



# Universidade: presente!



XXXI SIC

21. 25. OUTUBRO • CAMPUS DO VALE

## A Genética da Maratona: Diferenciação Genética em Etnias do Leste da África e sua Relação com o Sucesso na Corrida de Longa Distância

André Luís da Silva Zani

Orientador: Nelson Jurandi Rosa Fagundes

### Introdução

Atletas de elite competindo em provas de longa distância conquistam feitos impressionantes, tais como percorrer 42,195 km em pouco mais de duas horas, no caso da maratona. Desde a década de 1960, atletas do Leste da África, principalmente do Quênia e da Etiópia, têm dominado tais provas. Estudos demográficos mostram que, em cada um desses países, duas etnias são super-representadas entre os corredores de resistência de elite: os Kalenjin, no Quênia, e os Oromo, na Etiópia. Apesar de tanto fatores biológicos como culturais poderem ter um papel em explicar tal sucesso, nenhuma explicação simples surgiu ao se olhar para a história de vida ou loci classicamente associados a performance atlética. O objetivo deste estudo é contribuir para um melhor entendimento dos processos moleculares e fisiológicos, que, em última instância, resultam em sucesso na corrida de longa distância.

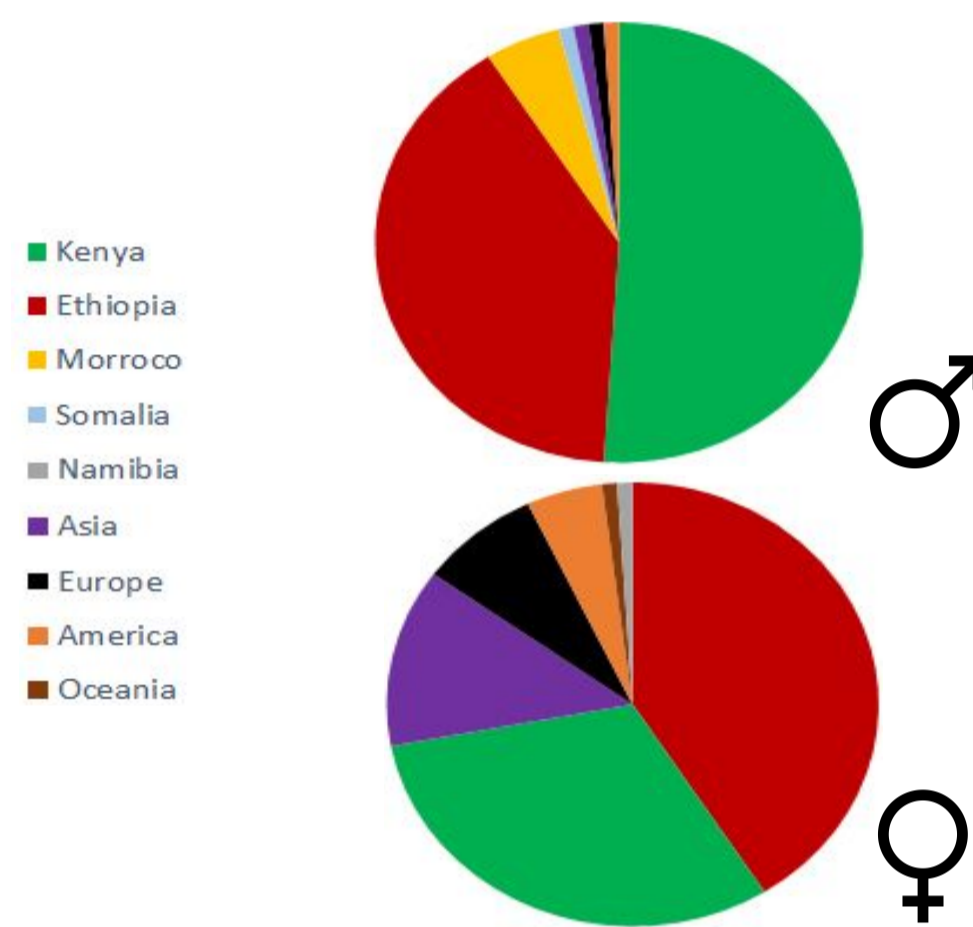


Fig. 1. Gráfico de pizza dos 100 melhores maratonistas de todos os tempos, homens e mulheres, por local de nascimento. Note a grande fração correspondente ao Quênia e a Etiópia, em ambos os sexos.

### Métodos

Nós começamos por hipotetizar que múltiplos fatores genéticos predisõem indivíduos das etnias Kalenjin e Oromo a serem a super-representados entre atletas de elite de provas de corrida de longa distância. A seguir, procuramos os picos genômicos de diferenciação genética nessas duas etnias utilizando dados públicos de frequência alélica para 1.152.000 polimorfismos de nucleotídeo único (SNPs, em inglês) obtidos do Projeto de Variação Genômica Africana (African Genome Variation Project - AGVP) e do Projeto 1000 Genomas (1000 Genomes Project). Para cada população focal, nós varremos o genoma utilizando a versão normalizada da estatística de tamanho de ramo populacional (Population Branch Statistics - PBS), o PBSn, em janelas de 20 SNPs com uma sobreposição 15 SNPs entre as janelas. O cálculo do PBSn requer duas populações de referência além das populações focais. Nós usamos os Luhya, do Quênia, e os Amhara, da Etiópia, como contrapartidas locais, e duas populações compostas representando o Oeste da África e a Eurásia, como populações distantemente relacionadas. Para cada comparação, nós mantivemos para análises posteriores o 0.1% das janelas com o maior valor de PBSn, e anotamos os genes associados a essas janelas em uma vizinhança de 5 kbp. As listas completas de genes foram submetidas a análise de conjuntos de genes (Gene-set analysis) através da plataforma Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) para identificar os processos biológicos (categoria GO Biological Processes) e fenótipos (categoria GWAS) que possam estar associados à corrida de resistência.

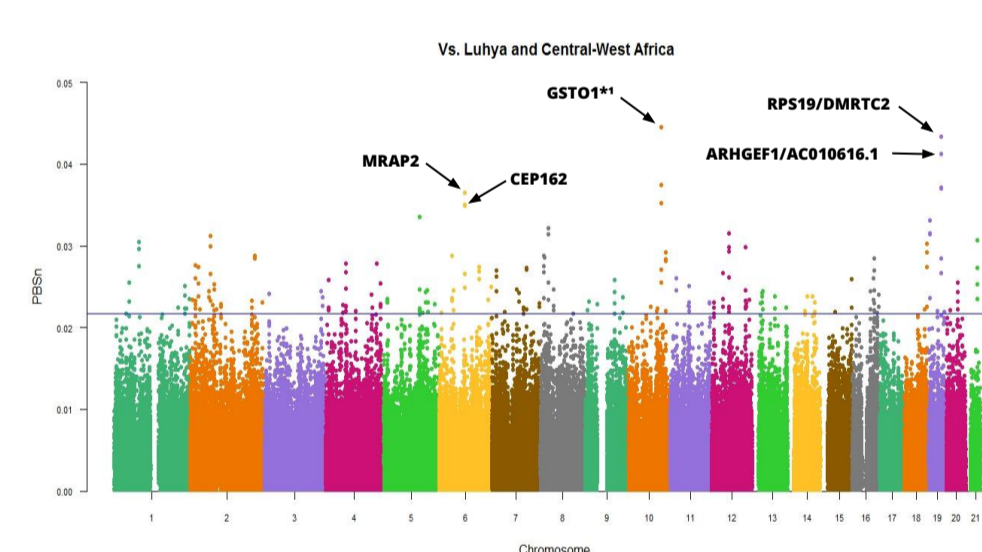
### Resultados e Discussão

Nossos resultados compreendem diversos fenótipos e processos, associados com coisas como antropometria e sistemas circulatório e respiratório. Algo marcante é a predominância de características e processos que envolvem cálcio, especialmente em Oromo, mas também em Kalenjin. Para essa segunda população, um dos maiores picos de PBSn envolve o gene *GSTO1*, o qual codifica para uma proteína que modula a atividade dos Receptores de Rianodina 1 and 2 (*RYR1* e *RYR2*), canais de cálcio do retículo sarcoplasmático, responsáveis pela contração do músculo esquelético e cardíaco, respectivamente. Curiosamente, o *RYR2* aparece na lista de genes mais diferenciados de Oromo.

Oromo			
Categoria GWAS catalog	Número de sets	Categoria GO_biological_process	Número de sets
<i>Pressão sanguínea</i>	6	GO_CALCIIUM_ION_TRANSPORT	2
Densidade mineral óssea	2	GO_CALCIIUM_MEDIATED_SIGNALING	2
<i>Altura</i>	1	GO_CARDIAC_CONDUCTION	1
<i>Circunferência do quadril</i>	2	<b><i>GO_CIRCULATORY_SYSTEM_DEVELOPMENT</i></b>	1
<i>Função pulmonar</i>	5	<b><i>GO_REGULATION_OF_CALCIIUM_ION_TRANSPORT</i></b>	10
Volume ventricular total	1	<b><i>GO_REGULATION_OF_CALCIIUM_MEDIATED_SIGNALING</i></b>	1
<i>Atividade física vigorosa</i>	1	GO_REGULATION_OF_DEVELOPMENTAL_GROWTH	1
<i>Circunferência da cintura</i>	7		
<i>Razão cintura-quadril</i>	6		

Tab. 1. Enriquecimentos para Kalenjin e Oromo que podem ser associados a corrida de resistência. Para poupar espaço, os sets de genes obtidos do FUMA-GWAS foram agrupados em categorias mais amplas. O número de sets em cada categoria está mostrado. Todos os sets têm um p-valor ajustado menor do que 0,05. Categorias pintadas têm sets em comum entre Kalenjin e Oromo. Categorias em negrito e itálico têm genes no top 5 de valores de PBSn para alguma comparação.

#### PBSn de Kalenjin



#### PBSn de Oromo

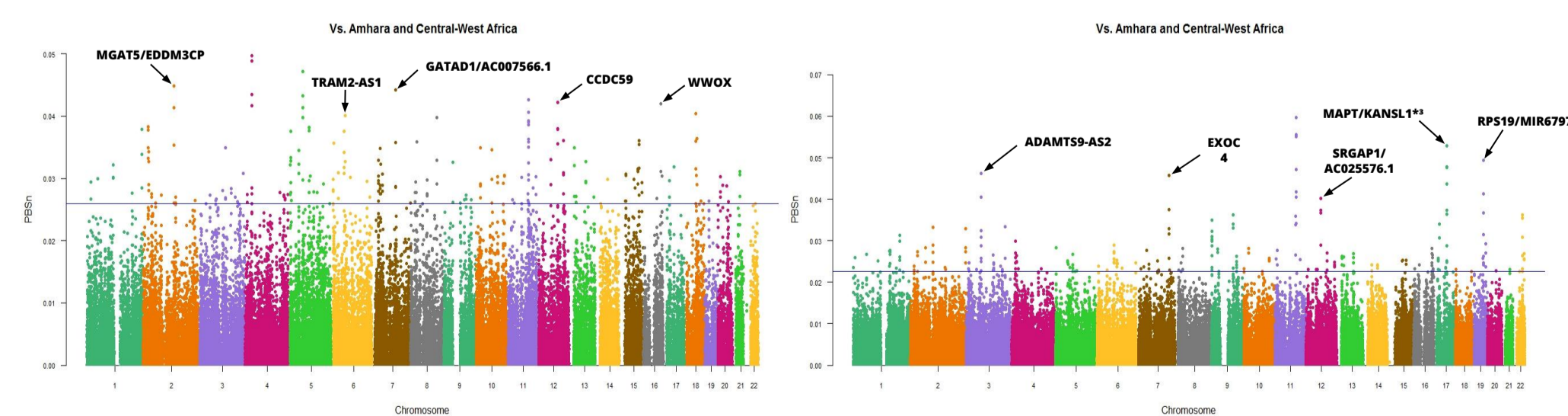
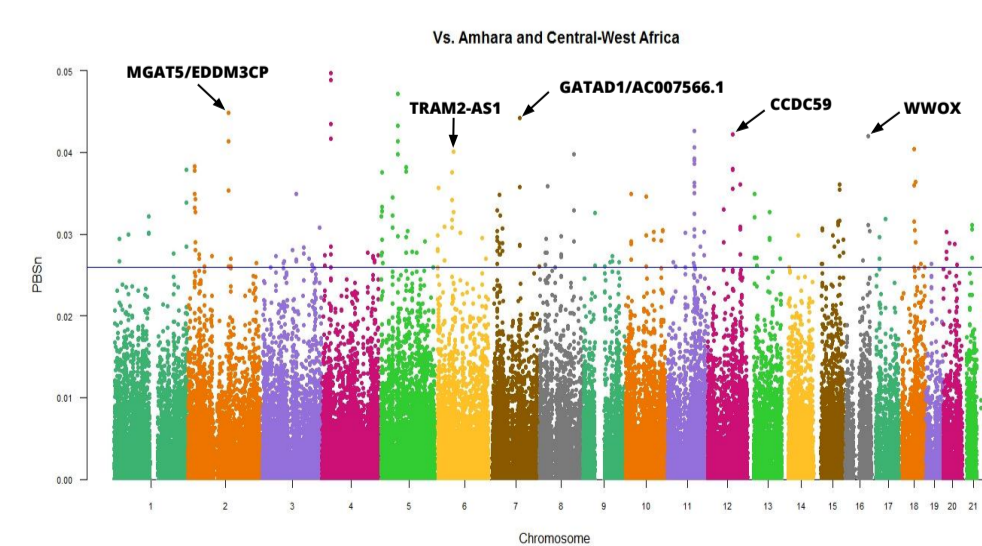
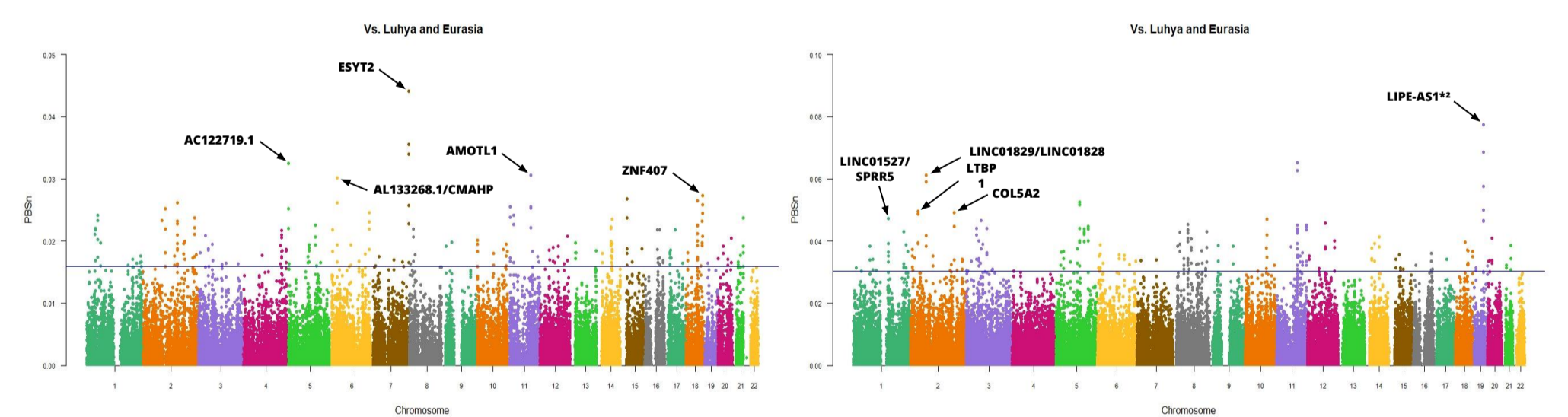
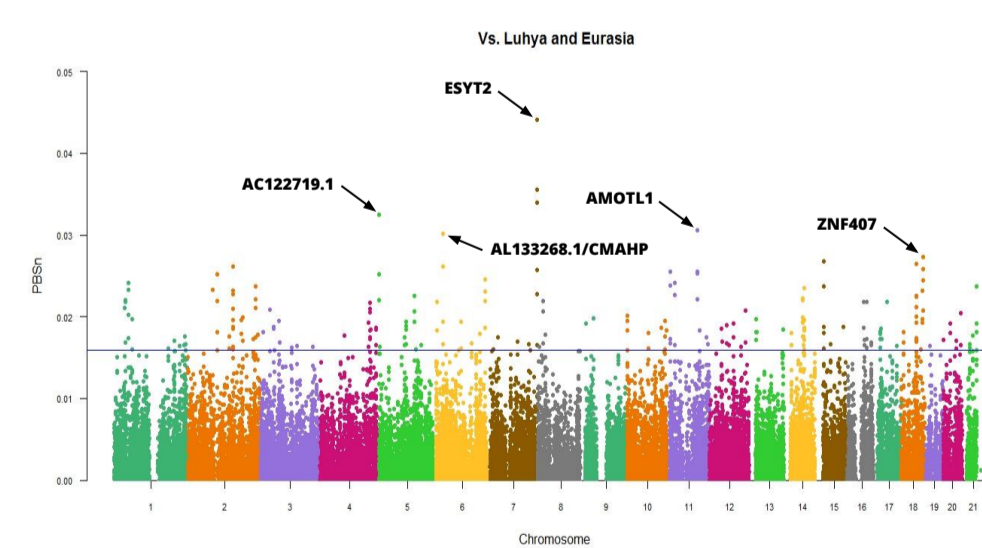
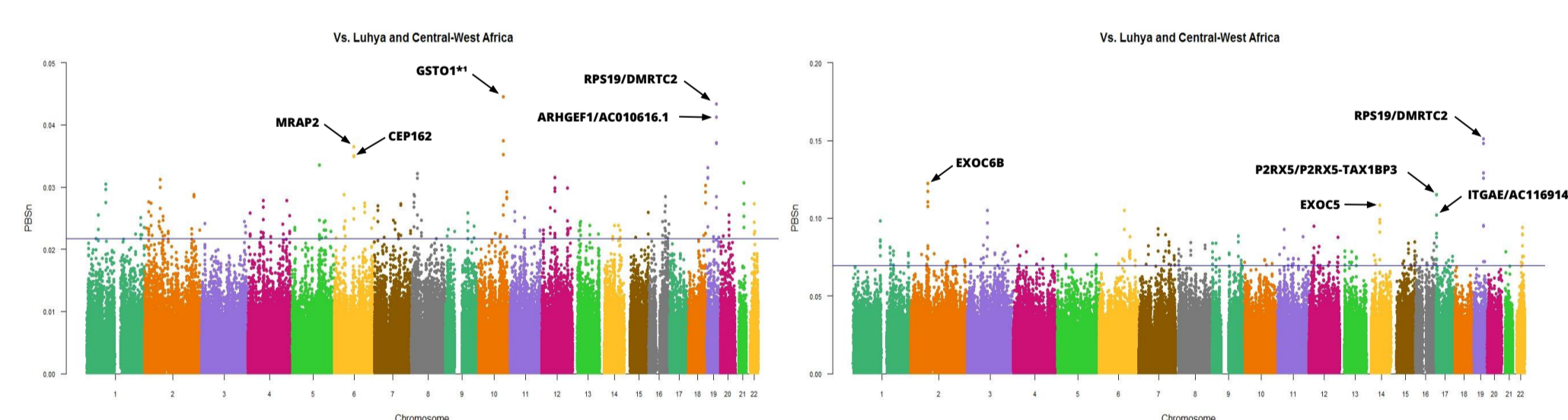


Fig. 2. Manhattan plots dos valores de PBSn para Kalenjin e Oromo quando comparadas com cada par de populações de referência. Cada ponto representa uma janela de 20 SNPs. A linha azul demarca os 0,1% mais diferenciados. Perceba que a escala não é a mesma para todas as comparações. As janelas não intergênicas com os 5 maiores valores de PBSn têm seus genes associados mostrados. \*Esse gene também aparece em janelas associadas a outros genes: <sup>1</sup>*GSTO1*, <sup>2</sup>*MIR4482*, <sup>3</sup>*CNFN*, *CEACAM1*; <sup>4</sup>*CR936218.1*, *AC217779.1*, *AC243658.1*.

Os Oromo também mostram um grande número de sets enriquecidos relacionados a sistema circulatório, que incluem o gene *LTBP1*, cuja proteína se liga à Fibrilina 1 na parede de artérias coronárias humanas. O *LTBP1* parece participar do desenvolvimento de características antropométricas também, já que ele se liga à proteína TGF- $\beta$ 1 na matriz da placa de crescimento, regulando a proliferação de condrocitos e a mineralização. Em estudos de GWAS, esse gene foi associado com altura e circunferência da cintura e do quadril. O set na categoria "forma do quadril" contém outro gene altamente diferenciado nos Kalenjin: o *WWOX*. Foi demonstrado, em ratos, que esse gene é essencial para o desenvolvimento dos ossos, também tendo sido relacionado a função pulmonar em estudos de GWAS. Esses resultados reforçam a ideia da corrida de resistência como uma atividade sistêmica complexa. Mesmo com vários genes altamente diferenciados nessas populações associados a características que conhecidamente afetam economia de corrida (traços antropométricos e respiratórios, assim como função de músculo estriado), serão necessários outros estudos para esclarecer a associação específica desses genes com a corrida de resistência, assim como suas características nas diferentes populações humanas.

#### Kalenjin

Categoria GWAS catalog	Número de sets	Categoria GO_biological_process	Número de sets
Densidade mineral óssea	1	GO_CALCIIUM_DEPENDENT_CELL_CELL_ADHESION	1
Homeostase de glicose	1	<b><i>GO_CIRCULATORY_SYSTEM_DEVELOPMENT</i></b>	1
<i>Altura</i>	1	<b><i>GO_REGULATION_OF_CALCIIUM_MEDIATED_SIGNALING</i></b>	1
<i>Forma do quadril</i>	1	GO_RESPIRATORY_SYSTEM_DEVELOPMENT	1
<i>Função pulmonar</i>	4		
<i>Circunferência da cintura</i>	1		
<i>Razão cintura-quadril</i>	2		