073/pnas.1920570117
75. Fan QR, Mosyak L, Winter CC, Wagtmann N, Long EO, Wiley DC. Structure of the Inhibitory Receptor for Human Natural Killer Cells Resembles Haematopoietic Receptors. *Nature* (1997) 389:96–100. doi: 10.1038/38028
76. Stewart-Jones GBE, di Gleria K, Kollnberger S, McMichael AJ, Jones EY, Bowness P. Crystal Structures and KIR3DL1 Recognition of Three Immunodominant Viral Peptides Complexed to HLA-B*2705. *Eur J Immunol* (2005) 35:341–51. doi: 10.1002/eji.200425724
77. Grantham R. Amino Acid Difference Formula to Help Explain Protein Evolution. *Science* (1974) 185:862–4. doi: 10.1126/science.185.4154.862
78. Vilches C, Rajalingam R, Uhrberg M, Gardiner CM, Young NT, Parham P. KIR2DL5, a Novel Killer-Cell Receptor With a D0-D2 Configuration of Ig-Like Domains. *J Immunol* (2000) 164:5797–804. doi: 10.4049/jimmunol.164.11.5797
79. Barten R, Torkar M, Haude A, Trowsdale J, Wilson MJ. Divergent and Convergent Evolution of NK-cell Receptors. *Trends Immunol* (2001) 22:52–7. doi: 10.1016/S1471-4906(00)01802-0
80. Gómez-Lozano N, Estefanía E, Williams F, Halfpenny I, Middleton D, Solís R, et al. The Silent KIR3DP1 Gene (CD158c) is Transcribed and Might Encode a Secreted Receptor in a Minority of Humans, in Whom the KIR3DP1, KIR2DL4 and KIR3DL1/KIR3DS1 Genes are Duplicated. *Eur J Immunol* (2005) 35:16–24. doi: 10.1002/eji.200425493
81. Norman PJ, Carrington CVF, Byng M, Maxwell LD, Curran MD, Stephens HAF, et al. Natural Killer Cell Immunoglobulin-Like Receptor (KIR) Locus Profiles in African and South Asian Populations. *Genes Immun* (2002) 3:86–95. doi: 10.1038/sj.gene.6363836
82. Williams F, Maxwell LD, Halfpenny IA, Meenagh A, Sleator C, Curran MD, et al. Multiple Copies of KIR 3DL/S1 and KIR 2DL4 Genes Identified in a Number of Individuals. *Hum Immunol* (2003) 64:729–32. doi: 10.1016/s0198-8859(03)00089-2
83. Solloch UV, Schefzyk D, Schäfer G, Massalski C, Kohler M, Pruschke J, et al. Estimation of German Kir Allele Group Haplotype Frequencies. *Front Immunol* (2020) 11:429. doi: 10.3389/fimmu.2020.00429
84. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A Global Reference for Human Genetic Variation. *Nature* (2015) 526:68–74. doi: 10.1038/nature15393
85. Hilton HG, Vago L, Older Aguilar A, Moesta AK, Graef T, Abi Rached L, et al. Mutation At Positively Selected Positions in the Binding Site for HLA-C Shows KIR2DL1 is a More Refined But Less Adaptable NK Cell Receptor Than KIR2DL3. *J Immunol* (2012) 189:1418–30. doi: 10.4049/jimmunol.1100431
86. Wagner I, Schefzyk D, Pruschke J, Schöfl G, Schöne B, Gruber N, et al. Allele-Level KIR Genotyping of More Than a Million Samples: Workflow, Algorithm, and Observations. *Front Immunol* (2018) 9:2843. doi: 10.3389/fimmu.2018.02843
87. Crum KA, Logue SE, Curran MD, Middleton D. Development of a PCR-SSOP Approach Capable of Defining the Natural Killer Cell Inhibitory Receptor (KIR) Gene Sequence Repertoires. *Tissue Antigens* (2000) 56:313–26. doi: 10.1034/j.1399-0039.2000.560403.x
88. Hollenbach JA, Augusto DG, Alaez C, Bubnova L, Fae I, Fischer G, et al. 16 (Th) IHIW: Population Global Distribution of Killer Immunoglobulin-Like Receptor (KIR) and Ligands. *Int J Immunogenet* (2013) 40:39–45. doi: 10.1111/iji.12028
89. Augusto DG, Piovezan BZ, Tsuneto LT, Callegari-Jacques SM, Petzl-Erler ML. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. *PloS One* (2013) 8:e56755. doi: 10.1371/journal.pone.0056755

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Amorim, Augusto, Nemat-Gorgani, Montero-Martin, Marin, Shams, Dandekar, Caillier, Parham, Fernández-Viña, Oksenberg, Norman and Hollenbach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

6.2 CAPÍTULO II: Natural Killer Cell Receptor Variation is Associated with More Aggressive Subtypes of Breast Cancer.

Natural killer cell receptor variation is associated with more aggressive subtypes of breast cancer

Leonardo M. Amorim[1], Enilze M. S. F. Ribeiro[1], Márcia H. Beltrame, Jill A. Hollenbach[2], Danillo G. Augusto[1,2,3*]
[1]Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil.
[2]Department of Neurology, University of California, San Francisco, CA, USA
[3] Department of Biological Sciences, University of North Carolina at Charlotte
Corresponding Author:
Danillo G. Augusto, PhD
University of North Carolina at Charlotte
Department of Biological Sciences
danillo.augusto@uncc.edu

Abstract
Breast cancer (BC) is the most common malignancy among women with tumors varying in aggressiveness, histological type, molecular subtype, among others. The natural killer cells (NK) are lymphocytes that act both in the innate and adaptive immune response recognizing and killing neoplastic cells. Among others, NK cells express the killer-cell immunoglobulin-like receptor (KIR) molecules on the cell surface, which modulate NK cell responses through the recognition of human leukocyte antigen (HLA) molecules. Despite the well-stablished role of KIR-HLA in diseases, the variation of these genes in high resolution has never been analysed in cancer. Here, we analysed a cohort of 550 patients with BC and 747 controls. We found that the presence of three copies of *KIR2DL5* was associated with protection (OR = 0.21, p^{corr} = 0.027), while the presence of three copies of *KIR3DL1S1* was associated with increased risk to BC (OR=2.44, p^{corr} =0.009). We also identified that the pair KIR2DL3*001+HLA-C1 ligand was significantly reduced in patients with the aggressive non-luminal HER2+ (ER and PR negative) compared with all other subtypes (OR = 0.23, p^{corr} = 0.016), in addition to a strong association of KIR2DS1*002+HLA-C2 in patients with triple negative tumors compared to all other subtypes (OR = 3.34, $p^{corr} < 1.4 \times 10^{-4}$). Finally, we observed a strong protective effect of the pair KIR2DL1*001+HLA-C2, significantly increased in patients with less severe BC grade I compared with the most severe grades II+III (OR = 0.25, p^{corr} = 0.038). In summary, we have provided solid evidence that the variation of *KIR* genes is implicated in BC.
Keywords: NK receptors, Breast Cancer, HER2, Triple Negative, KIR, HLA.

Introduction

Breast cancer (BC) is most common type of malignancy across the globe (1,2) with 2.2 million new cases diagnosed in 2020 and represents the leading cause of death by malignant disease among women worldwide (2,3). Mutations in genes

such as *BRCA1*, *BRCA2*, *CHEK2*, *PTEN,* and *TP53*, among others, have been identified as causing hereditary BC (4). However, only 5 to 10% of breast cancers are caused by germline mutations transmitted to offspring. Instead, sporadic BC is the most common form of the disease and results from the accumulation of acquired and uncorrected genetic alterations in somatic genes. Therefore, sporadic BC is a multifactorial disease in which multiple genes and environmental factors contribute to its etiology (5,6). Identifying its genetic associations can aid in predicting disease risk at the individual and population levels and provide information about underlying disease mechanisms.

Breast tumors are classified into grades I, II, and III by comparing the microscopic similarity of tumor cells to normal cells of breast tissue. Grade I is the least aggressive while grade III is the most severe (7–9). In the past, the choice of treatment was essentially based on the tumor's grade and stage. However, much has been studied to improve the classification of breast cancer using specific markers to account to the complexity and heterogeneity of this disease (10).

Immunohistochemical markers such as estrogen (ER) and progesterone (PR) receptors, HER2 (human epidermal growth factor receptor 2) overexpression, and KI67 (a nuclear protein that is associated with cellular proliferation), are used to classify BC into subtypes that are biologically distinct to guide treatment decisions. Luminal subtypes represents 50-70% of tumors, thus being the most common subtypes, and are characterized by the presence of estrogen (ER) and/or progesterone (PR) receptors, HER2 negative (Luminal A or B) or positive (only luminal B), and low (Luminal A) or high (Luminal B) expression of KI67 (11,12). Luminal A tumors are typically grade I, with has a better prognosis, and are typically treated with hormone therapy; chemotherapy is restricted to only a few cases. Luminal B have worse prognosis compared to Luminal A, and are treated with hormone therapy and chemotherapy (12,13).

HER2 (also known as ErbB2 or HER2/neu) is a transmembrane receptor with tyrosine kinase activity, capable of activating several pro-survivals intracellular signaling pathways (14). The HER2 enriched (non-luminal) subtype is characterized by an overexpression of HER2, high expression of KI67, and

absence of progesterone and estrogen receptors (12). HER2+ represents 15-20% of BC cases and are typically grade III, associated with a more aggressive disease. Patients with HER2+ tumor present high recurrence and lower survival rates (15).

Triple-negative BC is an aggressive type of tumor that is usually more difficult to treat, characterized by a faster growth rate, higher risk of metastasis and recurrence. Normal breast cells typically have receptors for estrogen and progesterone, which allows them to grow in response to these hormones. However, triple-negative breast cancer cells do not have those receptors or the overexpression of the HER2, which preclude the use of hormonal treatment or HER2-targeted therapeutic monoclonal antibodies (10).

In the context of cancer immunology, natural killer (NK) cells are in the front line of defense against neoplastic cells. NK cells are lymphocytes that act both in innate and adaptive immune responses (16) killing malignant and infected cells, also playing an essential role in reproduction (17). Their cytotoxicity against the target cells is controlled by the balance of signals from their activating and inhibitory surface receptors (18). Among these receptors are the killer cell immunoglobulin-like receptor (KIR) family, which recognizes human leukocyte antigen (HLA) class I molecules. HLA class I molecules are expressed on the surface of all nucleated cells (19) and encoded by genes within the major histocompatibility complex (MHC) located at the human chromosome 6. The abnormal expression commonly caused by malignancy and infection is one of the signals that lead to the extermination by NK cells (20,21).

KIR genes, on the other hand, are located at the chromosome 19 and exhibit an unusual polymorphism of presence or absence of genes, in addition to large structural deletions and insertions and an extreme sequence variability (22–26). Coupled with the homology and high sequency identity among genes, these characteristics have systematically precluded the analysis of *KIR* in high resolution in populations or disease cohorts.

Previous studies analyzing only the presence and absence variation suggested that *KIR* polymorphism might be involved in sporadic BC susceptibility, outcomes, and treatment (27–30). However, these past studies were limited by low sample sizes and reduced genotyping resolution, which reduce the power to detect susceptibility markers. Here, we fill this gap by applying next-generation sequencing technology in a large BC cohort to perform the first high-resolution analysis of *KIR* in cancer, and *KIR* alleles associated with the aggressive HER2+ and triple negative BC.

Material and Methods

Study Cohort

We analyzed genomic DNA from peripheral blood cells from 550 women with sporadic BC and 747 unrelated healthy women. Samples were collected by oncologists specialized in treating BC from Hospital Nossa Senhora das Graças, Curitiba, Brazil, who diagnosed and classified the patients according to clinical guidelines jointly established by the American Society of Clinical Oncology and the College of American Pathology. The diagnosis was combined with comprehensive genetics counseling, including family medical history, medical records, and genetic tests to identify patients with sporadic breast cancer. Volunteers older than 18 years of age diagnosed with sporadic BC were included. The median age in patients is 56 years old. As control, we included women ranging from 40 to 60 years old without history of cancer. All participants self-identified as Euro-descendants.

Information for histological type, breast cancer degree and immunohistochemical subtype, in addition to information on the presence of metastases in lymph nodes, was available for part of the patients (Table 1).

Genotyping

A total of 100 ng of high-quality DNA is fragmented using the KAPA HyperPlus kit (Roche), incubating for 5 minutes at 37 °C. Subsequently, the fragmented DNA

have their ends repaired, poly-A tail added, and are ligated through PCR to Illumina compatible dual index adapters uniquely barcoded. After ligation, fragments are purified with 0.8X ratio HyperCap magnetic beads (Roche) followed by double size selection (0.42X and 0.15X ratios) to select libraries of approximately 800 bp. Finally, libraries are amplified and purified with magnetic beads. After quantification by quantitative PCR, 30 ng of each sample are precisely pooled using ultrasonic acoustic energy, and the enrichment targeted capture is performed with hybridization kits from Twist Bioscience. Briefly, the libraries are bound to 10,456 biotinylated probes described previously (31). By using streptavidin magnetic beads, the fragments targeted by the probes are captured and then amplified and purified. Enriched libraries are analyzed in BioAnalyzer (Agilent) and quantified by digital-droplet PCR. Finally, enriched libraries are using NovaSeq6000 (Illumina) with paired-end 250 bp sequencing protocol. After sequencing, raw fastq files are analyzed using our custom bioinformatics pipeline (32) to obtain *KIR* copy number and high-resolution allelic data, and the software HLA Explore (Omixon) and AlloSeq Assign (CareDx) to obtain *HLA* genotypes.

Data Analysis

Gene-content haplotypes were identified manually, based on the precise copy number determination, the known linkage disequilibrium among *KIR* genes, and allelic information, as described previously (25).

HLA-A, HLA-B, and HLA-C were classified according to their known relevant epitopes for KIR recognition. We considered the following KIR-HLA pairs: HLA-A*11 and KIR2DS2 (33); HLA-A*03/HLA-A*11 and KIR3DL2 (34); HLA-C2 e KIR2DL1 (35); HLA-C1 and KIR2DL2 (35); HLA-C1 and KIR2DL3 (35); HLA-Bw4 and KIR3DL1 (36,37); HLA-C2 and KIR2DS1 (38).

Genetic association analysis was performed using the R package BIGDAWG (39), which handles multiallelic data to test for association at the haplotype, locus, and allele. Associations for presence and absence and *KIR* copy numbers were tested with the software IBM SPSS statistics software (40), using Mantel-

Haenszel method (41). The significance threshold was set at $p^{corr} < 0.05$ after correction for multiple tests using the Bonferroni method (42).

Results

KIR Copy Number Variation is Associated with Increased Risk to Sporadic Breast Cancer

Most previous studies of *KIR* in diseases, including cancer, were limited to analyzing presence and absence variation due to technical limitations to genotype *KIR* in high-resolution in a large scale (27,28,43–45). Leveraging the precise copy number (CN) determination of our method, we investigated if the CN variations of *KIR2DL1*, *KIR2DL23*, *KIR2DL4*, *KIR2DL5*, *KIR2DS1*, *KIR2DS2*, *KIR2DS35*, *KIR2DS4*, and *KIR3DL1S1* were associated with sporadic BC risk. For this analysis, we included all genes except *KIR3DL3* and *KIR3DL2*, which exhibit limited or no CN variation. The presence of three copies of *KIR2DL5* was associated with protection (OR = 0.21, 95%CI 0.07-0.618, p = 0.003, p^{corr} = 0.027), while the presence of three copies of *KIR3DL1S1* was associated with increased risk (OR=2.44, 95%CI 1.40-4.12, p = 0.001, p^{corr} =0.009; Table 2). We have previously demonstrated that the duplication of *KIR3DL1S1* is usually associated with the recently described duplication of genes *KIR3DP1~KIR2DL4~KIR3DS1* (25). We found that this recently discovered large structural variant significantly increases the risk of sporadic BC (OR = 2.42, 95%CI 1.30-4.61 p = 0.007) (Figure 1). We have not identified associations for other gene-content haplotypes (Suppl. table 1).

Despite the associations with *KIR* copy numbers, no associations were observed when we analyzed only the presence and absence of genes (suppl. table 2). The analysis of *KIR* CN variation in the immunohistochemical subtypes, tumor grade and presence of metastasis did not show significant results (supp. table 3)

KIR Allelic Variation is Associated with Aggressive Breast Cancer

Likewise, allelic variation of *KIR* in disease is poorly known in diseases due to the high levels of polymorphism and sequence identity among *KIR* genes (46–50). We compared the *KIR* allelic frequencies between patients and controls and among disease outcomes (Suppl table 4). Because the functionality of KIR molecules relies on the simultaneous presence of specific HLA ligands (REF), we furthered our analysis by comparing allelic frequencies in the presence of the known HLA ligands for each receptor. We have not identified associations when comparing patients with controls (Suppl. table 5), but KIR-HLA combinations were associated with more aggressive forms of BC.

*KIR2DL3*001 in the Presence of HLA-C1 is Associated with Non-Luminal HER2+ Breast Cancer*

The gene *KIR2DL23*, initially thought to be two genes (*KIR2DL2* and *KIR2DL3*) encode for potent inhibitory KIR that recognize subgroups of HLA-C molecules. KIR2DL2 and KIR2DL3 molecules recognize HLA-C1, a subgroup of HLA molecules carrying asparagine at residue at position 80 (reviewed by Augusto & Petzl-Erler, 2015 (51)). We identified that the allele *KIR2DL3*001* in the presence of HLA-C1 ligand was significantly reduced in patients with the aggressive non-luminal HER2+ (ER and PR negative) compared with all other subtypes (OR = 0.23, p = 0.002, p^{corr} = 0.016) (Figure 2; Suppl. Table 6) and with luminal (OR = 0.27, p = 0.004, p^{corr} = 0.022) (Suppl. Table 7). We did not find associations for ligand alone (Suppl. Table 8), indicating that this functional KIR-HLA pair was protective rather than the HLA ligand alone.

*KIR2DS1*002 in the Absence of HLA-C2 is Strongly Associated with Triple Negative Breast Cancer*

KIR2DS1 is an activating ligand that recognizes HLA-C2, the group of HLA molecules with lysine at the position 80 (38). We identified a strong association of *KIR2DS1*002* in the absence of HLA-C2 in TN patients compared to all other subtypes (OR = 3.34, $p < 2.7 \times 10^{-5}$, $p^{corr} < 1.4 \times 10^{-4}$) (Figure 3; suppl. Table 9) and with luminal cases only (OR = 3.18, $p < 6.6 \times 10^{-5}$, $p^{corr} < 3.3 \times 10^{-4}$) (suppl. table 10). No association was found for the ligand alone (results not shown).

Interestingly, we observed an opposite effect when we compared luminal patients versus other subtypes, (OR = 0.46, p = 0.006, p^{corr} = 0.03; Supp. Table 11), consistent with an overall association of KIR2DS1*002+ absence of HLA-C2 with the subtypes with a poor prognosis. No association was found for the ligand alone (results not shown).

KIR Associations with Breast Cancer Grades and Lymph Node Metastasis Status

KIR2DL1 also recognizes HLA-C2 molecules (35), and this pair has been described as the most potent for NK cell education against target cells (52–54). We observed a strong protective effect of the pair KIR2DL1*001+HLA-C2, with this combination significantly increased in patients with less severe BC grade I compared with the most severe grades II+III (OR = 0.25, p = 0.0048, p^{corr} = 0.038) (Figure 4; Suppl. Table 12).
Finally, we identified a borderline association of KIR2DL3*005 + HLA-C1 with presence or absence of metastasis in lymph nodes, being this combination increased in metastatic BC (OR = 3.9, p = 0.005, p^{corr} = 0.05) (Supp. table 13).

Discussion

To our knowledge, this is the first study to examine high-resolution *KIR* allelic variation and copy number in cancer. The few studies that have investigated the *KIR* in BC (27–30,55) have analyzed relatively small cohorts and were limited to identifying presence and absence of genes, which corresponds to the lowest genotyping resolution. However, copy-number directly affects the immune response (56), and the limitation of only assessing presence and absence results in loss of power to detect locus-level associations. Here, we found significant associations of *KIR* copy number and alleles with BC despite the complete lack of associations for presence and absence variation, a clear example that high resolution is required to assess the full extent of *KIR* associations with disease.
Applying our precise copy number determination allowed us to identify that the presence of three copies *of KIR3DL1S1* increases the risk of sporadic BC. Our previous work has demonstrated that the duplication of *KIR3DL1S1* is almost

always accompanied by duplication of the pseudogene *KIR3DP1* and *KIR2DL4*, forming a novel variant carried by 1% of European Americans from the U.S. (25). This recently discovered variant was strongly associated with susceptibility in our cohort of sporadic BC. Although our study cannot identify a mechanism, KIR3DL1S1 is a critical receptor for NK cell education (57–59), while KIR2DL4 expression is uniquely regulated by epigenetic mechanisms (60,61). Moreover, the duplication of these genes is associated with the presence of a chimeric *KIR3DP1* allele that is transcribed and putatively can produce a secreted protein (62).

We also found the presence of three copies of *KIR2DL5* associated with protection against sporadic BC. KIR2DL5 is a unique receptor that, differently from most KIR, do not seem to bind HLA (63). Recently, KIR2DL5 has been shown to bind the poliovirus receptor (PVR) on the cell surface, regulating the cytotoxicity of NK cells against tumor cells (64). Expressed in multiple malignancies, PVR is typically low expressed in normal tissues, and higher PVR levels have been associated with a poor prognosis in different cancers (65–67), including BC (68). PVR is also a ligand for TIGIT (T cell immunoreceptor with Ig and ITIM domains) (69) and blocking the TIGIT-PVR axis is promising for BC immunotherapy (68–71). Therefore, our findings that having more copies of KIR2DL5 protects against BC is intriguing and warrant further evaluation of how KIR2DL5 interacts with TIGIT-PVR in modulating BC susceptibility.

We identified that *KIR2DL3*001* in the presence of its ligand HLA-C1 was significantly reduced in HER2+ patients, a protective effect considering HER2+ BC tends to be more aggressive than other types. Because *KIR2DL3*001* is highly frequent, we also tested the presence of the gene KIR2DL3 (all alleles) in the presence of the ligand to check if the association could be driven by gene content instead of allelic variation. The presence of KIR2DL3 + HLA-C1 had a weaker effect (OR = 0.31) and was not significant after correction ($p = 0.014$, $p^{corr} = 0.154$), indicating an association with the allele *KIR2DL3*001*.

HER2 signaling can downregulate HLA class I expression in breast cancer cell lines in vitro (72–75), which was corroborated by immunohistochemistry (76) and

mRNA signatures in HER2+ tumors (77), increasing their susceptibility to NK cell cytotoxicity. It is widely accepted that cells containing inhibitory KIR receptors represent the main subset of educated NK lymphocytes (78). NK cells that developed without MHC class I are unable to kill tumor cell lines (79,80). Moreover, even in MHC class I-sufficient scenarios, NK cells lacking inhibitory receptors for self-MHC class I are hypo-responsive to in vitro stimulation through several activating receptors (19,81). Therefore, expression of self-MHC class I-reactive inhibitory receptors enhances the responsive potential of NK cells against infected and cancer cells (19,82). KIR2DL3*001 is known for a moderate avidity for HLA-C1 but with high specificity, with minimal cross-reactivity with C2 (52,83). Dunphy et al. (2015) has shown that the proportion of cells expressing KIR2DL3 is higher in *KIR2DL3*001* homozygotes in comparison to *KIR2DL3*002* homozygotes, therefore increasing the size of the NK cell subset that may interact with HLA-C1 (84). In addition, they also observed significantly higher expression levels on the cell surface of KIR2DL3*001 positive cells, biasing the toward higher expression in two ways. Interestingly, another *KIR2DL3* allele, *KIR2DL3*005*, was increased in metastatic patients. This allele encodes a receptor with increased affinity and avidity for HLA-C1 due to alterations in the interdomain hinge angle (95). Thus, the protective effect of KIR2DL3*001+HLA-C1 and KIR2DL3*005+HLA-C1 observed in patients with HER2+ tumors and with metastasis, respectively, might be in part explained by how this pair affects NK cell education in combination with the impact of HER2 signaling on HLA expression.

The importance of NK cells for early tumor immune surveillance is supported by studies showing increased cancer risk in individuals with low NK cell activity (85). HER2-specific antibodies, such as trastuzumab and pertuzumab, can activate NK cell-mediated antibody-dependent cellular cytotoxicity (ADCC), contributing to their antitumor efficacy (86). Antibody-dependent NK cell activation results in the release of cytotoxic granules as well as the secretion of pro-inflammatory cytokines and chemokines, resulting in cytolytic killing of tumor cells. Ehlers and collaborators (87) showed that KIR-ligand-mismatched NK cells degranulated stronger compared to the matched NK cells under trastuzumab treatment, which raised HER2+ cells cytotoxicity. However, many HER2 positive patients develop

resistance to mAB treatment, requiring additional alternative therapies. For example, anti-HER2 therapy benefits are higher in ER-negative tumors, and progressively decrease in tumors with increased ER expression (88). On the other hand, there is an increased resistance to NK cell mediated ADCC in estrogen positive HER2+ tumors. Estrogens also up regulate HLA class I in BC cell lines (89–91), potentially influencing susceptibility to NK cell activity. Interestingly, we observed a protective effect of KIR2DL3*001+HLA-C1 in HER2 positive/ ER negative tumors; however, we did not observe the same effect for ER positive tumors, which could be due to ER positive tumors' ability to evade NK cells cytotoxicity.

Finding effective treatment for triple negative breast cancer is challenging, and the gold standard treatments are surgery and chemotherapy (92). However, the prognosis is poor, and recurrence is common. Morales-Esteves and collaborators (93) previously showed that the presence of the gene *KIR2DS1* and HLA-C2 extended the efficacy of mAbs treatment in breast cancer patients. Interestingly, we observed that presence of the allele *KIR2DS1*002* without the ligand HLA-C2 increased the risk to triple negative BC approximately three times, a subtype of cancer that is not treatable with mAbs. Therefore, our results show that allelic variation of KIR2DS1 increases the risk of this aggressive type of BC and suggests this gene as a potential target for studies focusing on triple negative BC treatment.

We found the pair KIR2DL1*001+HLA-C2 associated with a lower risk of developing the more aggressive BC grades II and III, which could be a consequence of KIR2DL1 pivotal role in regulating NK cells. KIR2DL1 binds the ligand HLA-C2 with high avidity and affinity (38,52), and *KIR2DL1* alleles exhibit differential expression for both cell surface levels and proportion of NK cells expressing the receptor (52,84,94). Moreover, KIR2DL1 and HLA-C2 are responsible for the strongest regulatory signal among KIR, being the major educator of NK cell cytotoxicity (52–54). Our previous work has shown that a *KIR2DL1* variant that marks lower expression of this receptor is associated with higher HLA-C expression in global populations, possibly indicating a compensation mechanism driven by strong selective pressure (94).

A limitation of our study is the power to detect associations of variants with smaller effects in the subgroups of patients. Although we analyzed hundreds of patients and controls, not all subtypes are represented with large sample sizes. Despite this limitation, however, we were able to demonstrate that KIR variation plays a significant role in BC risk, and the genetic association is supported by previous functional studies.

Concluding Remarks

We have provided solid evidence that the variation of *KIR* genes is implicated in BC. More importantly, we have shown that high-resolution analysis of *KIR*, including allelic diversity and copy number analysis, is a requirement to fully assess the effect of *KIR* variation in disease. We hope our work contributes to identifying BC susceptibility markers that might be used for cancer prevention and treatment, particularly in patients from low- or middle-income countries typically neglected in major genomic efforts.

References

1. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, et al. SEER Cancer Statistics Review, 1975-2018. National Cancer Institute. Bethesda (2020). https://seer.cancer.gov/csr/1975_2018/index.html [Accessed March 21, 2022]
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* (2021) 71:209–249. doi: 10.3322/caac.21660
3. American Cancer Society. Cancer Facts and Figures 2022. Atlanta, GA: American Cancer Society (2022).
4. Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. *Annu Rev Genomics Hum Genet* (2008) 9:321–345. doi: 10.1146/annurev.genom.9.081307.164339
5. Jara L, Morales S, de Mayo T, Gonzalez-Hormazabal P, Carrasco V, Godoy R. Mutations in BRCA1, BRCA2 and other breast and ovarian cancer susceptibility genes in Central and South American populations. *Biol Res* (2017) 50:35. doi: 10.1186/s40659-017-0139-2
6. Kenemans P, Verstraeten RA, Verheijen RHM. Oncogenic pathways in hereditary and sporadic breast cancer. *Maturitas* (2004) 49:34–43. doi: 10.1016/j.maturitas.2004.06.005

7. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* (1991) 19:403–410. doi: 10.1111/j.1365-2559.1991.tb00229.x
8. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer* (1957) 11:359–377. doi: 10.1038/bjc.1957.43
9. Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, Blamey RW, Ellis IO. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol* (2008) 26:3153–3158. doi: 10.1200/JCO.2007.15.5986
10. Li J, Chen Z, Su K, Zeng J. Clinicopathological classification and traditional prognostic indicators of breast cancer. *Int J Clin Exp Pathol* (2015) 8:8500–8505.
11. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn H-J, Panel members. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* (2011) 22:1736–1747. doi: 10.1093/annonc/mdr304
12. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* (2009) 101:736–750. doi: 10.1093/jnci/djp082
13. Citron ML, Berry DA, Cirrincione C, Hudis C, Winer EP, Gradishar WJ, Davidson NE, Martino S, Livingston R, Ingle JN, et al. Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. *J Clin Oncol* (2003) 21:1431–1439. doi: 10.1200/JCO.2003.09.081
14. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene* (2007) 26:6469–6487. doi: 10.1038/sj.onc.1210477
15. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* (1989) 244:707–712. doi: 10.1126/science.2470152
16. Cooper MA, Colonna M, Yokoyama WM. Hidden talents of natural killers: NK cells in innate and adaptive immunity. *EMBO Rep* (2009) 10:1103–1110. doi: 10.1038/embor.2009.203
17. Parham P, Moffett A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat Rev Immunol* (2013) 13:133–144. doi: 10.1038/nri3370
18. Parham P. Killer cell immunoglobulin-like receptor diversity: balancing signals in the natural killer cell response. *Immunology Letters* (2004) 92:11–13. doi: 10.1016/j.imlet.2003.11.016
19. Anfossi N, André P, Guia S, Falk CS, Roetynck S, Stewart CA, Breso V, Frassati C, Reviron D, Middleton D, et al. Human NK cell education by inhibitory receptors for MHC class I. *Immunity* (2006) 25:331–342. doi: 10.1016/j.immuni.2006.06.013

20. Kärre K, Ljunggren HG, Piontek G, Kiessling R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* (1986) 319:675–678. doi: 10.1038/319675a0
21. Ljunggren HG, Kärre K. In search of the “missing self”: MHC molecules and NK cell recognition. *Immunol Today* (1990) 11:237–244. doi: 10.1016/0167-5699(90)90097-s
22. Uhrberg M, Parham P, Wernet P. Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. *Immunogenetics* (2002) 54:221–229. doi: 10.1007/s00251-002-0463-7
23. Uhrberg M, Valiante NM, Shum BP, Shilling HG, Lienert-Weidenbach K, Corliss B, Tyan D, Lanier LL, Parham P. Human Diversity in Killer Cell Inhibitory Receptor Genes. *Immunity* (1997) 7:753–763. doi: 10.1016/S1074-7613(00)80394-5
24. Vilches C, Rajalingam R, Uhrberg M, Gardiner CM, Young NT, Parham P. KIR2DL5, a Novel Killer-Cell Receptor with a D0-D2 Configuration of Ig-Like Domains. *The Journal of Immunology* (2000) 164:5797–5804. doi: 10.4049/jimmunol.164.11.5797
25. Amorim LM, Augusto DG, Nemat-Gorgani N, Montero-Martin G, Marin WM, Shams H, Dandekar R, Caillier S, Parham P, Fernández-Viña MA, et al. High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity. *Frontiers in Immunology* (2021) 12: https://www.frontiersin.org/article/10.3389/fimmu.2021.674778 [Accessed March 21, 2022]
26. Roe D, Vierra-Green C, Pyo C-W, Eng K, Hall R, Kuang R, Spellman S, Ranade S, Geraghty DE, Maiers M. Revealing complete complex KIR haplotypes phased by long-read sequencing technology. *Genes & Immunity* (2017) 18:127–134. doi: 10.1038/gene.2017.10
27. Ozturk OG, Gun FD, Polat G. Killer cell immunoglobulin-like receptor genes in patients with breast cancer. *Med Oncol* (2012) 29:511–515. doi: 10.1007/s12032-011-9932-x
28. Jobim MR, Jobim M, Salim PH, Portela P, Jobim LF, Leistner-Segal S, Bittelbrunn AC, Menke CH, Biazús JV, Roesler R, et al. Analysis of KIR gene frequencies and HLA class I genotypes in breast cancer and control group. *Human Immunology* (2013) 74:1130–1133. doi: 10.1016/j.humimm.2013.06.021
29. Hematian Larki M, Barani S, Talei A-R, Ghaderi A. Diversity of KIRs in invasive breast cancer patients and healthy controls along with the clinical significance in ER/PR/HER2+ patients. *Genes Immun* (2020) 21:380–389. doi: 10.1038/s41435-020-00117-1
30. Ashouri E, Rajalingam K, Barani S, Farjadian S, Ghaderi A, Rajalingam R. Coexistence of inhibitory and activating killer-cell immunoglobulin-like receptors to the same cognate HLA-C2 and Bw4 ligands confer breast cancer risk. *Sci Rep* (2021) 11:7932. doi: 10.1038/s41598-021-86964-y
31. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, Jayaraman J, Wroblewski EE, Trowsdale J, Rajalingam R, et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. *Am J Hum Genet* (2016) 99:375–391. doi: 10.1016/j.ajhg.2016.06.023

32. Marin WM, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, Lange V, Sauter J, Norman PJ, Hollenbach JA. High-throughput Interpretation of Killer-cell Immunoglobulin-like Receptor Short-read Sequencing Data with PING. *PLOS Computational Biology* (2021) 17:e1008904. doi: 10.1371/journal.pcbi.1008904
33. Liu J, Xiao Z, Ko HL, Shen M, Ren EC. Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A*11. *Proceedings of the National Academy of Sciences* (2014) 111:2662–2667. doi: 10.1073/pnas.1322052111
34. Hansasuta P, Dong T, Thananchai H, Weekes M, Willberg C, Aldemir H, Rowland-Jones S, Braud VM. Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. *Eur J Immunol* (2004) 34:1673–1679. doi: 10.1002/eji.200425089
35. Winter CC, Gumperz JE, Parham P, Long EO, Wagtmann N. Direct binding and functional transfer of NK cell inhibitory receptors reveal novel patterns of HLA-C allotype recognition. *J Immunol* (1998) 161:571–577.
36. Cella M, Longo A, Ferrara GB, Strominger JL, Colonna M. NK3-specific natural killer cells are selectively inhibited by Bw4-positive HLA alleles with isoleucine 80. *J Exp Med* (1994) 180:1235–1242. doi: 10.1084/jem.180.4.1235
37. Carr WH, Pando MJ, Parham P. KIR3DL1 polymorphisms that affect NK cell inhibition by HLA-Bw4 ligand. *J Immunol* (2005) 175:5222–5229. doi: 10.4049/jimmunol.175.8.5222
38. Stewart CA, Laugier-Anfossi F, Vély F, Saulquin X, Riedmuller J, Tisserant A, Gauthier L, Romagné F, Ferracci G, Arosa FA, et al. Recognition of peptide–MHC class I complexes by activating killer immunoglobulin-like receptors. *PNAS* (2005) 102:13224–13229. doi: 10.1073/pnas.0503594102
39. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. *Human Immunology* (2016) 77:283–287. doi: 10.1016/j.humimm.2015.12.006
40. IBM Corp. IBM SPSS Statistics for Windows, Version 27.0. (2020)
41. Mantel N, Haenszel W. Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *JNCI: Journal of the National Cancer Institute* (1959) 22:719–748. doi: 10.1093/jnci/22.4.719
42. Bonferroni CE. *Teoria statistica delle classi e calcolo delle probabilità*. Seeber (1936). 62 p.
43. Verheyden S, Bernier M, Demanet C. Identification of natural killer cell receptor phenotypes associated with leukemia. *Leukemia* (2004) 18:2002–2007. doi: 10.1038/sj.leu.2403525
44. Middleton D, Diler AS, Meenagh A, Sleator C, Gourraud PA. Killer immunoglobulin-like receptors (KIR2DL2 and/or KIR2DS2) in presence of their ligand (HLA-C1 group) protect against chronic myeloid leukaemia. *Tissue Antigens* (2009) 73:553–560. doi: 10.1111/j.1399-0039.2009.01235.x
45. Besson C, Roetynck S, Williams F, Orsi L, Amiel C, Lependeven C, Antoni G, Hermine O, Brice P, Ferme C, et al. Association of killer cell immunoglobulin-like receptor genes with Hodgkin's lymphoma in a familial study. *PLoS One* (2007) 2:e406. doi: 10.1371/journal.pone.0000406

46. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res* (2020) 48:D948–D955. doi: 10.1093/nar/gkz950
47. Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, Graef T, McQueen KL, Guethlein LA, Carrington CVF, et al. Meiotic recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. *Genome Res* (2009) 19:757–769. doi: 10.1101/gr.085738.108
48. Jiang W, Johnson C, Jayaraman J, Simecek N, Noble J, Moffatt MF, Cookson WO, Trowsdale J, Traherne JA. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. *Genome Res* (2012) 22:1845–1854. doi: 10.1101/gr.137976.112
49. Martin MP, Bashirova A, Traherne J, Trowsdale J, Carrington M. Cutting Edge: Expansion of the KIR Locus by Unequal Crossing Over. *J Immunol* (2003) 171:2192. doi: 10.4049/jimmunol.171.5.2192
50. Martin AM, Kulski JK, Gaudieri S, Witt CS, Freitas EM, Trowsdale J, Christiansen FT. Comparative genomic analysis, diversity and evolution of two KIR haplotypes A and B. *Gene* (2004) 335:121–131. doi: 10.1016/j.gene.2004.03.018
51. Augusto DG, Petzl-Erler ML. KIR and HLA under pressure: evidences of coevolution across worldwide populations. *Hum Genet* (2015) 134:929–940. doi: 10.1007/s00439-015-1579-9
52. Hilton HG, Guethlein LA, Goyos A, Nemat-Gorgani N, Bushnell DA, Norman PJ, Parham P. Polymorphic HLA-C receptors balance the functional characteristics of KIR haplotypes. *J Immunol* (2015) 195:3160–3170. doi: 10.4049/jimmunol.1501358
53. David G, Djaoud Z, Willem C, Legrand N, Rettman P, Gagne K, Cesbron A, Retière C. Large spectrum of HLA-C recognition by killer Ig-like receptor (KIR)2DL2 and KIR2DL3 and restricted C1 SPECIFICITY of KIR2DS2: dominant impact of KIR2DL2/KIR2DS2 on KIR2D NK cell repertoire formation. *J Immunol* (2013) 191:4778–4788. doi: 10.4049/jimmunol.1301580
54. Horowitz A, Djaoud Z, Nemat-Gorgani N, Blokhuis J, Hilton HG, Béziat V, Malmberg K-J, Norman PJ, Guethlein LA, Parham P. Class I HLA haplotypes form two schools that educate NK cells in different ways. *Sci Immunol* (2016) 1:eaag1672. doi: 10.1126/sciimmunol.aag1672
55. Alomar SY, Alkhuriji A, Trayhyrn P, Alhetheel A, Al-jurayyan A, Mansour L. Association of the genetic diversity of killer cell immunoglobulin-like receptor genes and HLA-C ligand in Saudi women with breast cancer. *Immunogenetics* (2017) 69:69–76. doi: 10.1007/s00251-016-0950-x
56. Pelak K, Need AC, Fellay J, Shianna KV, Feng S, Urban TJ, Ge D, De Luca A, Martinez-Picado J, Wolinsky SM, et al. Copy Number Variation of KIR Genes Influences HIV-1 Control. *PLoS Biol* (2011) 9:e1001208. doi: 10.1371/journal.pbio.1001208
57. Boudreau JE, Mulrooney TJ, Le Luduec J-B, Barker E, Hsu KC. KIR3DL1 and HLA-B Density and Binding Calibrate NK Education and Response to HIV. *J Immunol* (2016) 196:3398–3410. doi: 10.4049/jimmunol.1502469
58. Saunders PM, Pymm P, Pietra G, Hughes VA, Hitchen C, O'Connor GM, Loiacono F, Widjaja J, Price DA, Falco M, et al. Killer cell immunoglobulin-

like receptor 3DL1 polymorphism defines distinct hierarchies of HLA class I recognition. *J Exp Med* (2016) 213:791–807. doi: 10.1084/jem.20152023
59. Boulet S, Kleyman M, Kim JY, Kamya P, Sharafi S, Simic N, Bruneau J, Routy J-P, Tsoukas CM, Bernard NF. A combined genotype of KIR3DL1 high expressing alleles and HLA-B*57 is associated with a reduced risk of HIV infection. *AIDS* (2008) 22:1487–1491. doi: 10.1097/QAD.0b013e3282ffde7e
60. Trompeter H-I, Gómez-Lozano N, Santourlidis S, Eisermann B, Wernet P, Vilches C, Uhrberg M. Three Structurally and Functionally Divergent Kinds of Promoters Regulate Expression of Clonally Distributed Killer Cell Ig-Like Receptors (KIR), of KIR2DL4, and of KIR3DL3. *The Journal of Immunology* (2005) 174:4135–4143. doi: 10.4049/jimmunol.174.7.4135
61. Chan H-W, Kurago ZB, Stewart CA, Wilson MJ, Martin MP, Mace BE, Carrington M, Trowsdale J, Lutz CT. DNA Methylation Maintains Allele-specific KIR Gene Expression in Human Natural Killer Cells. *Journal of Experimental Medicine* (2003) 197:245–255. doi: 10.1084/jem.20021127
62. Gómez-Lozano N, Estefanía E, Williams F, Halfpenny I, Middleton D, Solís R, Vilches C. The silent KIR3DP1 gene (CD158c) is transcribed and might encode a secreted receptor in a minority of humans, in whom the KIR3DP1, KIR2DL4 and KIR3DL1/KIR3DS1 genes are duplicated. *European Journal of Immunology* (2005) 35:16–24. doi: 10.1002/eji.200425493
63. Cisneros E, Moraru M, Gómez-Lozano N, López-Botet M, Vilches C. KIR2DL5: An Orphan Inhibitory Receptor Displaying Complex Patterns of Polymorphism and Expression. *Front Immunol* (2012) 3:289. doi: 10.3389/fimmu.2012.00289
64. Husain B, Ramani SR, Chiang E, Lehoux I, Paduchuri S, Arena TA, Patel A, Wilson B, Chan P, Franke Y, et al. A Platform for Extracellular Interactome Discovery Identifies Novel Functional Binding Partners for the Immune Receptors B7-H3/CD276 and PVR/CD155. *Molecular & Cellular Proteomics* (2019) 18:2310–2323. doi: 10.1074/mcp.TIR119.001433
65. Masson D, Jarry A, Baury B, Blanchardie P, Laboisse C, Lustenberger P, Denis MG. Overexpression of the CD155 gene in human colorectal carcinoma. *Gut* (2001) 49:236–240. doi: 10.1136/gut.49.2.236
66. Gromeier M, Lachmann S, Rosenfeld MR, Gutin PH, Wimmer E. Intergeneric poliovirus recombinants for the treatment of malignant glioma. *Proc Natl Acad Sci U S A* (2000) 97:6803–6808. doi: 10.1073/pnas.97.12.6803
67. Lee BR, Chae S, Moon J, Kim MJ, Lee H, Ko HW, Cho BC, Shim HS, Hwang D, Kim HR, et al. Combination of PD-L1 and PVR determines sensitivity to PD-1 blockade. *JCI Insight* (2020) 5: doi: 10.1172/jci.insight.128633
68. Stamm H, Oliveira-Ferrer L, Grossjohann E-M, Muschhammer J, Thaden V, Brauneck F, Kischel R, Müller V, Bokemeyer C, Fiedler W, et al. Targeting the TIGIT-PVR immune checkpoint axis as novel therapeutic option in breast cancer. *OncoImmunology* (2019) 8:e1674605. doi: 10.1080/2162402X.2019.1674605
69. Bowers JR, Readler JM, Sharma P, Excoffon KJDA. Poliovirus Receptor: More than a simple viral receptor. *Virus Research* (2017) 242:1–6. doi: 10.1016/j.virusres.2017.09.001

70. Chauvin J-M, Zarour HM. TIGIT in cancer immunotherapy. *J Immunother Cancer* (2020) 8:e000957. doi: 10.1136/jitc-2020-000957
71. Chiang EY, Mellman I. TIGIT-CD226-PVR axis: advancing immune checkpoint blockade for cancer immunotherapy. *J Immunother Cancer* (2022) 10:e004711. doi: 10.1136/jitc-2022-004711
72. Herrmann F, Lehr H-A, Drexler I, Sutter G, Hengstler J, Wollscheid U, Seliger B. HER-2/neu-mediated regulation of components of the MHC class I antigen-processing pathway. *Cancer Res* (2004) 64:215–220. doi: 10.1158/0008-5472.can-2522-2
73. Okita R, Mougiakakos D, Ando T, Mao Y, Sarhan D, Wennerberg E, Seliger B, Lundqvist A, Mimura K, Kiessling R. HER2/HER3 signaling regulates NK cell-mediated cytotoxicity via MHC class I chain-related molecule A and B expression in human breast cancer cell lines. *J Immunol* (2012) 188:2136–2145. doi: 10.4049/jimmunol.1102237
74. Choudhury A, Charo J, Parapuram SK, Hunt RC, Hunt DM, Seliger B, Kiessling R. Small interfering RNA (siRNA) inhibits the expression of the Her2/neu gene, upregulates HLA class I and induces apoptosis of Her2/neu positive tumor cell lines. *Int J Cancer* (2004) 108:71–77. doi: 10.1002/ijc.11497
75. Vertuani S, Triulzi C, Roos AK, Charo J, Norell H, Lemonnier F, Pisa P, Seliger B, Kiessling R. HER-2/neu mediated down-regulation of MHC class I antigen processing prevents CTL-mediated tumor recognition upon DNA vaccination in HLA-A2 transgenic mice. *Cancer Immunol Immunother* (2009) 58:653–664. doi: 10.1007/s00262-008-0587-1
76. Inoue M, Mimura K, Izawa S, Shiraishi K, Inoue A, Shiba S, Watanabe M, Maruyama T, Kawaguchi Y, Inoue S, et al. Expression of MHC Class I on breast cancer cells correlates inversely with HER2 expression. *Oncoimmunology* (2012) 1:1104–1110. doi: 10.4161/onci.21056
77. Mamessier E, Sylvain A, Thibult M-L, Houvenaeghel G, Jacquemier J, Castellano R, Gonçalves A, André P, Romagné F, Thibault G, et al. Human breast cancer cells enhance self tolerance by promoting evasion from NK cell antitumor immunity. *J Clin Invest* (2011) 121:3609–3622. doi: 10.1172/JCI45816
78. Di Vito C, Mikulak J, Mavilio D. On the Way to Become a Natural Killer Cell. *Frontiers in Immunology* (2019) 10: https://www.frontiersin.org/article/10.3389/fimmu.2019.01812 [Accessed March 30, 2022]
79. Liao NS, Bix M, Zijlstra M, Jaenisch R, Raulet D. MHC class I deficiency: susceptibility to natural killer (NK) cells and impaired NK activity. *Science* (1991) 253:199–202. doi: 10.1126/science.1853205
80. Jonsson AH, Yokoyama WM. Natural killer cell tolerance licensing and other mechanisms. *Adv Immunol* (2009) 101:27–79. doi: 10.1016/S0065-2776(08)01002-X
81. Kim S, Poursine-Laurent J, Truscott SM, Lybarger L, Song Y-J, Yang L, French AR, Sunwoo JB, Lemieux S, Hansen TH, et al. Licensing of natural killer cells by host major histocompatibility complex class I molecules. *Nature* (2005) 436:709–713. doi: 10.1038/nature03847
82. Orr MT, Lanier LL. Natural killer cell education and tolerance. *Cell* (2010) 142:847–856. doi: 10.1016/j.cell.2010.08.031

83. Moesta AK, Norman PJ, Yawata M, Yawata N, Gleimer M, Parham P. Synergistic polymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C than KIR2DL3. *J Immunol* (2008) 180:3969–3979. doi: 10.4049/jimmunol.180.6.3969
84. Dunphy SE, Guinan KJ, Chorcora CN, Jayaraman J, Traherne JA, Trowsdale J, Pende D, Middleton D, Gardiner CM. 2DL1, 2DL2 and 2DL3 all contribute to KIR phenotype variability on human NK cells. *Genes Immun* (2015) 16:301–310. doi: 10.1038/gene.2015.15
85. Imai K, Matsuyama S, Miyake S, Suga K, Nakachi K. Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. *Lancet* (2000) 356:1795–1799. doi: 10.1016/S0140-6736(00)03231-1
86. Bhat R, Watzl C. Serial killing of tumor cells by human natural killer cells--enhancement by therapeutic antibodies. *PLoS One* (2007) 2:e326. doi: 10.1371/journal.pone.0000326
87. Ehlers FAI, Beelen NA, van Gelder M, Evers TMJ, Smidt ML, Kooreman LFS, Bos GMJ, Wieten L. ADCC-Inducing Antibody Trastuzumab and Selection of KIR-HLA Ligand Mismatched Donors Enhance the NK Cell Anti-Breast Cancer Response. *Cancers (Basel)* (2021) 13:3232. doi: 10.3390/cancers13133232
88. Bhargava R, Dabbs DJ, Beriwal S, Yildiz IA, Badve P, Soran A, Johnson RR, Brufsky AM, Lembersky BC, McGuire KP, et al. Semiquantitative hormone receptor level influences response to trastuzumab-containing neoadjuvant chemotherapy in HER2-positive breast cancer. *Mod Pathol* (2011) 24:367–374. doi: 10.1038/modpathol.2010.209
89. Hamada K, Gleason SL, Levi BZ, Hirschfeld S, Appella E, Ozato K. H-2RIIBP, a member of the nuclear hormone receptor superfamily that binds to both the regulatory element of major histocompatibility class I genes and the estrogen response element. *Proceedings of the National Academy of Sciences* (1989) 86:8289–8293. doi: 10.1073/pnas.86.21.8289
90. Sim BC, Hui KM. A HLA class I cis-regulatory element whose activity can be modulated by hormones. *International Journal of Cancer* (1994) 59:646–654. doi: 10.1002/ijc.2910590512
91. Rodríguez F, Perán F, Garrido F, Ruiz-Cabello F. Upmodulation by estrogen of HLA class I expression in breast tumor cell lines. *Immunogenetics* (1994) 39:161–167. doi: 10.1007/BF00241256
92. Kaufmann M, Hortobagyi GN, Goldhirsch A, Scholl S, Makris A, Valagussa P, Blohmer J-U, Eiermann W, Jackesz R, Jonat W, et al. Recommendations from an international expert panel on the use of neoadjuvant (primary) systemic treatment of operable breast cancer: an update. *J Clin Oncol* (2006) 24:1940–1949. doi: 10.1200/JCO.2005.02.6187
93. Morales-Estevez C, De la Haba-Rodriguez J, Manzanares-Martin B, Porras-Quintela I, Rodriguez-Ariza A, Moreno-Vega A, Ortiz-Morales MJ, Gomez-España MA, Cano-Osuna MT, Lopez-Gonzalez J, et al. KIR Genes and Their Ligands Predict the Response to Anti-EGFR Monoclonal Antibodies in Solid Tumors. *Front Immunol* (2016) 7:561. doi: 10.3389/fimmu.2016.00561
94. Vargas L de B, Dourado RM, Amorim LM, Ho B, Calonga-Solís V, Issler HC, Marin WM, Beltrame MH, Petzl-Erler ML, Hollenbach JA, et al. Single Nucleotide Polymorphism in KIR2DL1 Is Associated With HLA-C Expression

in Global Populations. *Front Immunol* (2020) 11:1881. doi: 10.3389/fimmu.2020.01881

95. Frazier WR, Steiner N, Hou L, Dakshanamurthy S, Hurley CK. Allelic Variation in KIR2DL3 Generates a KIR2DL2-like Receptor with Increased Binding to Its HLA-C Ligand. *J Immunol* (2013) 190:6198–6208. doi: 10.4049/jimmunol.1300464

Figure Legends

Figure 1. *KIR3DP1*, *KIR2DL4* and *KIR3DL1S1* duplications are associated with breast cancer.

Figure 2. KIR2DL3*001 in the presence of HLA-C1 is associated with protection against HER2 overexpression cases

Figure 3. KIR2DS2*001 in the absence of HLA-C2 is associated with susceptibility for triple negative breast cancer.

Figure 4. KIR2DL1*001 in the presence of HLA-C2 is protective against aggressive breast cancer (Grades 2 and 3).

Table legends

Table 1. Breast cancer patients stratification. Summary of breast cancer patients subtypes. f = frequency; NA = samples without subtype information.

Table 2. *KIR* copy number variation in breast cancer patients and controls. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies.

Supplementary information

Supplementary table 1. *KIR* haplotypes in patients versus controls. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; cA = Centromeric A haplotypes grouped; cB = Centromeric B haplotypes grouped; tA = Telomeric A haplotypes grouped; tB = Telomeric B haplotypes grouped.

Supplementary table 2. *KIR* presence and absence polymorphism frequencies in breast cancer patients, controls and subtypes. f = frequencies; TN = Triple Negative Breast Cancer; HER2+ = HER2 overexpression; Metastasis = Metastasis in lymph nodes.

Supplementary table 3. Copy number variation frequencies in BC patients, controls and subtypes. f = frequencies; TN = Triple Negative Breast Cancer; HER2+ = HER2 overexpression; Metastasis = Metastasis in lymph nodes.

Supplementary table 4. *KIR* allelic frequencies in breast cancer patients, controls and subtypes. f = frequencies; TN = Triple Negative Breast Cancer; HER2+ = HER2 overexpression; Metastasis = Metastasis in lymph nodes.

Supplementary table 5. *KIR* allelic variation and HLA ligand in breast cancer patients and controls. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies.

Supplementary table 6. *KIR* allelic variation and HLA ligand in HER2 patients vs other subtypes. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; HER2+ = HER2 overexpression.

Supplementary table 7. *KIR* allelic variation and HLA ligand in HER2+ patients versus luminal patients. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; HER2+ = HER2 overexpression.

Supplementary table 8. KIR ligand frequencies. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; TN = Triple Negative Breast Cancer.

Supplementary table 9. *KIR* allelic variation and HLA ligand in triple negative patients versus other subtypes. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; TN = Triple Negative Breast Cancer.

Supplementary table 10. *KIR* allelic variation and HLA ligand in triple negative patients vs luminal patients. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; TN = Triple Negative Breast Cancer.

Supplementary table 11. *KIR* allelic variation and HLA ligand in Luminal patients versus other subtypes. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.Upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies.

Supplementary table 12. *KIR* allelic variation and HLA ligand in Grade 2 and 3 breast cancer patients versus Grade 1. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.correcteded = p-value correct with Bonferroni; f = frequencies.

Supplementary table 13. *KIR* allelic variation and HLA ligand in patients with and without lymph node metastasis. OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies; Metastasis = Metastasis in lymph nodes.

Figure 1. *KIR3DP1*, *KIR2DL4* and *KIR3DL1S1* duplications are associated with breast cancer.

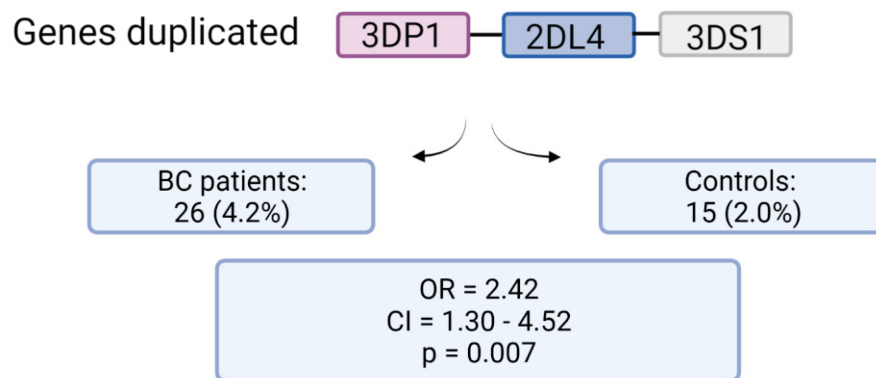


Figure 2. KIR2DL3*001 in the presence of HLA-C1 is associated with protection against HER2 overexpression cases

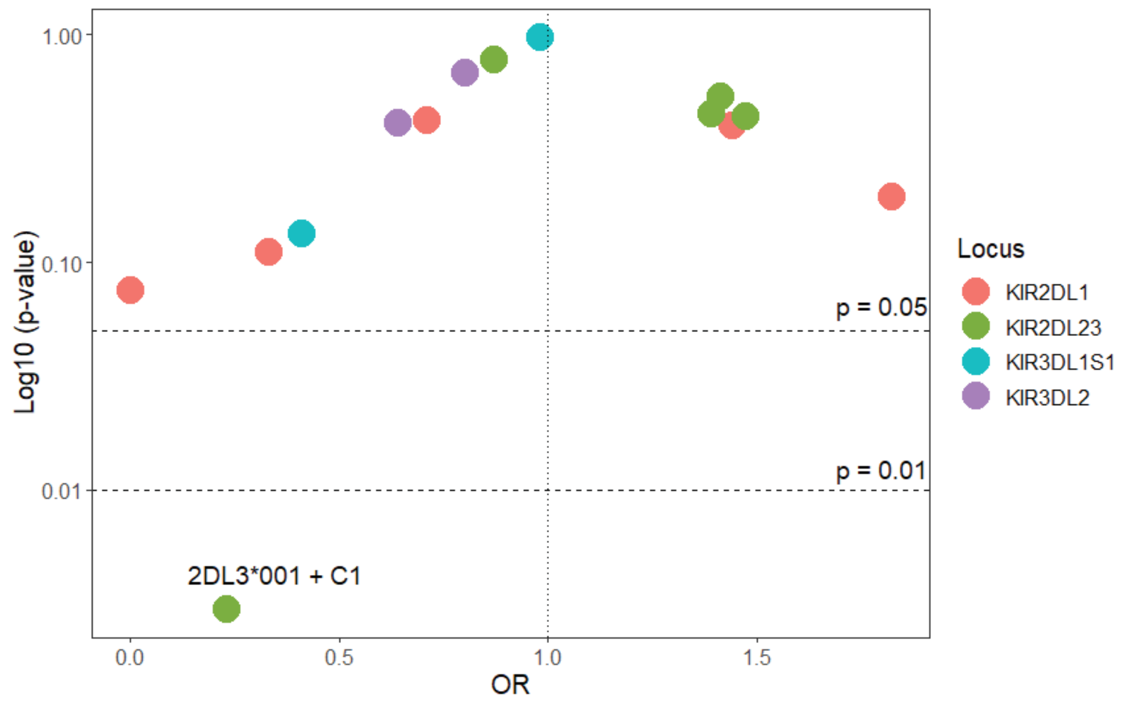


Figure 3. KIR2DS2*001 in the absence of HLA-C2 is associated with susceptibility for triple negative breast cancer.

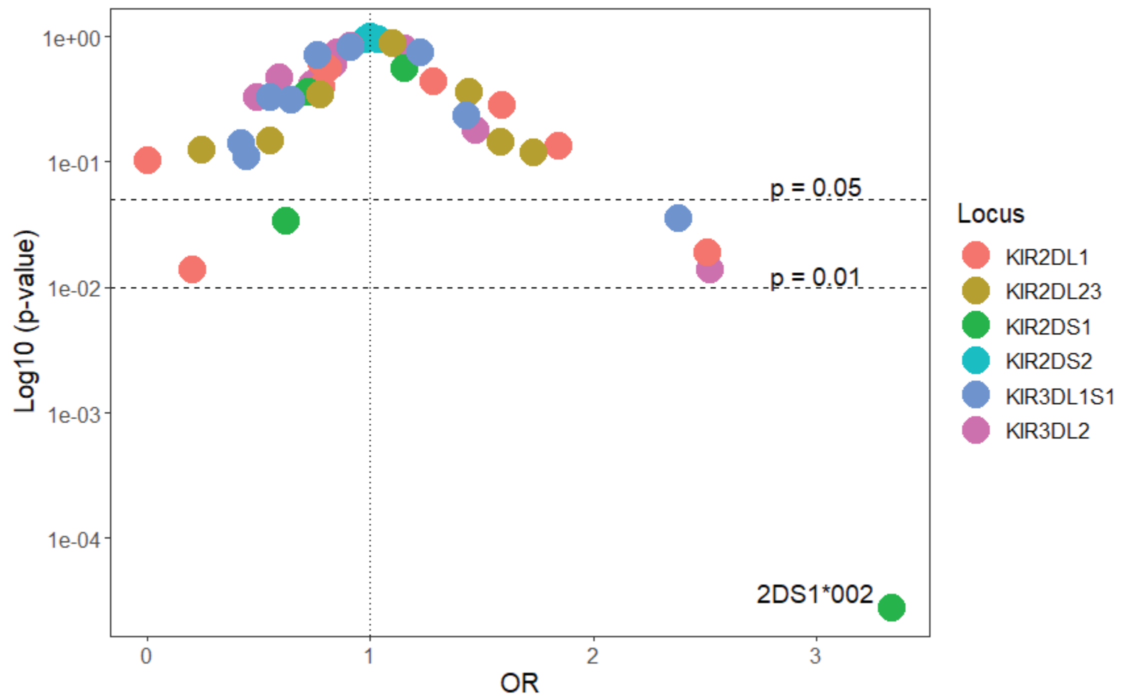
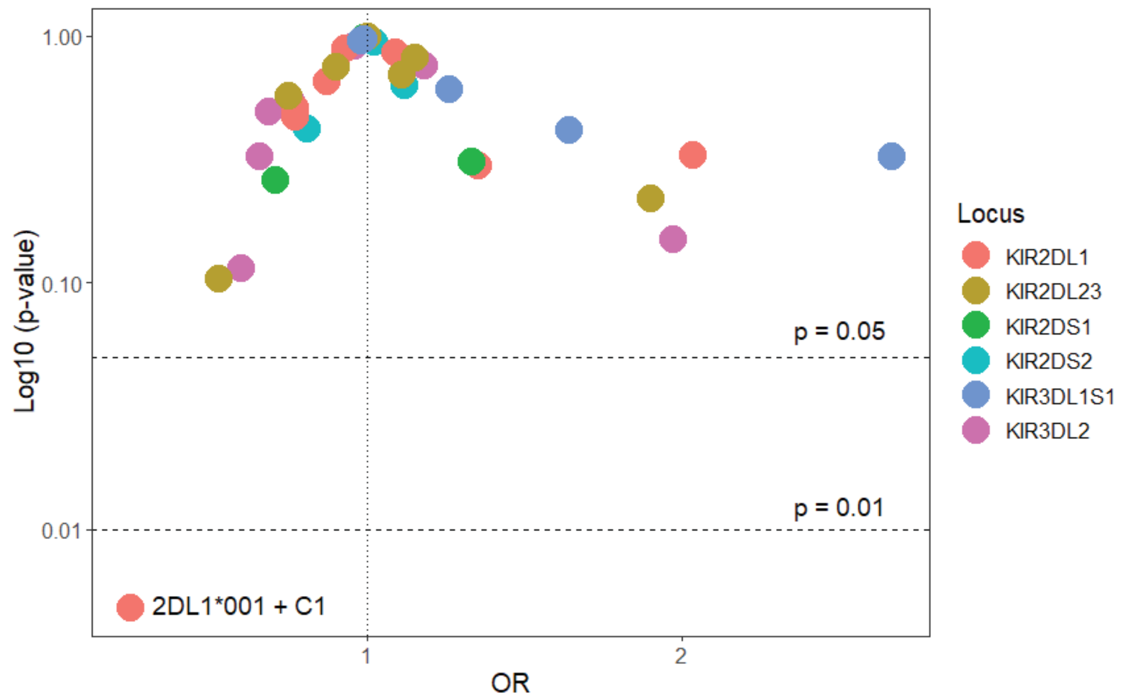


Figure 4. KIR2DL1*001 in the presence of HLA-C2 is protective against aggressive breast cancer (Grades 2 and 3).



Breast Cancer Cases (n = 550)

Degree of Aggressiviness	f	n	Immunohistochemical Subtype	f	n
Grade I	0.118	44	Luminal	0.815	317
Grade II	0.589	218	HER2 positive	0.062	24
Grade III	0.293	107	Triple Negative	0.123	48
NA	-	180	NA	-	161
Total	1	370	Total	1	389

Metastasis in lymph nodes	f	n
Positive	0.441	195
Negative	0.559	247
NA	-	108
Total	1	442

Table 1. Breast cancer patients stratification. Summary of breast cancer patients subtypes. f = frequency; NA = samples without subtype information.

Locus	Copy number	Cases	Controls	OR	CI.lower	CI.upper	***p***	***p.corrected***
		f (n)	*f* (n)					
KIR2DL23	2	0.972 (719)	0.942 (508)	0.48	0.27	0.84	0.009	0.081
	3	0.009 (7)	0.028 (15)	3	1.21	7.4	0.012	0.108
KIR2DL4	2	0.885 (477)	0.922 (682)	0.65	0.44	0.88	0.027	0.243
	3	0.05 (27)	0.027 (20)	1.88	1.05	3.38	0.033	0.297
KIR3DL1S1	2	0.916 (678)	0.859 (463)	0.55	0.39	0.79	0.001	0.009
	3	0.03 (22)	0.069 (37)	2.40	1.40	4.12	0.001	0.009
KIR2DL5	3	0.007 (4)	0.034 (25)	0.21	0.07	0.62	0.003	0.027

Table 2. KIR copy number variation in breast cancer patients and controls.

OR = Odds Ratio; CI.lower = Lower Confidence Interval; CI.Upper = Upper Confidence Interval; p.corrected = p-value corrected with Bonferroni; f = frequencies.

6.3 CAPÍTULO III: *HLA* Alleles Are Associated with Aggressive Tumors in Brazilian Women Diagnosed with Sporadic Breast Cancer.

HLA alleles are associated with aggressive tumors in Brazilian women diagnosed with sporadic breast cancer

Leonardo M. Amorim[1], Enilze Ribeiro[1], Márcia H. Beltrame[1], Jill A. Hollenbach[2], Danillo G. Augusto[1,2,3*]

[1]Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil.

[2]Department of Neurology, University of California, San Francisco, CA, USA

[3] Department of Biological Sciences, University of North Carolina at Charlotte

Corresponding Author:

Danillo G. Augusto, PhD

University of North Carolina at Charlotte

Department of Biological Sciences

danillo.augusto@uncc.edu

Abstract

Although HLA molecules play a critical role in the immune response, their implication in cancer susceptibility is less known. Here, we explore the HLA allelic variation in 576 patients with sporadic breast cancer (BC) and 743 controls. The allele *HLA-C*12:03* was associated with triple negative BC, as well as variations in amino acid positions 11 and 12 of *HLA-B*. *HLA-DPB1*17:01* was associated with protection to the development of lymph node metastasis. Moreover, amino acid positions 151 (H and R) of *HLA-A* and 9 (M) and 11 (V) in *HLA-DRB1* were associated with the more aggressive BC. Here, we show strong evidence that HLA polymorphism is implicated in the risk of aggressive BC subtypes and outcomes.

Introduction

The genes encoding human leukocyte antigen (HLA) molecules are the most polymorphic in the human genome and are located in the major histocompatibility complex in the short arm of chromosome six at position 6p21.3 (1). HLA molecules play a critical role in the immune response by presenting antigenic peptides to T cells (2). Despite the substantial evidence showing that *HLA* variation is pivotal in modulating the risk to viral and autoimmune diseases (3–12), the mechanisms behind its implications in cancer susceptibility are less known. Here, we explore the HLA allelic risk and outcomes in a cohort of Brazilian women diagnosed with sporadic breast cancer (BC).

HLA class I molecules are expressed on all nucleated and present peptides derived from cytosolic antigens to CD8+ T cells to eliminate tumor and infected cells (2), and interact with killer cell immunoglobulin-like receptor (KIR) expressed by NK cells (13,14). HLA class II, on the other hand, are expressed on antigen-presenting cells, which include B cells, activating T cells, macrophages, and dendritic cells, and present peptides from exogenous antigens to CD4+ T cells (2). In the context of cancer, HLA variation can affect immune evasion if certain variants present tumor antigens less efficiently (15).

Breast tumor is the most common type of cancer among women (16–21). The classification of breast tumors is important for the prognosis and treatment direction, and they are classified according to the degree of aggressiveness, lymph node metastasis status, immunohistochemical subtype, among others (22). BC patients can be classified according to the degree of aggressiveness, which classifies tumor cells based on their similarity to normal cells (23,24). This classification divides the tumors into grades I, II and III, with grade I being the least aggressive and grade III the worst prognosis (24,25). BC cases can also be classified by their immunohistochemical subtype; in this classification, the presence or absence of estrogen and progesterone receptors in the tumor cell is analyzed, as well as the amplification of the human epidermal growth factor receptor-type 2 gene (ERBB2 or HER2) and the Ki-67 labeling index, a marker of cell proliferation (26–29).

Luminal A is characterized by the presence of estrogen (ER) and progesterone (PR) receptors, HER2 negative, and having a KI67 expression below 14% (26,27). Luminal B subtype is divided into Luminal B HER2 negative, which is positive for ER and can also be positive for PR, negative for HER2 amplification, showing a high expression of KI67 (14% or more), or Luminal B HER2 positive type, being ER positive and having HER2 amplification (26–28). The fourth subtype is called non-luminal enriched HER2, which has both ER and PR negative and HER2 positive (26,27). Ki67 expression is high, and these tumors are usually grade III and have an unfavorable prognosis (29). The last subtype considered in our study is the triple negative (TN) which is negative for ER, PR and HER2 amplification (26,27,30). This subtype is the most aggressive of all and has the worst prognosis.

Only a few studies analyzed the association of *HLA* alleles with sporadic BC, even fewer the context of BC subtypes (31–35), and none in a South American population. Recent studies have shown that while the BC mortality is reducing in high-income income countries (36–39), the mortality rates in Brazil is rapidly increasing every year (40). Therefore, understanding how HLA variation affects BC risk in Brazilian women might provide novel insights about the disease mechanisms that could ultimately contribute to reducing health disparities. Here, we analyzed HLA class I and II allelic variation in a well characterized Brazilian BC cohort and healthy controls, as well as subtypes of BC and the development of lymph node metastasis.

Methods

We analyzed genomic DNA from peripheral blood cells from 576 women with sporadic BC and 743 unrelated healthy women. Samples were collected by oncologists from the Federal University of Paraná, Curitiba, Brazil, who diagnosed and classified the patients according to clinical guidelines jointly established by the American Society of Clinical Oncology and the College of American Pathology. The median age in patients is 56 years old. Women ranging from 40 to 60 years old without history of cancer were included as control group.

BC cases were stratified according to degree of aggressiveness, immunohistochemical subtype, and the presence of metastasis in lymph nodes (Suppl. Table 1). In this study we considered the immunohistochemical subtypes defined in the St. Gallen conference (26,27).

A total of 100 ng of high-quality DNA is fragmented using the Twist EF Kit 2.0 I (Twist Bioscience), incubating for 5 minutes at 37 °C. Subsequently, the fragmented DNA have their ends repaired, poly-A tail added, and are ligated through PCR to Illumina compatible dual index adapters uniquely barcoded. After ligation, fragments are purified with 0.8X ratio Ampure XP magnetic beads (Beckman Coulter) followed by double size selection (0.42X and 0.15X ratios) to select libraries of approximately 800 bp. Finally, libraries are amplified and purified with magnetic beads. After quantification by quantitative PCR, 30 ng of each sample are precisely pooled using ultrasonic acoustic energy, and the enrichment targeted capture is performed with hybridization kits from Twist Bioscience. Briefly, the libraries are bound to 2,064 biotinylated 120 bp probes. By using streptavidin magnetic beads, the targeted fragments are captured and then amplified and purified. Enriched libraries are analyzed in BioAnalyzer (Agilent) and quantified by digital-droplet PCR. Finally, enriched libraries are sequenced using NovaSeq6000 or HiSeq4000 (Illumina) with paired end 150bp sequencing protocol. After sequencing, raw fastq files are processed simultaneously by the software HLA Explore (Omixon) and AlloSeq Assign (CareDx) to obtain HLA genotypes. Genetic association analysis was performed using the R package BIGDAWG (41), which handles multiallelic data to test for association at the haplotype, locus, and allele. The significance threshold was set at $p^{corr} < 0.05$ after correction for multiple tests using the Bonferroni method (42).

Results

We analyzed HLA class I and II allelic variation and their corresponding amino acid variation in BC patients and controls. We also compared subgroups of patients regarding the degree of aggressiveness, immunohistochemical subtypes, and presence of lymph node metastasis. Although we did not find any association for allelic or amino acid variation when comparing BC patients with

controls (Suppl Table 2 and 3), we identified strong associations when comparing patients with different tumor subtypes.

We observed that *HLA-C*12:03* was associated with TN cases compared to other subtypes (OR = 2.7, CI 1.27 - 5.43, p = 0.0025, p^{corr} = 0.033) (Figure 1A; Suppl. Table 2). For the amino acid analysis, we identified a few positions in different HLA genes associated with TN cases. Amino acid positions 11 and 12 in HLA-B are associated with susceptibility or protection in TN cases. An alanine (A) in position 11 of HLA-B confers protection to the development of TN breast cancer (OR = 0.53, CI 0.33 - 0.84, p = 0.0037, p^{corr} = 0.037), while the presence of a serine (S) in the same position confers risk (OR = 1.90, CI 1.19 - 3.00, p = 0.0037, p^{corr} = 0.037). Regarding position 12, a methionine (M) confers protection (OR = 0.54, CI 0.34 - 0.84, p = 0.0038, p^{corr} = 0.038) while a valine (V) confers risk (OR = 1.88, CI 1.19 - 2.98, p = 0.0038, p^{corr} = 0.038) (Figure 1B; Suppl. Table 3).

We also observed that *HLA-DPB1*17:01* was associated with protection against the development of lymph node metastasis. The frequency of this allele was significantly increased in patients that did not develop lymph node metastasis (OR = 0.21, CI 0.05 - 0.63, p = 0.0019, p^{corr} = 0.037) (Figure 1C; Suppl. Table 2). Moreover, we compared the frequencies of *HLA* class I and II alleles between BC cases with and without the development of lymph nodes metastasis. The presence of histidine (H) in position 151 of HLA-A was associated with protection to grade III BC (OR = 0.3, CI 0.12 - 0.67, p = 0.0016, p^{corr} = 0.016) while an arginine (R) confers risk (OR = 3.36, CI 1.49 - 8.51, p = 0.0016, p^{corr} = 0.016), when compared to grade I. In addition, two positions in HLA-DRB1 were also associated with protection against grade III BC: the presence of a lysine (K) in position 9 and an aspartic acid (D) in position 11 were increased in grade I (OR = 0.13, CI 0.01 - 0.73, p = 0.0033, p^{corr} = 0.033) (Figure 1D; Suppl Table 3).

Discussion

To our knowledge, this is the first study to examine *HLA* class I and II allelic variation in sporadic BC immunohistochemical subtypes, cancer grade, and the development of lymph node metastasis.

In our study, *HLA-C*12:03* was associated to protection against TN breast cancer. This allele was also associated in a recent meta-analysis that included 12,901 breast cancer cases and 12,583 controls (33). Although the authors did not observe any associations between any *HLA* allele and breast cancer risk using the stringent p-value cutoff of $p < 5x10^{-8}$, *HLA-C*12:03* was the allele with the lowest *p*-value (OR = 1.29, $p = 1.08x10^{-3}$).

In our study, *HLA-DPB1*17:01* was shown to be protective against the development of lymph node metastasis. HLA-DPB1 is a classical class II molecule that plays a central role in the immune system by presenting peptides derived from extracellular proteins. Variation of HLA-DPB1 has been previously associated with different types of cancer such as Hodkign`s lymphoma (43), acute lymphoblastic leukemia (44) and cervical cancer (45). Previously, clinical observations and animal experiments have shown that, although tumor cells are derived from host cells, they elicit immune responses. Histopathological studies show that many tumors are surrounded by mononuclear cell infiltrates composed of T lymphocytes, natural killer (NK) cells, and activated macrophages, and that activated lymphocytes and macrophages are present in lymph nodes that drain tumor growth sites. Therefore, the presence of specific HLA molecules can affect the immune response in different ways.

The use of genetically modified T cells is being explored as a treatment for patients with metastatic cancer. Recently, Lu and collaborators demonstrated the safety and efficacy of administering autologous CD4+ T cells (a method that involves the administration of a patient’s own autologous immune cells) that are genetically engineered to express HLA*DPB1*0401-restricted antitumor TCR that targets MAGE-A3 as therapy for patients with metastatic cancer (46). So, our findings may present a new target molecule that could be used as cancer immunotherapy.

In our study, we identified a polymorphism in two positions of *HLA-B* that were associated with TN breast cancer cases. Amino acids alanine and methionine present in positions 11 and 12, respectively, define a subgroup of HLA-B alleles

(*HLA-B*08, HLA-B *13, HLA-B *15, HLA -B *35, HLA-B *40, HLA-B *44-47, HLA-B *49, HLA-B *50, HLA-B *51-59, HLA-B *78, HLA-B *82 and HLA-B *83*). A previous study demonstrated that *HLA-B*13:01* was associated to protection against BC, whilst *HLA-B*55:01* was associated with aggravated risk (31). De Léon *et al* (2009) performed a study in a mestizo Mexican population, and found that *HLA-B*15:01* was associated with risk (OR, 3.714; p = 0.031) (35). While these authors found that these alleles confer risk to the development of BC, we identified that the amino acid polymorphisms identified in theses alleles actually confer protection against the development of aggressive TN breast cancer.

Regarding the amino acid changes in positions 11 and 12 of HLA-B, we found that an alanine (A) in position 11 of HLA-B confers protection to the development of TN breast cancer, while the presence of a serine (S) in the same position confers risk. According to the Grantham score the change A>S is considered moderately conservative (Score = 99). The Grantham score is a prediction of the effect of substitutions between amino acids based on chemical properties, including polarity and molecular volume, characterized into classes of increasing chemical dissimilarity: conservative (0-50), moderately conservative (51-100), moderately radical (101-150), or radical (≥151) (48).

Regarding position 12, a methionine (M) confers protection while a valine (V) confers risk. According to the Grantham score the change M>V is considered conservative (Score = 21). Therefore, these changes are not likely to affect the protein function.

We also identified polymorphisms in *HLA-A* and *HLA-DRB1* associated to protection or risk to the development of grade III BC. In *HLA-A* position 151, we identified an arginine that defines a subgroup of alleles: *HLA-A*23, HLA-A*29-33, HLA-A*74*. HLA-A was not associated with breast cancer in previous studies. Position 151 is part of the alpha 2 chain and is not predicted to affect protein binding (47). According to the Grantham score the change H>R is considered conservative (Score = 29). Therefore, regarding our results, the substitution is not predicted to cause any deleterious effect in the molecule.

In *HLA-DRB1*, the presence of a lysine and an aspartic acid in positions 9 and 11 respectively, were associated to protection against grade III BC. These two amino acids distinguish the *HLA-DRB1*09* group of alleles. *HLA-DRB1*09:01* was recently associated to severe cases of COVID-19 in the Japanese population (49). This specific allele group was not previously associated with BC. However, there are many reports of *HLA-DRB1* alleles associated with BC. Bayraktar et al found that *HLA-DRB1*1801* was associated to protection against BC (31). Aureli and collaborators found *HLA-DRB1*10:01* and *HLA-DRB1*11:01* associated with risk of development of BC (34). Moreover, Gun et al reported that *HLA-DRB1*13* was associated with protection against BC (32). However, HLA-*DRB1*13:01* was associated with risk of developing BC in a Mexican mestizo population (35).
Previous studies identified polymorphisms in *HLA-G* associated to breast cancer, where in a meta-analysis the authors identified the association of *HLA-G* 14-bp Insertion/Deletion and *HLA-G* +3142 C/G polymorphisms with BC (50). *HLA-G* was also identified as a pivotal mediator of breast cancer resistance to trastuzumab, as HLA-G binding to KIR2DL4 impairs the cytotoxicity of NK cells in HER2 positive breast cancer microenvironment (51). Despite the possible importance of *HLA-G* to BC, we did not find any polymorphism of *HLA-G* associated to BC and its subtypes in our study.

Many studies have been performed for HLA and BC (31–35), however none of them explored BC grade, immunohistochemical subtypes and lymph nodes metastasis status, especially in South American women. HLA has an important role in presenting antigens to T lymphocytes, and for fighting against tumor cells. In our study we did not find any HLA allele or amino acid position associated with sporadic BC. However, here we show that HLA is associated with more aggressive subtypes such as TN and grade III BC and is also associated with lymph node metastasis.

References

1. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC, Wright MW, et al. Gene map of the extended human MHC. *Nat Rev Genet* (2004) 5:889–899. doi: 10.1038/nrg1489

2. Klein J, Sato A. The HLA system. First of two parts. *N Engl J Med* (2000) 343:702–709. doi: 10.1056/NEJM200009073431006

3. Singal DP, Blajchman MA. Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes* (1973) 22:429–432. doi: 10.2337/diab.22.6.429

4. Cudworth AG, Woodrow JC. HL-A system and diabetes mellitus. *Diabetes* (1975) 24:345–349. doi: 10.2337/diab.24.4.345

5. van der Horst-Bruinsma IE, Visser H, Hazes JM, Breedveld FC, Verduyn W, Schreuder GM, de Vries RR, Zanelli E. HLA-DQ-associated predisposition to and dominant HLA-DR-associated protection against rheumatoid arthritis. *Hum Immunol* (1999) 60:152–158. doi: 10.1016/s0198-8859(98)00101-3

6. Milicic A, Lee D, Brown MA, Darke C, Wordsworth BP. HLA-DR/DQ haplotype in rheumatoid arthritis: novel allelic associations in UK Caucasians. *J Rheumatol* (2002) 29:1821–1826.

7. Newton JL, Harney SMJ, Wordsworth BP, Brown MA. A review of the MHC genetics of rheumatoid arthritis. *Genes Immun* (2004) 5:151–157. doi: 10.1038/sj.gene.6364045

8. Simmonds MJ, Howson JMM, Heward JM, Carr-Smith J, Franklyn JA, Todd JA, Gough SCL. A novel and major association of HLA-C in Graves' disease that eclipses the classical HLA-DRB1 effect. *Hum Mol Genet* (2007) 16:2149–2153. doi: 10.1093/hmg/ddm165

9. Lozano JM, González R, Kindelán JM, Rouas-Freiss N, Caballos R, Dausset J, Carosella ED, Peña J. Monocytes and T lymphocytes in HIV-1-positive patients express HLA-G molecule. *AIDS* (2002) 16:347–351. doi: 10.1097/00002030-200202150-00005

10. Bardeskar NS, Mania-Pramanik J. HIV and host immunogenetics: unraveling the role of HLA-C. *HLA* (2016) 88:221–231. doi: 10.1111/tan.12882

11. Wang S-F, Chen K-H, Chen M, Li W-Y, Chen Y-J, Tsao C-H, Yen M, Huang JC, Chen Y-MA. Human-leukocyte antigen class I Cw 1502 and class II DR 0301 genotypes are associated with resistance to severe acute respiratory syndrome (SARS) infection. *Viral Immunol* (2011) 24:421–426. doi: 10.1089/vim.2011.0024

12. Wang W, Zhang W, Zhang J, He J, Zhu F. Distribution of HLA allele frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). *HLA* (2020) 96:194–196. doi: 10.1111/tan.13941

13. Trowsdale J. Genetic and functional relationships between MHC and NK receptor genes. *Immunity* (2001) 15:363–374. doi: 10.1016/s1074-7613(01)00197-2

14. Parham P. Killer cell immunoglobulin-like receptor diversity: balancing signals in the natural killer cell response. *Immunology Letters* (2004) 92:11–13. doi: 10.1016/j.imlet.2003.11.016

15. Hazini A, Fisher K, Seymour L. Deregulation of HLA-I in cancer and its central importance for immunotherapy. *J Immunother Cancer* (2021) 9:e002899. doi: 10.1136/jitc-2021-002899

16. Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, Khaled H, Liu M-C, Martin M, Namer M, et al. The Global Breast Cancer Burden: Variations in Epidemiology and Survival. *Clinical Breast Cancer* (2005) 6:391–401. doi: 10.3816/CBC.2005.n.043

17. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, et al. SEER Cancer Statistics Review, 1975-2018. National Cancer Institute. Bethesda (2020). https://seer.cancer.gov/csr/1975_2018/index.html [Accessed March 21, 2022]

18. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72:7–33. doi: 10.3322/caac.21708

19. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* (2021) 71:209–249. doi: 10.3322/caac.21660

20. American Cancer Society. Cancer Facts and Figures 2022. Atlanta, GA: American Cancer Society (2022).

21. Mamessier E, Sylvain A, Thibult M-L, Houvenaeghel G, Jacquemier J, Castellano R, Gonçalves A, André P, Romagné F, Thibault G, et al. Human breast cancer cells enhance self tolerance by promoting evasion from NK cell antitumor immunity. *J Clin Invest* (2011) 121:3609–3622. doi: 10.1172/JCI45816

22. Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. *Annu Rev Genomics Hum Genet* (2008) 9:321–345. doi: 10.1146/annurev.genom.9.081307.164339

23. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer* (1957) 11:359–377. doi: 10.1038/bjc.1957.43

24. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* (1991) 19:403–410. doi: 10.1111/j.1365-2559.1991.tb00229.x

25. Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, Blamey RW, Ellis IO. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol* (2008) 26:3153–3158. doi: 10.1200/JCO.2007.15.5986

26. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn H-J, Panel members. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* (2011) 22:1736–1747. doi: 10.1093/annonc/mdr304

27. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn H-J, Panel members. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* (2013) 24:2206–2223. doi: 10.1093/annonc/mdt303

28. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* (2009) 101:736–750. doi: 10.1093/jnci/djp082

29. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* (1989) 244:707–712. doi: 10.1126/science.2470152

30. Anders C, Carey LA. Understanding and treating triple-negative breast cancer. *Oncology (Williston Park)* (2008) 22:1233–1239; discussion 1239-1240, 1243.

31. Bayraktar B, Yilmaz E, Bayraktar O, Apaydin BB, Erguner IE, Kayabasi B, Ozcelik AA, Eren B. Distribution of HLA antigens in breast cancer. *Bratisl Lek Listy* (2012) 113:372–375. doi: 10.4149/bll_2012_084

32. Gun FD, Ozturk OG, Polat A, Polat G. HLA class-II allele frequencies in Turkish breast cancer patients. *Med Oncol* (2012) 29:466–471. doi: 10.1007/s12032-011-9873-4

33. Ho PJ, Khng AJ, Tan BK-T, Tan EY, Tan S-M, Tan VKM, Lim GH, Aronson KJ, Chan TL, Choi J-Y, et al. Relevance of the MHC region for breast cancer susceptibility in Asians. *Breast Cancer* (2022) 29:869–879. doi: 10.1007/s12282-022-01366-w

34. Aureli A, Canossi A, Del Beato T, Buonomo O, Rossi P, Roselli M, Papola F, Marziani B, Sconocchia G. Breast Cancer Is Associated with Increased HLA-DRB1*11:01 and HLA-DRB1*10:01 Allele Frequency in a Population

of Patients from Central Italy. *Immunol Invest* (2020) 49:489–497. doi: 10.1080/08820139.2020.1737539

35. Cantú de León D, Pérez-Montiel D, Villavicencio V, García Carranca A, Mohar Betancourt A, Acuña-Alonzo V, López-Tello A, Vargas-Alarcón G, Barquera R, Yu N, et al. High resolution human leukocyte antigen (HLA) class I and class II allele typing in Mexican mestizo women with sporadic breast cancer: case-control study. *BMC Cancer* (2009) 9:48. doi: 10.1186/1471-2407-9-48

36. Guo F, Kuo Y-F, Shih YCT, Giordano SH, Berenson AB. Trends in breast cancer mortality by stage at diagnosis among young women in the United States. *Cancer* (2018) 124:3500–3509. doi: 10.1002/cncr.31638

37. Autier P, Boniol M, LaVecchia C, Vatten L, Gavin A, Héry C, Heanue M. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. *BMJ* (2010) 341:c3620. doi: 10.1136/bmj.c3620

38. Tabar L, Yen M-F, Vitak B, Chen H-HT, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* (2003) 361:1405–1410. doi: 10.1016/S0140-6736(03)13143-1

39. Anderson WF, Jatoi I, Devesa SS. Assessing the impact of screening mammography: Breast cancer incidence and mortality rates in Connecticut (1943-2002). *Breast Cancer Res Treat* (2006) 99:333–340. doi: 10.1007/s10549-006-9214-z

40. Silva JDDE, de Oliveira RR, da Silva MT, Carvalho MD de B, Pedroso RB, Pelloso SM. Breast Cancer Mortality in Young Women in Brazil. *Frontiers in Oncology* (2021) 10: https://www.frontiersin.org/articles/10.3389/fonc.2020.569933 [Accessed September 4, 2022]

41. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. *Human Immunology* (2016) 77:283–287. doi: 10.1016/j.humimm.2015.12.006

42. Bonferroni CE. *Teoria statistica delle classi e calcolo delle probabilità*. Seeber (1936). 62 p.

43. Taylor GM, Gokhale DA, Crowther D, Woll P, Harris M, Alexander F, Jarrett R, Cartwright RA. Increased frequency of HLA-DPB1*0301 in Hodgkin’s disease suggests that susceptibility is HVR-sequence and subtype-associated. *Leukemia* (1996) 10:854–859.

44. Taylor M, Harrison C, Eden T, Birch J, Greaves M, Lightfoot T, Hussain A, on behalf of UKCCS Investigators. HLA-DPB1 supertype-associated protection from childhood leukaemia: relationship to leukaemia karyotype

and implications for prevention. *Cancer Immunol Immunother* (2008) 57:53–61. doi: 10.1007/s00262-007-0349-5

45. Cheng L, Guo Y, Zhan S, Xia P. Association between HLA-DP Gene Polymorphisms and Cervical Cancer Risk: A Meta-Analysis. *Biomed Res Int* (2018) 2018:7301595. doi: 10.1155/2018/7301595

46. Lu Y-C, Parker LL, Lu T, Zheng Z, Toomey MA, White DE, Yao X, Li YF, Robbins PF, Feldman SA, et al. Treatment of Patients With Metastatic Cancer Using a Major Histocompatibility Complex Class II–Restricted T-Cell Receptor Targeting the Cancer Germline Antigen MAGE-A3. *JCO* (2017) 35:3322–3329. doi: 10.1200/JCO.2017.74.5463

47. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research* (2021) 49:D480–D489. doi: 10.1093/nar/gkaa1100

48. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* (1974) 185:862–864. doi: 10.1126/science.185.4154.862

49. Anzurez A, Naka I, Miki S, Nakayama-Hosoya K, Isshiki M, Watanabe Y, Nakamura-Hoshi M, Seki S, Matsumura T, Takano T, et al. Association of HLA-DRB1*09:01 with severe COVID-19. *HLA* (2021) 98:37–42. doi: 10.1111/tan.14256

50. Tizaoui K, Jalouli M, Ouzari H, Harrath AH, Rizzo R, Boujelbene N, Zidi I. 3′UTR-HLA-G polymorphisms and circulating sHLA-G are associated with breast cancer: Evidence from a meta-analysis. *Immunology Letters* (2022) 248:78–89. doi: 10.1016/j.imlet.2022.06.010

51. Zheng G, Guo Z, Li W, Xi W, Zuo B, Zhang R, Wen W, Yang A-G, Jia L. Interaction between HLA-G and NK cell receptor KIR2DL4 orchestrates HER2-positive breast cancer resistance to trastuzumab. *Sig Transduct Target Ther* (2021) 6:1–15. doi: 10.1038/s41392-021-00629-w

Figure Legends

Figure 1A. *HLA-C*12:03* is associated with increased risk to the development of triple negative breast cancer.

Figure 1B. Changes in amino acid positions 11 and 12 of HLA-B are associated with triple negative breast cancer.

Figure 1C. *HLA-DPB1*17:01* is associated with protection gainst lymph node metastasis.

Figure 1D. Positions 9 and 11 of HLA-DRB1, and 151 of HLA-A are associated with risk and protection against grade III BC

Figure 1A. *HLA-C*12:03* is associated with increased risk to the development of triple negative breast cancer.

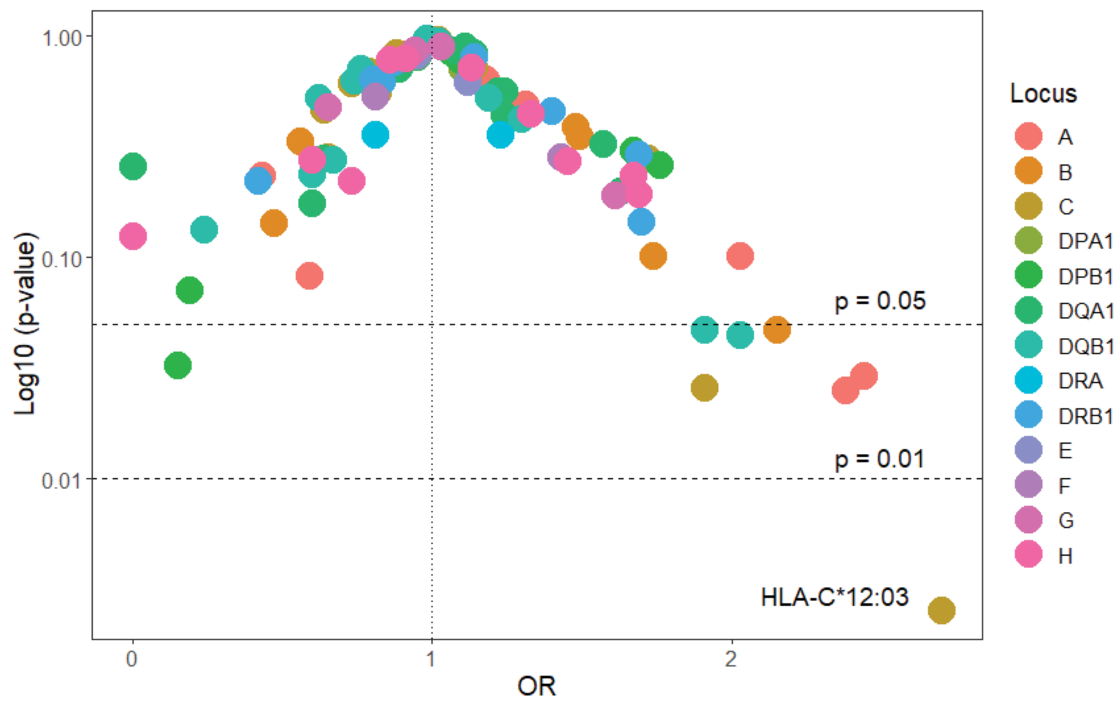


Figure 1B. Changes in amino acid positions 11 and 12 of HLA-B are associated with triple negative breast cancer.

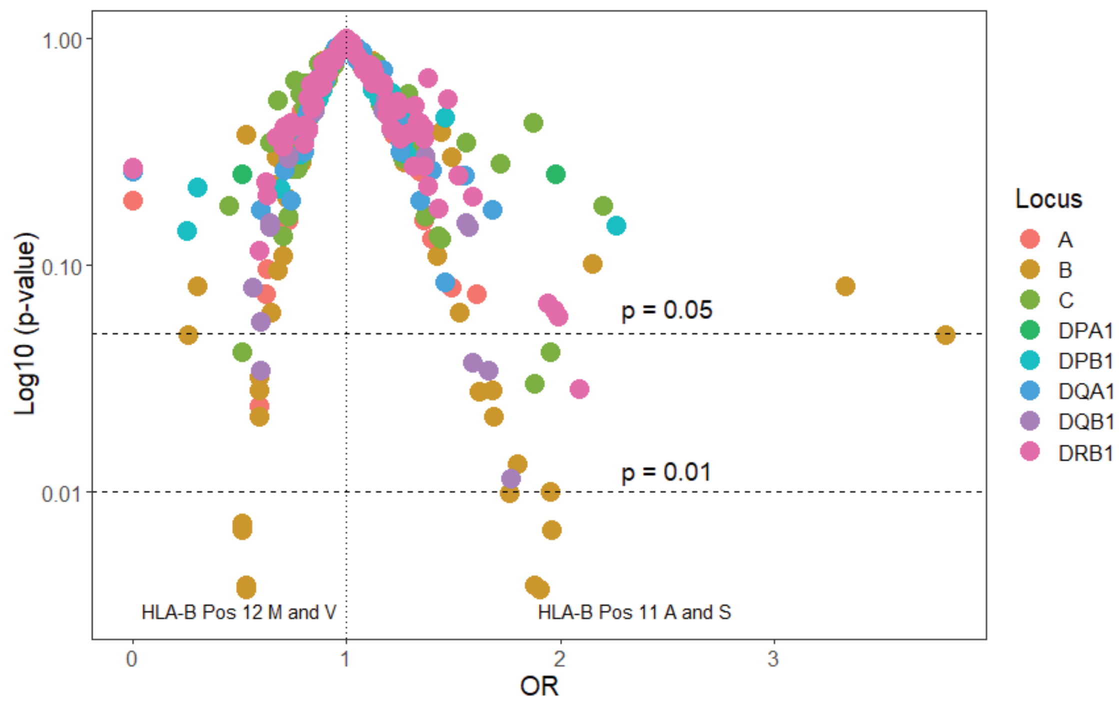
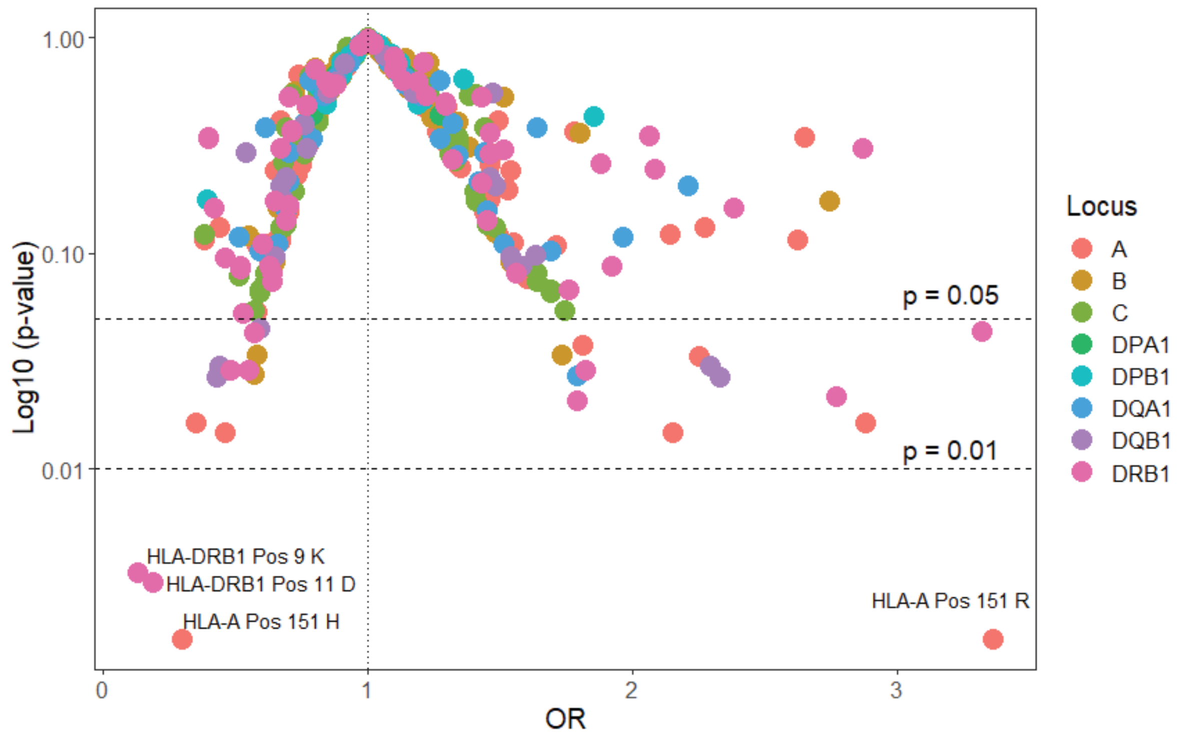


Figure 1C. *HLA-DPB1*17:01* is associated with protection gainst lymph node metastasis.

Figure 1D. Positions 9 and 11 of HLA-DRB1, and 151 of HLA-A are associated with risk and protection against grade III BC



7.CONCLUSÕES

Neste trabalho exploramos a variação de presença e ausência, haplotípica, de número de cópias e alélica de *KIR*, utilizando sequenciamento de nova geração e uma *pipeline* bioinformática projetada especificamente para resolver os obstáculos técnicos na determinação de genótipos *KIR*.

No primeiro capítulo, exploramos a diversidade de *KIR* em uma amostra populacional de 2130 indivíduos euro descendentes dos Estados Unidos. A determinação precisa do número de cópias de *KIR* nos permitiu identificar um conjunto de haplótipos de conteúdo gênico incomuns, responsáveis por 5,2% da variação estrutural. Além disso, nós mostramos que a maior parte da variação de haplótipos centroméricos e teloméricos ocorre em poucos haplótipos (cA01, cB01, cB02, tA01 e tB01). Neste estudo, *KIR2DL4* foi o gene moldura que mais variou em número de cópias (6,5% de todos os indivíduos). Além disso, observamos 250 alelos com resolução de 5 dígitos, dos quais 90 têm frequências ≥1%. Com os dados de haplótipos de conteúdo gênico e informações alélicas, nós conseguimos demonstrar o desequilíbrio de ligação existente entre genes *KIR*, dado que poderá ser usado por outros pesquisadores na hora de resolver ambiguidades alélicas em seus estudos.

Nós também encontramos padrões de sequência que eram consistentes com a presença de novos alelos em 398 (18,7%) indivíduos, e contextualizamos vários polimorfismos de nucleotídeo únicos (*SNPs*) órfãos dentro do complexo. Também identificamos uma nova variante de *KIR2DL1* (Pro151Arg), e demonstramos por dinâmica molecular que esta substituição pode afetar a interação com HLA-C. Esse foi o primeiro estudo a descrever a variação de *KIR* em alta resolução para todos os genes em uma população, no qual estabelecemos bases importantes que servirão de exemplo para futuros estudos que venham a explorar a diversidade de *KIR*.

No segundo capítulo desta tese, nós exploramos a variação de KIR e seus ligantes HLA em uma amostra de 550 paciente com câncer de mama esporádico e 747 controles. Utilizando a mesma metodologia do primeiro capítulo, nós identificamos que a presença de três cópias de *KIR2DL5* estava associada à proteção (OR = 0,21, p^{corr} = 0,027), enquanto a presença de três cópias de *KIR3DL1S1* estava associada ao risco aumentado para BC (OR = 2,44, p^{corr} =

0,009). Também identificamos que o par KIR2DL3*001+HLA-C1 estava significativamente reduzido em pacientes com o subtipo HER2+ não luminal (ER e PR negativo) em comparação com todos os outros subtipos (OR = 0,23, p^{corr} = 0,016), além de uma forte associação de KIR2DS1*002+HLA-C2 em pacientes com tumores triplo negativos em comparação com todos os outros subtipos (OR = 3,34, p^{corr} < 1,4 x10-4). Por fim, observamos um forte efeito protetor do par KIR2DL1*001+HLA-C2, significativamente aumentado em pacientes com BC de grau I em comparação com os graus mais agressivos II+III (OR = 0,25, p^{corr} = 0,038).

No terceiro e último capítulo, nós exploramos a diversidade alélica de HLA classe I e classe II em BC e seus subtipos. Nós encontramos o alelo *HLA-C*12:03* associado à BC triplo negativo, bem como variações em amino ácidos nas posições 11 e 12 de HLA-B. *HLA-C*12:03* já havia sido associado à BC em uma meta-análise realizada com dados de mulheres Asiáticas. *HLA-DPB1*17:01* foi associado à proteção ao desenvolvimento de metástases em linfonodos. E por fim, a posição 151 de HLA-A estava associada ao desenvolvimento de câncer de mama grau 3, e as posições 9 e 11 de HLA-DRB1 estavam associadas à proteção contra o câncer de mama de grau 3.

Nenhum estudo anterior explorou totalmente a gama de variação estrutural e de sequência de *KIR* como apresentamos neste trabalho. Demonstramos que a junção do sequenciamento de alto rendimento com ferramentas computacionais de última geração permite a exploração de todos os aspectos da variação de *KIR,* incluindo a determinação da diversidade de haplótipos em nível populacional, melhorando a compreensão dessa complexa família gênica e fornecendo uma referência importante para estudos futuros.

REFERÊNCIAS

1. Turnbull C, Rahman N. Genetic Predisposition to Breast Cancer: Past, Present, and Future. Annual Review of Genomics and Human Genetics. 2008 Sep;9(1):321–45.

2. Parham P. Taking license with natural killer cell maturation and repertoire development. Immunological reviews. 2006 Dec;214:155–60.

3. Uhrberg M, Parham P, Wernet P. Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. Immunogenetics. 2002 Jul;54(4):221–9.

4. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. Nucleic Acids Res. 2020 Jan 8;48(D1):D948–55.

5. Khakoo SI, Carrington M. KIR and disease: A model system or system of models? Immunological Reviews. 2006;214(1):186–201.

6. Augusto DG. The Impact of KIR Polymorphism on the Risk of Developing Cancer: Not as Strong as Imagined? Frontiers in Genetics. 2016 Jun 28;7:121.

7. Ozturk OG, Gun FD, Polat G. Killer cell immunoglobulin-like receptor genes in patients with breast cancer. Medical Oncology. 2012 Jun 9;29(2):511–5.

8. Jobim MR, Jobim M, Salim PH, Portela P, Jobim LF, Leistner-Segal S, et al. Analysis of KIR gene frequencies and HLA class I genotypes in breast cancer and control group. Human Immunology. 2013;74(9):1130–3.

9. Alomar SY, Alkhuriji A, Trayhyrn P, Alhetheel A, Al-jurayyan A, Mansour L. Association of the genetic diversity of killer cell immunoglobulin-like receptor genes and HLA-C ligand in Saudi women with breast cancer. Immunogenetics. 2016;

10. Hematian Larki M, Barani S, Talei AR, Ghaderi A. Diversity of KIRs in invasive breast cancer patients and healthy controls along with the clinical significance in ER/PR/HER2+ patients. Genes Immun. 2020 Dec;21(6–8):380–9.

11. Ashouri E, Rajalingam K, Barani S, Farjadian S, Ghaderi A, Rajalingam R. Coexistence of inhibitory and activating killer-cell immunoglobulin-like receptors to the same cognate HLA-C2 and Bw4 ligands confer breast cancer risk. Sci Rep. 2021 Apr 12;11(1):7932.

12. Li J, Chen Z, Su K, Zeng J. Clinicopathological classification and traditional prognostic indicators of breast cancer. 2015;8(January 2013):8500–5.

13. Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis C a, et al. The global breast cancer burden: variations in epidemiology and survival. Clinical breast cancer. 2005;6(5):391–401.

14. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. SEER Cancer Statistics Review, 1975-2018 [Internet]. National Cancer Institute. Bethesda; 2020 [cited 2022 Mar 21]. Available from: https://seer.cancer.gov/csr/1975_2018/index.html

15. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022 Jan;72(1):7–33.

16. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021;71(3):209–49.

17. American Cancer Society. Cancer Facts and Figures 2022. Atlanta, GA: American Cancer Society; 2022.

18. Mamessier E, Sylvain A, Thibult M laure, Houvenaeghel G, Jacquemier J, Castellano R, et al. Human breast cancer cells enhance self tolerance by promoting evasion from NK cell antitumour immunity. Jnci. 2011;121(9):3609–22.

19. Kenemans P, Verstraeten RA, Verheijen RHM. Oncogenic pathways in hereditary and sporadic breast cancer. Maturitas. 2004 Sep 24;49(1):34–43.

20. Wooster R, Weber BL. Breast and Ovarian Cancer. New England Journal of Medicine. 2003 Jun 5;348(23):2339–47.

21. PDQ® Adult Treatment Editorial Board. Breast Cancer Treatment (Adult) (PDQ®)–Health Professional Version - NCI [Internet]. National Cancer Institute; 2022 [cited 2022 Jul 22]. Available from: https://www.cancer.gov/types/breast/hp/breast-treatment-pdq

22. Colditz GA, Kaphingst KA, Hankinson SE, Rosner B. Family history and risk of breast cancer: nurses' health study. Breast Cancer Res Treat. 2012 Jun;133(3):1097–104.

23. Wolin KY, Carson K, Colditz GA. Obesity and cancer. Oncologist. 2010;15(6):556–65.

24. Kaaks R, Berrino F, Key T, Rinaldi S, Dossus L, Biessy C, et al. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). J Natl Cancer Inst. 2005 May 18;97(10):755–65.

25. Endogenous Hormones and Breast Cancer Collaborative Group, Key TJ, Appleby PN, Reeves GK, Roddam AW, Helzlsouer KJ, et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. Br J Cancer. 2011 Aug 23;105(5):709–22.

26. Ritte R, Lukanova A, Tjønneland A, Olsen A, Overvad K, Mesrine S, et al. Height, age at menarche and risk of hormone receptor-positive and -negative breast cancer: a cohort study. Int J Cancer. 2013 Jun 1;132(11):2619–29.

27. Kotsopoulos J, Chen WY, Gates MA, Tworoger SS, Hankinson SE, Rosner BA. Risk factors for ductal and lobular breast cancer: results from the nurses' health study. Breast Cancer Res. 2010;12(6):R106.

28. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022 Jan;12(1):31–46.

29. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009 May 20;101(10):736–50.

30. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science (New York, NY). 1987 Jan 9;235(4785):177–82.

31. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011 Aug;22(8):1736–47.

32. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013 Sep;24(9):2206–23.

33. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science (New York, NY). 1989 May 12;244(4905):707–12.

34. Wu G, Xing M, Mambo E, Huang X, Liu J, Guo Z, et al. Somatic mutation and gain of copy number of PIK3CA in human breast cancer. Breast Cancer Res. 2005;7(5):R609-616.

35. Dubik D, Dembinski TC, Shiu RP. Stimulation of c-myc oncogene expression associated with estrogen-induced proliferation of human breast cancer cells. Cancer Res. 1987 Dec 15;47(24 Pt 1):6517–21.

36. Park KJ, Krishnan V, O'Malley BW, Yamamoto Y, Gaynor RB. Formation of an IKKalpha-dependent transcription complex is required for estrogen receptor-mediated gene activation. Mol Cell. 2005 Apr 1;18(1):71–82.

37. Gillett C, Fantl V, Smith R, Fisher C, Bartek J, Dickson C, et al. Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. Cancer Res. 1994 Apr 1;54(7):1812–7.

38. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. The oncologist. 2004;9(4):361–77.

39. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994 Oct 7;266(5182):66–71.

40. Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, et al. BRCA1 mutations in primary breast and ovarian carcinomas. Science. 1994 Oct 7;266(5182):120–2.

41. Szabo CI, King MC. Population genetics of BRCA1 and BRCA2. Am J Hum Genet. 1997 May;60(5):1013–20.

42. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995 Dec 21;378(6559):789–92.

43. Marsh DJ, Coulon V, Lunetta KL, Rocca-Serra P, Dahia PL, Zheng Z, et al. Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. Hum Mol Genet. 1998 Mar;7(3):507–15.

44. Børresen-Dale AL. TP53 and breast cancer. Hum Mutat. 2003 Mar;21(3):292–300.

45. Afzal S, Hassan M, Ullah S, Abbas H, Tawakkal F, Khan MA. Breast Cancer; Discovery of Novel Diagnostic Biomarkers, Drug Resistance, and Therapeutic Implications. Front Mol Biosci. 2022;9:783450.

46. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012 Oct 4;490(7418):61–70.

47. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer. 1957 Sep;11(3):359–77.

48. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991 Nov;19(5):403–10.

49. Argani P, Cimino-Mathews A. Grade of Breast Cancer [Internet]. Johns Hopkins Medicine. 2015 [cited 2016 Aug 20]. Available from: http://pathology.jhu.edu/breast/grade.php

50. Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. J Clin Oncol. 2008 Jul 1;26(19):3153–8.

51. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000 Aug;406(6797):747–52.

52. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences. 2001 Sep 11;98(19):10869–74.

53. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res. 2004 Aug 15;10(16):5367–74.

54. Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med. 2010 May 25;7(5):e1000279.

55. Gao JJ, Swain SM. Luminal A Breast Cancer and Molecular Assays: A Review. Oncologist. 2018 May;23(5):556–65.

56. Prat A, Cheang MCU, Martín M, Parker JS, Carrasco E, Caballero R, et al. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. J Clin Oncol. 2013 Jan 10;31(2):203–9.

57. Li Z, Wei H, Li S, Wu P, Mao X. The Role of Progesterone Receptors in Breast Cancer. Drug Des Devel Ther. 2022;16:305–14.

58. Citron ML, Berry DA, Cirrincione C, Hudis C, Winer EP, Gradishar WJ, et al. Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. J Clin Oncol. 2003 Apr 15;21(8):1431–9.

59. Collins DM, Conlon NT, Kannan S, Verma CS, Eli LD, Lalani AS, et al. Preclinical Characteristics of the Irreversible Pan-HER Kinase Inhibitor Neratinib Compared with Lapatinib: Implications for the Treatment of HER2-Positive and HER2-Mutated Breast Cancer. Cancers (Basel). 2019 May 28;11(6):E737.

60. Kulukian A, Lee P, Taylor J, Rosler R, de Vries P, Watson D, et al. Preclinical Activity of HER2-Selective Tyrosine Kinase Inhibitor Tucatinib as a Single Agent or in Combination with Trastuzumab or Docetaxel in Solid Tumor Models. Mol Cancer Ther. 2020 Apr;19(4):976–87.

61. Martin M, Holmes FA, Ejlertsen B, Delaloge S, Moy B, Iwata H, et al. Neratinib after trastuzumab-based adjuvant therapy in HER2-positive breast cancer (ExteNET): 5-year analysis of a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2017 Dec;18(12):1688–700.

62. Modi S, Park H, Murthy RK, Iwata H, Tamura K, Tsurutani J, et al. Antitumor Activity and Safety of Trastuzumab Deruxtecan in Patients With HER2-Low-Expressing Advanced Breast Cancer: Results From a Phase Ib Study. J Clin Oncol. 2020 Jun 10;38(17):1887–96.

63. Nordstrom JL, Gorlatov S, Zhang W, Yang Y, Huang L, Burke S, et al. Anti-tumor activity and toxicokinetics analysis of MGAH22, an anti-HER2 monoclonal antibody with enhanced Fcγ receptor binding properties. Breast Cancer Res. 2011;13(6):R123.

64. Anders C, Carey LA. Understanding and treating triple-negative breast cancer. Oncology (Williston Park, NY). 2008 Oct;22(11):1233–9; discussion 1239-40, 1243.

65. Hudis CA, Gianni L. Triple-negative breast cancer: an unmet medical need. The oncologist. 2011;16 Suppl 1:1–11.

66. Weigelt B, Peterse JL, van't Veer LJ. Breast cancer metastasis: markers and models. Nature Reviews Cancer. 2005 Aug;5(8):591–602.

67. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S. Functions of natural killer cells. Nature Immunology. 2008 May;9(5):503–10.

68. Robertson MJ, Ritz J. Biology and clinical relevance of human natural killer cells. Blood. 1990 Dec 15;76(12):2421–38.

69. Sun JC, Lanier LL. Natural killer cells remember: an evolutionary bridge between innate and adaptive immunity? European journal of immunology. 2009 Aug;39(8):2059–64.

70. Vivier E, Raulet DH, Moretta A, Caligiuri MA, Zitvogel L, Lanier LL, et al. Innate or adaptive immunity? The example of natural killer cells. Science (New York, NY). 2011 Jan 7;331(6013):44–9.

71. Waggoner SN, Reighard SD, Gyurova IE, Cranert SA, Mahl SE, Karmele EP, et al. Roles of natural killer cells in antiviral immunity. Curr Opin Virol. 2016 Feb;16:15–23.

72. Moffett A, Colucci F. Uterine NK cells: active regulators at the maternal-fetal interface. J Clin Invest. 2014 May;124(5):1872–9.

73. Bulmer JN, Lash GE. The Role of Uterine NK Cells in Normal Reproduction and Reproductive Disorders. Adv Exp Med Biol. 2015;868:95–126.

74. Kiessling R, Klein E, Wigzell H. „Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. European Journal of Immunology. 1975 Feb;5(2):112–7.

75. Herberman RB, Holden HT. Natural cell-mediated immunity. Adv Cancer Res. 1978;27:305–77.

76. Herberman RB, Ortaldo JR. Natural killer cells: their roles in defenses against disease. Science. 1981 Oct 2;214(4516):24–30.

77. Agaugué S, Marcenaro E, Ferranti B, Moretta L, Moretta A. Human natural killer cells exposed to IL-2, IL-12, IL-18, or IL-4 differently modulate priming of naive T cells by monocyte-derived dendritic cells. Blood. 2008 Sep 1;112(5):1776–83.

78. Moretta A. The dialogue between human natural killer cells and dendritic cells. Curr Opin Immunol. 2005 Jun;17(3):306–11.

79. Vitale M, Della Chiesa M, Carlomagno S, Pende D, Aricò M, Moretta L, et al. NK-dependent DC maturation is mediated by TNFalpha and IFNgamma released upon engagement of the NKp30 triggering receptor. Blood. 2005 Jul 15;106(2):566–71.

80. Mattiola I, Pesant M, Tentorio PF, Molgora M, Marcenaro E, Lugli E, et al. Priming of Human Resting NK Cells by Autologous M1 Macrophages via the Engagement of IL-1β, IFN-β, and IL-15 Pathways. J Immunol. 2015 Sep 15;195(6):2818–28.

81. Cooper MA, Fehniger TA, Caligiuri MA, Robertson MJ, Ritz J, Lanier LL, et al. The biology of human natural killer-cell subsets. Trends in Immunology. 2001 Nov;22(11):633–40.

82. Lanier LL, Testi R, Bindl J, Phillips JH. Identity of Leu-19 (CD56) leukocyte differentiation antigen and neural cell adhesion molecule. J Exp Med. 1989 Jun 1;169(6):2233–8.

83. Dong D, Zheng L, Lin J, Zhang B, Zhu Y, Li N, et al. Structural basis of assembly of the human T cell receptor–CD3 complex. Nature. 2019 Sep;573(7775):546–52.

84. Mandelboim O, Malik P, Davis DM, Jo CH, Boyson JE, Strominger JL. Human CD16 as a lysis receptor mediating direct natural killer cell cytotoxicity. Proc Natl Acad Sci U S A. 1999 May 11;96(10):5640–4.

85. Walzer T, Dalod M, Robbins SH, Zitvogel L, Vivier E. Natural-killer cells and dendritic cells: "l'union fait la force". Blood. 2005 Oct 1;106(7):2252–8.

86. Trowsdale J. Genetic and Functional Relationships between MHC and NK Receptor Genes. Immunity. 2001;15(3):363–74.

87. Raulet DH, Guerra N. Oncogenic stress sensed by the immune system: role of natural killer cell receptors. Nature reviews Immunology. 2009 Aug;9(8):568–80.

88. Brown MG, Scalzo AA, Yokoyama WM. The NKC and regulation of natural killer cell-mediated immunity. In: Major Histocompatibility Complex. Tokyo: Springer Japan; 2000. p. 287–301.

89. Lanier LL. Up on the tightrope: natural killer cell activation and inhibition. Nat Immunol. 2008 May;9(5):495–502.

90. Gasser S, Raulet DH. Activation and self-tolerance of natural killer cells. Immunological Reviews. 2006;214(1):130–42.

91. Kärre K, Ljunggren HG, Piontek G, Kiessling R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. Nature. 1986;319(6055):675–8.

92. Kelley J, Walter L, Trowsdale J. Comparative Genomics of Natural Killer Cell Receptor Gene Clusters. PLOS Genetics. 2005 de ago de;1(2):e27.

93. Raulet DH, Vance RE. Self-tolerance of natural killer cells. Nature Reviews Immunology. 2006 Jul;6(7):520–31.

94. Vilches C, Parham P. K IR : Diverse, Rapidly Evolving Receptors of Innate and Adaptive Immunity. Annual Review of Immunology. 2002 Apr;20(1):217–51.

95. Colonna M, Samaridis J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. Science. 1995 Apr 21;268(5209):405–8.

96. Wagtmann N, Biassoni R, Cantoni C, Verdiani S, Malnati MS, Vitale M, et al. Molecular clones of the p58 NK cell receptor reveal immunoglobulin-related molecules with diversity in both the extra- and intracellular domains. Immunity. 1995 May 1;2(5):439–49.

97. Vilches C, Rajalingam R, Uhrberg M, Gardiner CM, Young NT, Parham P. KIR2DL5, a Novel Killer-Cell Receptor with a D0-D2 Configuration of Ig-Like Domains. The Journal of Immunology. 2000 Jun 1;164(11):5797–804.

98. Vilches C, Pando MJ, Parham P. Genes encoding human killer-cell Ig-like receptors with D1 and D2 extracellular domains all contain untranslated pseudoexons encoding a third Ig-like domain. Immunogenetics. 2000 Jul 1;51(8–9):639–46.

99. Kikuchi-Maki A, Yusa S ichi, Catina TL, Campbell KS. KIR2DL4 is an IL-2-regulated NK cell receptor that exhibits limited expression in humans but triggers strong IFN-gamma production. Journal of immunology (Baltimore, Md : 1950). 2003 Oct 1;171(7):3415–25.

100. Long EO. Regulation of immune responses through inhibitory receptors. Annual review of immunology. 1999;17:875–904.

101. Olcese L, Cambiaggi A, Semenzato G, Bottino C, Moretta A, Vivier E. Human killer cell activatory receptors for MHC class I molecules are included in a multimeric complex expressed by natural killer cells. J Immunol. 1997 Jun 1;158(11):5083–6.

102. Lanier LL, Corliss BC, Wu J, Leong C, Phillips JH. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. Nature. 1998 Feb 12;391(6668):703–7.

103. Bléry M, Olcese L, Vivier E. Early signaling via inhibitory and activating NK receptors. Hum Immunol. 2000 Jan;61(1):51–64.

104. Long EO, Barber DF, Burshtyn DN, Faure M, Peterson M, Rajagopalan S, et al. Inhibition of natural killer cell activation signals by killer cell immunoglobulin-like receptors (CD158). Immunol Rev. 2001 Jun;181:223–33.

105. Béziat V, Hilton HG, Norman PJ, Traherne JA. Deciphering the killer-cell immunoglobulin-like receptor system at super-resolution for natural killer and T-cell biology. Immunology. 2017 Mar;150(3):248–64.

106. Long EO, Colonna M, Lanier LL. Inhibitory MHC class I receptors on NK and T cells: a standard nomenclature. Immunology today. 1996 Feb;17(2):100.

107. Gómez-Lozano N, Gardiner CM, Parham P, Vilches C. Some human KIR haplotypes contain two KIR2DL5 genes: KIR2DL5A and KIR2DL5B. Immunogenetics. 2002 Aug;54(5):314–9.

108. Uhrberg M, Valiante NM, Shum BP, Shilling HG, Lienert-Weidenbach K, Corliss B, et al. Human Diversity in Killer Cell Inhibitory Receptor Genes. Immunity. 1997 Dec;7(6):753–63.

109. Amorim LM, Augusto DG, Nemat-Gorgani N, Montero-Martin G, Marin WM, Shams H, et al. High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity. Frontiers in Immunology [Internet]. 2021 [cited 2022 Mar 21];12. Available from: https://www.frontiersin.org/article/10.3389/fimmu.2021.674778

110. Roe D, Vierra-Green C, Pyo CW, Eng K, Hall R, Kuang R, et al. Revealing complete complex KIR haplotypes phased by long-read sequencing technology. Genes & Immunity. 2017 Sep;18(3):127–34.

111. Martin AM, Kulski JK, Gaudieri S, Witt CS, Freitas EM, Trowsdale J, et al. Comparative genomic analysis, diversity and evolution of two KIR haplotypes A and B. Gene. 2004 Jun 23;335:121–31.

112. Marsh SGE, Parham P, Dupont B, Geraghty DE, Trowsdale J, Middleton D, et al. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. Hum Immunol. 2003 Jun;64(6):648–54.

113. Maxwell LD, Wallace A, Middleton D, Curran MD. A common KIR2DS4 deletion variant in the human that predicts a soluble KIR molecule analogous to the KIR1D molecule observed in the rhesus monkey. Tissue Antigens. 2002 Sep;60(3):254–8.

114. Augusto DG, Zehnder-Alves L, Pincerati MR, Martin MP, Carrington M, Petzl-Erler ML. Diversity of the KIR gene cluster in an urban Brazilian population. Immunogenetics. 2012 Feb;64(2):143–52.

115. Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, et al. Maternal activating KIRs protect against human reproductive failure mediated by fetal HLA-C2. J Clin Invest. 2010 Nov 1;120(11):4102–10.

116. Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, et al. Plasticity in the organization and sequences of human KIR/ILT gene families. Proc Natl Acad Sci USA. 2000 Apr 25;97(9):4778.

117. Augusto DG, Amorim LM, Farias TDJ, Petzl-Erler ML. KIR and HLA genotyping of Japanese descendants from Curitiba, a city of predominantly European ancestry from Southern Brazil. Vol. 77, Human Immunology. 2016.

118. Jiang K, Zhu FM, Lv QF, Yan LX. Distribution of killer cell immunoglobulin-like receptor genes in the Chinese Han population. Tissue Antigens. 2005;65(6):556–63.

119. Whang DH, Park H, Yoon JA, Park MH. Haplotype analysis of killer cell immunoglobulin-like receptor genes in 77 Korean families. Hum Immunol. 2005 Feb;66(2):146–54.

120. Suzuki Y, Hamamoto Y, Ogasawara Y, Ishikawa K, Yoshikawa Y, Sasazuki T, et al. Genetic polymorphisms of killer cell immunoglobulin-like receptors are associated with susceptibility to psoriasis vulgaris. J Invest Dermatol. 2004 May;122(5):1133–6.

121. Yawata M, Yawata N, McQueen KL, Cheng NW, Guethlein LA, Rajalingam R, et al. Predominance of group A KIR haplotypes in Japanese associated with diverse NK cell repertoires of KIR expression. Immunogenetics. 2002 Nov;54(8):543–50.

122. Amorim LM, van Tong H, Hoan NX, Vargas L de B, Ribeiro EM de SF, Petzl-Erler ML, et al. KIR-HLA distribution in a Vietnamese population from Hanoi. Hum Immunol. 2018 Feb;79(2):93–100.

123. Chaisri S, Kitcharoen K, Romphruk AV, Romphruk A, Witt CS, Leelayuwat C. Polymorphisms of killer immunoglobulin-like receptors (KIRs) and HLA ligands in northeastern Thais. Immunogenetics. 2013 Sep 1;65(9):645–53.

124. Toneva M, Lepage V, Lafay G, Dulphy N, Busson M, Lester S, et al. Genomic diversity of natural killer cell receptor genes in three populations. Tissue antigens. 2001 Apr;57(4):358–62.

125. Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res. 2020 08;48(D1):D783–8.

126. Pyo CW, Guethlein LA, Vu Q, Wang R, Abi-Rached L, Norman PJ, et al. Different Patterns of Evolution in the Centromeric and Telomeric Regions of Group A and B

Haplotypes of the Human Killer Cell Ig-Like Receptor Locus. Matsunami H, editor. PLoS ONE. 2010 Dec 29;5(12):e15115.

127. Vierra-Green C, Roe D, Hou L, Hurley CK, Rajalingam R, Reed E, et al. Allele-Level Haplotype Frequencies and Pairwise Linkage Disequilibrium for 14 KIR Loci in 506 European-American Individuals. Tang J, editor. PLoS ONE. 2012 Nov 5;7(11):e47491.

128. Alicata C, Ashouri E, Nemat-Gorgani N, Guethlein LA, Marin WM, Tao S, et al. KIR Variation in Iranians Combines High Haplotype and Allotype Diversity With an Abundance of Functional Inhibitory Receptors. Front Immunol. 2020 Apr 2;11:556.

129. Nemat-Gorgani N, Guethlein LA, Henn BM, Norberg SJ, Chiaroni J, Sikora M, et al. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. JI. 2019 May 1;202(9):2636–47.

130. Bruijnesteijn J, de Groot NG, Bontrop RE. The Genetic Mechanisms Driving Diversification of the KIR Gene Cluster in Primates. Frontiers in Immunology [Internet]. 2020 [cited 2022 Sep 2];11. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2020.582804

131. Martin MP, Bashirova A, Traherne J, Trowsdale J, Carrington M. Cutting edge: expansion of the KIR locus by unequal crossing over. Journal of immunology (Baltimore, Md : 1950). 2003 Sep 1;171(5):2192–5.

132. Jiang W, Johnson C, Jayaraman J, Simecek N, Noble J, Moffatt MF, et al. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. Genome Res. 2012 Oct;22(10):1845–54.

133. Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, et al. Meiotic recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. Genome Res. 2009 May;19(5):757–69.

134. Parham P. Killer cell immunoglobulin-like receptor diversity: Balancing signals in the natural killer cell response. Immunology Letters. 2004;92(1–2):11–3.

135. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. Am J Hum Genet. 2016 Aug 4;99(2):375–91.

136. Norman PJ, Chandanayingyong D, Vaughan RW, Verity DH, Stephens HAF. Distribution of natural killer cell immunoglobulin-like receptor sequences in three ethnic groups. Immunogenetics. 2001 Jan 24;52(3–4):195–205.

137. Sun JY, Gaidulis L, Miller MM, Goto RM, Rodriguez R, Forman SJ, et al. Development of a multiplex PCR-SSP method for Killer-cell immunoglobulin-like receptor genotyping. Tissue Antigens. 2004 Oct;64(4):462–8.

138. Vilches C, Castaño J, Gómez-Lozano N, Estefanía E. Facilitation of KIR genotyping by a PCR-SSP method that amplifies short DNA fragments. Tissue Antigens. 2007 Nov;70(5):415–22.

139. Martin MP, Carrington M. KIR Locus Polymorphisms. Methods in Molecular Biology. 2008;415.

140. Kulkarni S, Martin MP, Carrington M. KIR Genotyping by Multiplex PCR-SSP. In: Campbell KS, editor. Natural Killer Cell Protocols [Internet]. Totowa, NJ: Humana Press; 2010 [cited 2020 Feb 26]. p. 365–75. (Methods in Molecular Biology; vol. 612). Available from: http://link.springer.com/10.1007/978-1-60761-362-6_25

141. Ashouri E, Ghaderi A, Reed EF, Rajalingam R. A novel duplex SSP-PCR typing method for *KIR* gene profiling. Tissue Antigens. 2009 Jul;74(1):62–7.

142. Abalos AT, Eggers R, Hogan M, Nielson CM, Giuliano AR, Harris RB, et al. Design and validation of a multiplex specific primer-directed polymerase chain reaction assay for killer-cell immunoglobulin-like receptor genetic profiling. Tissue Antigens. 2011 Feb;77(2):143–8.

143. Crum KA, Logue SE, Curran MD, Middleton D. Development of a PCR-SSOP approach capable of defining the natural killer cell inhibitory receptor (KIR) gene sequence repertoires. Tissue Antigens. 2000 Oct;56(4):313–26.

144. Nong T, Saito K, Blair L, Tarsitani C, Lee JH. KIR genotyping by reverse sequence-specific oligonucleotide methodology. Tissue Antigens. 2007 Apr;69:92–5.

145. Koehler RN, Walsh AM, Moqueet N, Currier JR, Eller MA, Eller LA, et al. High-throughput genotyping of KIR2DL2/L3, KIR3DL1/S1, and their HLA class I ligands using real-time PCR. Tissue Antigens. 2009 Jul;74(1):73–80.

146. Alves LGT, Rajalingam R, Canavez F. A novel real-time PCR method for KIR genotyping. Tissue Antigens. 2009 Feb;73(2):188–91.

147. Hong HA, Loubser AS, de Assis Rosa D, Naranbhai V, Carr W, Paximadis M, et al. Killer-cell immunoglobulin-like receptor genotyping and HLA killer-cell immunoglobulin-like receptor-ligand identification by real-time polymerase chain reaction. Tissue Antigens. 2011 Sep;78(3):185–94.

148. Amorim LM, Santos THS, Hollenbach JA, Norman PJ, Marin WM, Dandekar R, et al. Cost-effective and fast *KIR* gene-content genotyping by multiplex melting curve analysis. HLA. 2018 Dec;92(6):384–91.

149. Houtchens KA, Nichols RJ, Ladner MB, Boal HE, Sollars C, Geraghty DE, et al. High-throughput killer cell immunoglobulin-like receptor genotyping by MALDI-TOF mass spectrometry with discovery of novel alleles. Immunogenetics. 2007 May 30;59(7):525–37.

150. Jiang W, Johnson C, Simecek N, López-Álvarez MR, Di D, Trowsdale J, et al. qKAT: a high-throughput qPCR method for KIR gene copy number and haplotype determination. Genome Med. 2016 Dec;8(1):99.

151. Marin W, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, et al. High-throughput Interpretation of Killer-cell Immunoglobulin-like Receptor Short-read Sequencing Data with PING. bioRxiv. 2021 Mar 24;2021.03.24.436770.

152. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. Nat Rev Genet. 2004 Dec;5(12):889–99.

153. Campbell RD, Trowsdale J. Map of the human MHC. Immunology Today. 1993;14(7):349–52.

154. Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J. Advances in the Evolutionary Understanding of MHC Polymorphism. Trends Genet. 2020 Apr;36(4):298–311.

155. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. J Immunol. 2005 Jan 1;174(1):6–19.

156. Klein J, Sato A. The HLA system. First of two parts. N Engl J Med. 2000 Sep 7;343(10):702–9.

157. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature. 1993 Jul 1;364(6432):33–9.

158. Müller CA, Engler-Blum G, Gekeler V, Steiert I, Weiss E, Schmidt H. Genetic and serological heterogeneity of the supertypic HLA-B locus specificities Bw4 and Bw6. Immunogenetics. 1989;30(3):200–7.

159. Gumperz JE, Barber LD, Valiante NM, Percival L, Phillips JH, Lanier LL, et al. Conserved and variable residues within the Bw4 motif of HLA-B make separable contributions to recognition by the NKB1 killer cell-inhibitory receptor. Journal of immunology (Baltimore, Md : 1950). 1997 Jun 1;158(11):5237–41.

160. Rauch A, Nolan D, Furrer H, McKinnon E, John M, Mallal S, et al. *HLA-Bw4* Homozygosity Is Associated with an Impaired CD4 T Cell Recovery after Initiation of Antiretroviral Therapy. Clinical Infectious Diseases. 2008 Jun 15;46(12):1921–5.

161. Kostyu DD, Cresswell P, Amos DB. A public HLA antigen associated with HLA-A9, Aw32, and Bw4. Immunogenetics. 1980 Oct;10(5):433–42.

162. Stern M, Ruggeri L, Capanni M, Mancusi A, Velardi A, Valiante NM, et al. Human leukocyte antigens A23, A24, and A32 but not A25 are ligands for KIR3DL1. Blood. 2008 Aug 1;112(3):708–10.

163. Winter CC, Gumperz JE, Parham P, Long EO, Wagtmann N. Direct binding and functional transfer of NK cell inhibitory receptors reveal novel patterns of HLA-C

allotype recognition. Journal of immunology (Baltimore, Md : 1950). 1998 Jul 15;161(2):571–7.

164. Cella M, Longo A, Ferrara GB, Strominger JL, Colonna M. NK3-specific natural killer cells are selectively inhibited by Bw4-positive HLA alleles with isoleucine 80. The Journal of experimental medicine. 1994 Oct 1;180(4):1235–42.

165. Liu J, Xiao Z, Ko HL, Shen M, Ren EC. Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A*11. Proceedings of the National Academy of Sciences. 2014 Feb 18;111(7):2662–7.

166. Moesta AK, Norman PJ, Yawata M, Yawata N, Gleimer M, Parham P. Synergistic polymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C than KIR2DL3. J Immunol. 2008 Mar 15;180(6):3969–79.

167. Hansasuta P, Dong T, Thananchai H, Weekes M, Willberg C, Aldemir H, et al. Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. European Journal of Immunology. 2004 Jun;34(6):1673–9.

168. Kollnberger S, Chan A, Sun MY, Chen LY, Wright C, di Gleria K, et al. Interaction of HLA-B27 homodimers with KIR3DL1 and KIR3DL2, unlike HLA-B27 heterotrimers, is independent of the sequence of bound peptide. Eur J Immunol. 2007 May;37(5):1313–22.

169. Carr WH, Pando MJ, Parham P. KIR3DL1 polymorphisms that affect NK cell inhibition by HLA-Bw4 ligand. J Immunol. 2005 Oct 15;175(8):5222–9.

170. O'Connor GM, Guinan KJ, Cunningham RT, Middleton D, Parham P, Gardiner CM. Functional polymorphism of the KIR3DL1/S1 receptor on human NK cells. Journal of immunology (Baltimore, Md : 1950). 2007 Jan 1;178(1):235–41.

171. Carr WH, Rosen DB, Arase H, Nixon DF, Michaelsson J, Lanier LL. Cutting Edge: KIR3DS1, a gene implicated in resistance to progression to AIDS, encodes a DAP12-associated receptor expressed on NK cells that triggers NK cell activation. Journal of immunology (Baltimore, Md : 1950). 2007 Jan 15;178(2):647–51.

172. Gillespie GMA, Bashirova A, Dong T, McVicar DW, Rowland-Jones SL, Carrington M. Lack of KIR3DS1 Binding to MHC Class I Bw4 Tetramers in Complex with CD8 $^{+}$ T Cell Epitopes. AIDS Research and Human Retroviruses. 2007 Mar;23(3):451–5.

173. O'Connor GM, Yamada E, Rampersaud A, Thomas R, Carrington M, McVicar DW. Analysis of binding of KIR3DS1*014 to HLA suggests distinct evolutionary history of KIR3DS1. Journal of immunology (Baltimore, Md : 1950). 2011 Sep 1;187(5):2162–71.

174. Burian A, Wang KL, Finton KAK, Lee N, Ishitani A, Strong RK, et al. HLA-F and MHC-I Open Conformers Bind Natural Killer Cell Ig-Like Receptor KIR3DS1. PloS one. 2016 Jan 20;11(9):e0163297.

175. Stewart CA, Laugier-Anfossi F, Vély F, Saulquin X, Riedmuller J, Tisserant A, et al. Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. Proc Natl Acad Sci U S A. 2005 Sep 13;102(37):13224–9.

176. Graef T, Moesta AK, Norman PJ, Abi-Rached L, Vago L, Older Aguilar AM, et al. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. The Journal of experimental medicine. 2009 Oct 26;206(11):2557–72.

177. Augusto DG, Petzl-Erler ML. KIR and HLA under pressure: evidences of coevolution across worldwide populations. Human Genetics. 2015;134(9):929–40.

178. Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF, Petzl-Erler ML. Activating KIR and HLA Bw4 Ligands Are Associated to Decreased Susceptibility to Pemphigus Foliaceus, an Autoimmune Blistering Skin Disease. Zimmer J, editor. PLoS ONE. 2012 Jul 2;7(7):e39991.

179. Augusto DG, O'Connor GM, Lobo-Alves SC, Bass S, Martin MP, Carrington M, et al. Pemphigus is associated with *KIR3DL2* expression levels and provides evidence that KIR3DL2 may bind HLA-A3 and A11 in vivo. European Journal of Immunology. 2015 Jul;45(7):2052–60.

180. Beltrame LM, Sell AM, Moliterno RA, Clementino SL, Cardozo DM, Dalalio MM, et al. Influence of *KIR* genes and their HLA ligands in susceptibility to dengue in a population from southern Brazil. Tissue Antigens. 2013 Dec;82(6):397–404.

181. Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR, et al. Global diversity and evidence for coevolution of KIR and HLA. Nature Genetics. 2007 Sep 12;39(9):1114–9.

182. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, et al. Co-evolution of human leukocyte antigen (HLA) class I ligands with killer-cell immunoglobulin-like receptors (KIR) in a genetically diverse population of sub-Saharan Africans. PLoS genetics. 2013 Oct;9(10):e1003938.

183. Hilton H, Norman P, Nemat-Gorgani N, Goyos A, Gignoux C, Mountain J, et al. LBP10: CO-EVOLUTION OF KIR AND HLA CLASS I IN A SOUTHERN AFRICAN HUNTER-GATHERER POPULATION. Human Immunology. 2015;76(4).

184. Trowsdale J, Moffett A. NK receptor interactions with MHC class I molecules in pregnancy. Seminars in Immunology. 2008 Dec 1;20(6):317–20.

185. Hao L, Nei M. Rapid expansion of killer cell immunoglobulin-like receptor genes in primates and their coevolution with MHC Class I genes. Gene. 2005 Mar 14;347(2):149–59.

186. Aguilar AMO, Guethlein LA, Adams EJ, Abi-Rached L, Moesta AK, Parham P. Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major

Variable Regulators of Human NK Cells. The Journal of Immunology. 2010 Oct 1;185(7):4238–51.

187. Abi-Rached L, Parham P. Natural selection drives recurrent formation of activating killer cell immunoglobulin-like receptor and Ly49 from inhibitory homologues. J Exp Med. 2005 Apr 18;201(8):1319–32.

188. Hollenbach JA, Augusto DG, Alaez C, Bubnova L, Fae I, Fischer G, et al. 16(th) IHIW: population global distribution of killer immunoglobulin-like receptor (KIR) and ligands. Int J Immunogenet. 2013 Feb;40(1):39–45.

189. Hiby SE, Walker JJ, O'Shaughnessy KM, Redman CWG, Carrington M, Trowsdale J, et al. Combinations of Maternal KIR and Fetal HLA-C Genes Influence the Risk of Preeclampsia and Reproductive Success. Journal of Experimental Medicine. 2004 Oct 11;200(8):957–65.

190. de Brito Vargas L, Beltrame MH, Ho B, Marin WM, Dandekar R, Montero-Martín G, et al. Remarkably Low KIR and HLA Diversity in Amerindians Reveals Signatures of Strong Purifying Selection Shaping the Centromeric KIR Region. Mol Biol Evol. 2021 Oct 11;39(1):msab298.

191. Gendzekhadze K, Norman PJ, Abi-Rached L, Graef T, Moesta AK, Layrisse Z, et al. Co-evolution of KIR2DL3 with HLA-C in a human population retaining minimal essential diversity of KIR and HLA class I ligands. PNAS. 2009 Nov 3;106(44):18692–7.

192. Dewan MdZ, Terunuma H, Ahmed S, Ohba K, Takada M, Tanaka Y, et al. Natural killer cells in breast cancer cell growth and metastasis in SCID mice. Biomedicine & Pharmacotherapy. 2005;59:S375–9.

193. Boyton RJ, Altmann DM. Natural killer cells, killer immunoglobulin-like receptors and human leucocyte antigen class I in disease. Clinical and experimental immunology. 2007 Jul;149(1):1–8.

194. Goedert JJ, Martin MP, Vitale F, Lauria C, Whitby D, Qi Y, et al. Risk of Classic Kaposi Sarcoma With Combinations of Killer Immunoglobulin-Like Receptor and Human Leukocyte Antigen Loci: A Population-Based Case-control Study. The Journal of infectious diseases. 2016 Feb 1;213(3):432–8.

195. Middleton D, Vilchez JR, Cabrera T, Meenagh A, Williams F, Halfpenny I, et al. Analysis of KIR gene frequencies in HLA class I characterised bladder, colorectal and laryngeal tumours. Tissue Antigens. 2007 Mar;69(3):220–6.

196. Kim HJ, Choi HB, Jang JP, Baek IC, Choi EJ, Park M, et al. HLA-Cw polypmorphism and killer cell immunoglobulin-like receptor (KIR) gene analysis in Korean colorectal cancer patients. International Journal of Surgery. 2014;12(8):815–20.

197. Hollenbach JA, Norman PJ, Creary LE, Damotte V, Montero-Martin G, Caillier S, et al. A specific amino acid motif of HLA-DRB1 mediates risk and interacts with

smoking history in Parkinson’s disease. Proc Natl Acad Sci USA. 2019 09;116(15):7419–24.

198. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012 Oct 1;28(19):2537–9.

199. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010 May;10(3):564–7.

200. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. PNAS. 1977 Dec 1;74(12):5463–7.

201. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, et al. Co-evolution of Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. PLoS Genet [Internet]. 2013 Oct 31 [cited 2020 Sep 29];9(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3814319/

202. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. Human Immunology. 2016 Mar 1;77(3):283–7.

203. IBM Corp. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp; 2020.

204. Mantel N, Haenszel W. Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. JNCI: Journal of the National Cancer Institute. 1959 Apr 1;22(4):719–48.

205. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. Seeber; 1936. 62 p.