

## Article

# Ileal Lactase Expression Associates with Lactase Persistence Genotypes

Jan Krzysztof Nowak <sup>1,\*</sup>, Emilia Dybska <sup>1</sup>, Marzena Dworacka <sup>2</sup>, Natallia Tsikhan <sup>3</sup>, Victoria Kononets <sup>4</sup>, Saule Bermagambetova <sup>4</sup> and Jarosław Walkowiak <sup>1</sup>

- <sup>1</sup> Department of Pediatric Gastroenterology and Metabolic Diseases, Poznan University of Medical Sciences, Szpitalna 27/33, 60-572 Poznan, Poland; emilia.dybska@student.ump.edu.pl (E.D.); jarwalk@ump.edu.pl (J.W.)
- <sup>2</sup> Department of Pharmacology, Poznan University of Medical Sciences, 60-572 Poznan, Poland; mdworac@ump.edu.pl
- <sup>3</sup> Department of Pediatrics, Grodno State Medical University, 230009 Grodno, Belarus; tsikhannat@gmail.com
- <sup>4</sup> West Kazakhstan Marat Ospanov Medical University, Aktobe 030012, Kazakhstan; micropaleontolog@yandex.kz (V.K.); b.saule74@mail.ru (S.B.)
- \* Correspondence: jan.nowak@ump.edu.pl

**Abstract:** (1) Background: Lactose digestion depends on persistence genotypes (including rs4988235), the frequency of which exhibits broad geographical variability. However, little is known about the relationship between lactase (*LCT*) genotypes and intestinal expression of *LCT*. We aimed to investigate ileal expression of *LCT* depending on main genetic polymorphisms (rs4988235, rs3754689, rs3739022), age, sex, smoking status, body mass index (BMI), and the expression of other genes; (2) Methods: phenotype, array-based genotype, and ileal mucosal biopsy expression data were obtained from the CEDAR study; (3) Results: analyses included 196 healthy Europeans (53.6% women) aged  $53.0 \pm 13.6$  years with a mean BMI of  $25.6 \pm 4.2$  kg/m<sup>2</sup>, of whom 17.4% were smoking. Ileal *LCT* expression was mostly independent of age, sex, BMI, or smoking. Rs4988235 homozygous minor allele (GG) associated with lower *LCT* expression (vs. AG  $p = 2.2 \times 10^{-6}$ , vs. AA  $p = 1.1 \times 10^{-7}$ ). Homozygous major allele of rs3754689 (GG) was related to higher *LCT* expression (vs. AG  $p = 1.7 \times 10^{-5}$ , vs. AA  $p = 0.0074$ ). Rs3754689 genotype did not modify *LCT* expression (GG vs. AG  $p = 0.051$ ) in rs4988235-heterozygous subgroup. Interestingly, *CD14*, which is a marker of monocytes and macrophages, was the strongest negative transcriptomic correlate of *LCT* expression ( $r = -0.57$ ,  $p_{FDR} = 1.1 \times 10^{-14}$ ); (4) Conclusions: both rs4988235 and rs3754689 associated with ileal *LCT* expression, which did not seem related to age, sex, smoking, or BMI. The inverse correlation between *LCT* and *CD14* expression in the ileum is striking and requires further investigation.

**Keywords:** lactase persistence; lactose intolerance; lactose; milk; macrophage; monocyte



**Citation:** Nowak, J.K.; Dybska, E.; Dworacka, M.; Tsikhan, N.; Kononets, V.; Bermagambetova, S.; Walkowiak, J. Ileal Lactase Expression Associates with Lactase Persistence Genotypes. *Nutrients* **2021**, *13*, 1340. <https://doi.org/10.3390/nu13041340>

Academic Editor: M. Andrea Azcarate-Peril

Received: 10 March 2021  
Accepted: 12 April 2021  
Published: 17 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Individuals with lactase persistence (LP) phenotype digest lactose till senility [1]. This is enabled by the sustained activity of lactase (*LCT*), a  $\beta$ -d-galactosidase also known as lactase phlorizin hydrolase (LPH) [2,3]. Its expression is observed on the brush border of the small intestine, with the highest values in jejunal microvilli [2]. Patients with non-persistence phenotype suffer from indigestion after the intake of lactose-containing products [3,4]. In this group, *LCT* activity usually drops after the age of five, following reduced *LCT* expression due to the presence of the homozygous minor allele C (G) at *LCT* position-13910 (rs4988235) [5,6].

The autosomal dominant inheritance pattern has enabled dairy intake in various populations around the world, probably reflecting adaptation to evolutionary pressures of changing lifestyles, including animal domestication [7]. The highest frequency of LP characterizes descendants of northern Europeans and some ethnic groups in Africa, which

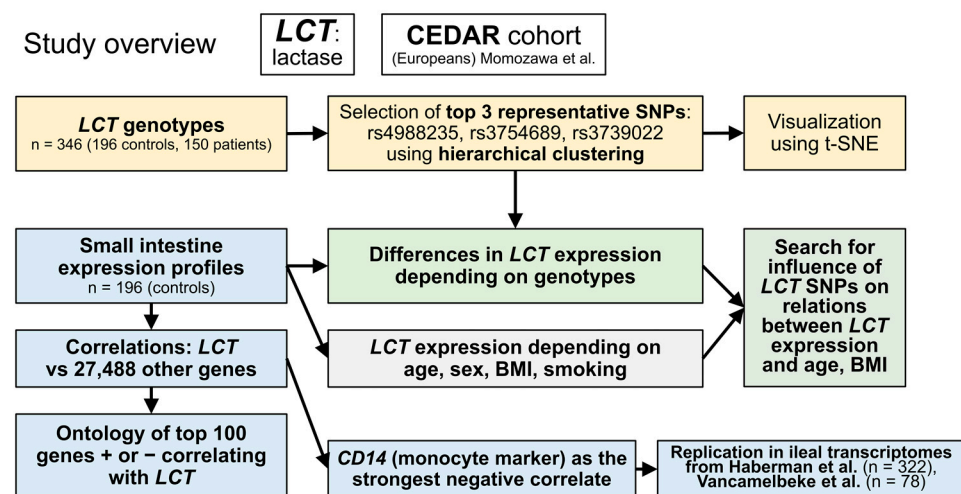
harbor distinct persistence polymorphisms [7–11]. LP is rarely found in South-East Asia and generally, the frequency of variants enabling lactose digestion in the adulthood falls from northwest to southeastern regions of Europe, Asia, and Africa [11].

Previous studies showed the applicability of rs4988235 genetic testing to exclude the risk of lactose intolerance in adults [8,12]. Currently, apart from standard clinical tests, the combined genetic and epigenetic assessment is advocated. DNA methylation in the *LCT* promoter region was found to associate with rs4988235 genotype, *LCT* gene expression, *LCT* activity, and LP [13]. Interestingly, the relationship between enzymatic activity and *LCT* mRNA level has not been confirmed in adults, even though the underlying polymorphism is known to act via a transcription regulation element [14].

*LCT* gene expression is insufficiently studied, with most investigations focusing on LP genetics and clinical assessment. This is also due to the difficulty in obtaining intestinal biopsies. In this study, we used a large ileal transcriptomic dataset from Correlated Expression & Disease Association Research (CEDAR) cohort to verify hypotheses that ileal *LCT* expression: (a) associates with *LCT* genotypes; (b) correlates with age and body mass index (BMI); (c) depends on sex and smoking status; and (d) correlates with expression of other transcripts.

## 2. Materials and Methods

CEDAR was a large study aimed at identifying expression-quantitative trait loci related to inflammatory bowel disease risk in a number of tissues, including leukocyte subsets and intestinal biopsies [15], both in healthy controls and patients with various diseases. This study focused on the healthy controls (n = 196), most of whom were middle-aged non-smokers with good nutritional status or overweight (more details in Section 3.2). The scheme of the study is presented in Figure 1.



**Figure 1.** Scheme of the study. BMI: body mass index, CD14: cluster of differentiation 14, CEDAR: Correlated Expression & Disease Association Research (by Momozawa et al. [15]), LCT: lactase, SNP: single nucleotide polymorphism, t-SNE: t-distributed stochastic neighbor embedding.

### 2.1. Genotypes

Complete raw data (Illumina HumanOmniExpress-12 v1.0) were obtained from the Array Express (accession ID: E-MTAB-6666) and processed with GenomeStudio 2.0.5 using the adequate array design file. After clustering and quality check, SNPs from a region on chromosome 2 including the *LCT* gene with 100 kbp margin (chr2:136,445,410–136,694,750; GRCh37/hg19) were exported to a genotype matrix, which was further processed to represent the count of the major allele in the genotype. Full *LCT* locus genotype data were used to explore the genetic structure of the population by performing t-distributed stochastic neighbor embedding of study participants (t-SNE, Rtsne). Only variants with

minor allele frequency (MAF) > 10% and variance > 10% were included. The obtained matrix was used for hierarchical clustering with the fastcluster package in R (R Software Foundation, Vienna, Austria), using a Ward D dissimilarity matrix. After dendrogram inspection (ggdendro) and setting a cut-off level, a most representative SNP was selected for each of the *three* resulting clusters. This choice considered minimal distance to other cluster members (representativeness) or prominence in the literature (rs4988235). The genotypes were subject to Pearson's correlation analysis to illustrate their relationships.

## 2.2. Transcriptomic Data

Raw data from the analysis of terminal ileum mucosal RNA using Illumina's HT12v4 expression microarray were downloaded from the Array Express (accession ID: E-MTAB-6667) and read using limma. Arrays with no variance and mean expression level of 0 were excluded, and the others underwent a quality check. The data were quantile-normalized using the qnc function, which adds an offset of 16 prior to log<sub>2</sub> transformation. Since the mean log<sub>2</sub> (signal + 16) from control probes was ~4.5, only genes with mean expression > 5 were included, bringing the number of available probes down from 47,323 to 27,490. Samples reporting gene expression in the terminal ileum of healthy participants were selected. The dataset obtained in this way was further used for downstream analyses.

Two probes binding *LCT* transcripts correlated with each other ( $r = 0.87$ ) and matched the same transcript with no other predicted complementary sequences in the human transcriptome (Basic Local Alignment Search Tool, BLAST) and were therefore averaged. Correlations of *LCT* with other transcripts were investigated and top genes positively and negatively associated with *LCT* were analyzed using Gene Set Enrichment Analysis biological process ontology (the Broad Institute, Cambridge, MA, USA). Phenotype data were drawn from the ArrayExpress (accession ID: E-MTAB-6667). All the statistical analyses were conducted using R directly (tidyverse, ggpubr) or via JASP as an interface (University of Amsterdam, Amsterdam, the Netherlands).

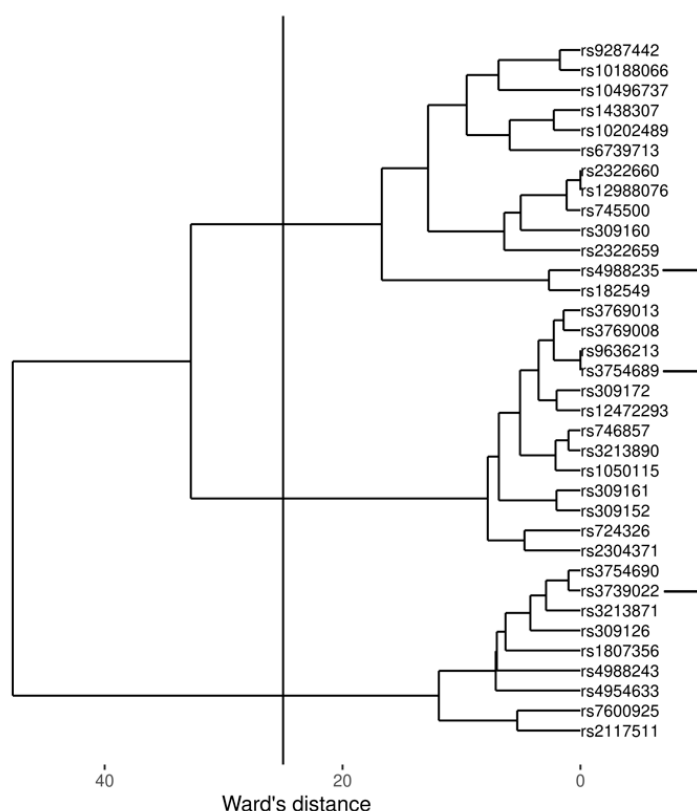
A complementary correlation analysis of *LCT* and *CD14* (cluster of differentiation 14) was conducted in data from Haberman et al. [16] (Gene Expression Omnibus accession ID: GSE57945) and Vancamelbeke et al. (GSE102133) [17] directly at the R2 genomics analysis and visualization platform of the Amsterdam Medical Center (<http://r2.amc.nl> (accessed on 10 March 2021)).

## 3. Results

### 3.1. Genotypes

Fifty-two SNPs were identified in *LCT* locus ( $\pm 100$  kbp) of 346 CEDAR participants (controls and patients), who are characterized in the original study [15]. In brief, the subjects were Europeans subject to screening colonoscopy, most of whom had no diseases but with a fraction having different diagnoses, including asthma. Seventeen polymorphisms with low MAF and variance were excluded ( $n = 17$ ) and the remaining 35 were clustered using the Ward's distance (Figure 2).

Three clusters were identified. For each of the three clusters, a representative SNP was chosen: rs4988235, rs3754689, and rs3739022. Rs4988235 was preferred because it is the best-known variant associated with lactose intolerance. The remaining two SNPs were selected because of short distance to all other variants included in their clusters, i.e., being most representative for the largest number of polymorphisms within the clusters. It should be noted that genes in clusters were similar (e.g., correlation between rs4988235 and rs182549 was  $r = 0.98$ ). Correlations between the three SNPs selected for further analysis are listed in Table 1.



**Figure 2.** Clustering of *LCT* locus polymorphisms with minor allele frequency > 10% and variance > 10% using a Ward's dissimilarity matrix. Three polymorphisms selected as representative for respective clusters are indicated with horizontal lines on the right.

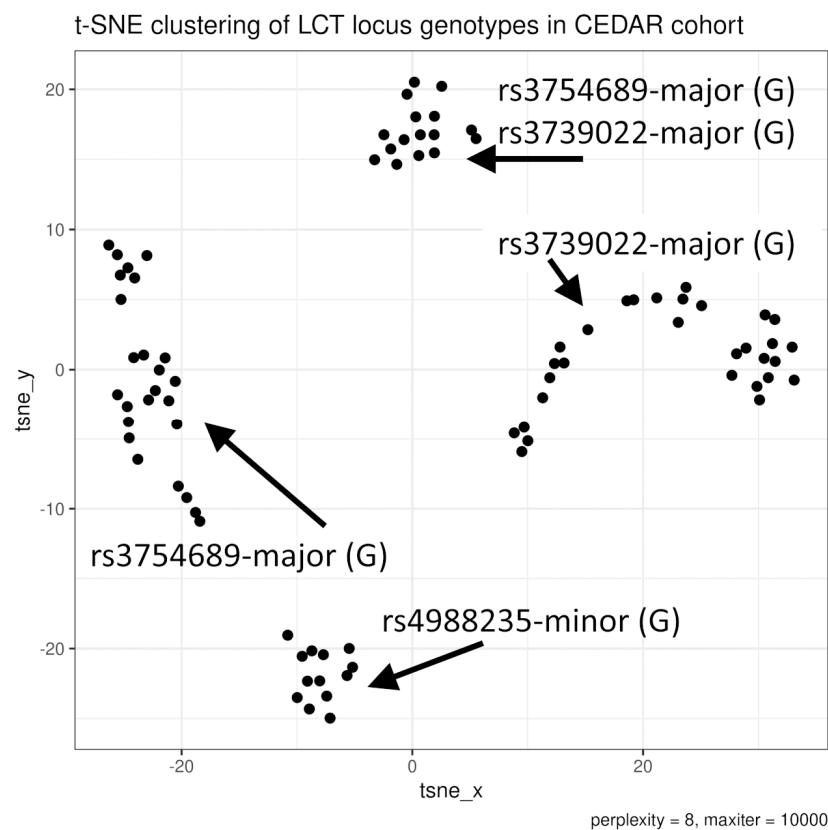
**Table 1.** Correlations between analyzed polymorphisms representative of genetic variance in *LCT* locus.

SNP Pair	Pearson's r	P
rs4988235–rs3754689	0.60	$2.2 \times 10^{-16}$
rs4988235–rs3739022	0.38	$2.6 \times 10^{-13}$
rs3754689–rs3739022	−0.18	$6.8 \times 10^{-4}$

Eighty-four CEDAR participants exhibiting unique combinations of *LCT* loci genotypes were clustered using t-SNE (Figure 3). The identified clusters exhibited specific combinations (haplotypes) of rs4988235, rs3754689, and rs3739022 genotypes, confirming that the selection of the three SNPs may be used to describe a considerable share of *LCT* locus genetic diversity. Table 2 presents allele frequencies for patients included in expression analyses. Please note that the genotypes are presented as per Illumina upper strand output, and therefore changes often known as C/T are presented herein as G/A, a convention also adopted in reports from other teams [18].

**Table 2.** *LCT* locus allele frequency in healthy CEDAR participants with paired ileal expression data.

Polymorphism	Homozygous Major	Heterozygous	Homozygous Minor
rs4988235	AA 57 (29.1%)	AG 98 (50.0%)	GG 41 (20.9%)
rs3754689	GG 121 (61.7%)	AG 66 (33.7%)	AA 9 (4.6%)
rs3739022	GG 141 (72.3%)	AG 50 (25.6%)	AA 1 (0.5%)



**Figure 3.** Visualization of t-distributed stochastic neighbor embedding (t-SNE) clustering of sets of *LCT* locus genotypes in CEDAR cohort based on 52 SNPs from HumanOmniExpress–12 v1.0. High prevalence of major/minor alleles is indicated.

### 3.2. Expression

Terminal ileum biopsy expression data were available for 196 healthy participants of the CEDAR study. The average age  $\pm$  SD was  $53.0 \pm 13.6$  years (range 17–80 years). The group included 105 women (53.6%) and 91 men (46.4%). Most of the subjects were non-smokers ( $n = 162/196$ , 82.6%). Participants mostly exhibited good nutritional status or were overweight, with a mean body mass index (BMI) of  $25.6 \pm 4.2$  kg/m<sup>2</sup> (range 15.9–40.6).

Two probes matching *LCT* had mean expression  $> 5$ : ILMN\_1749026 and ILMN\_2046073. BLAST revealed that they both should bind only one and the same transcript NM\_002299.2. For this reason and because of a high degree of correlation ( $r = 0.87$ ), the levels of the two probes were averaged.

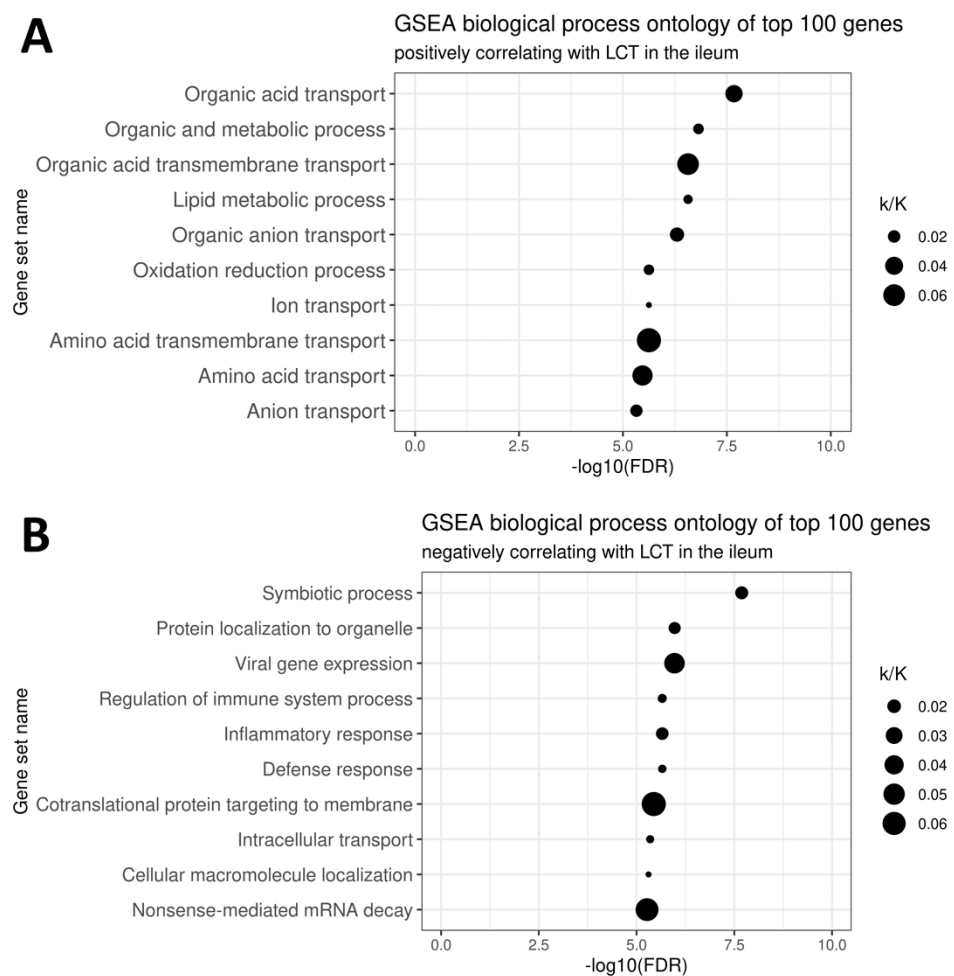
Ileal *LCT* did not significantly correlate with age ( $r = -0.10$ ,  $p = 0.169$ ) or BMI ( $r = -0.055$ ,  $p = 0.445$ ). It also did not depend on sex (female  $6.22 \pm 1.32$  vs. male  $6.09 \pm 1.16$ , Mann-Whitney  $p = 0.677$ ) or smoking status (smoker  $6.00 \pm 1.10$  vs. non-smoker  $6.19 \pm 1.28$ ,  $p = 0.505$ ).

The top transcriptomic correlates of the expression of *LCT* are reported in Table 3. Interestingly, the expression of *CD14*, which is a marker of monocytes and macrophages, was the strongest negative correlate of *LCT* ( $r = -0.57$ ,  $p_{\text{FDR}} = 1.1 \times 10^{-14}$ ). *LCT* expression most strongly positively associated with mitochondrial acylating methylmalonate-semialdehyde dehydrogenase ( $r = 0.56$ ,  $p_{\text{FDR}} = 2.4 \times 10^{-14}$ ). Other noteworthy correlates of *LCT* included *TLR3* (toll-like receptor 3;  $r = 0.51$ ,  $p = 2.4 \times 10^{-11}$ ), *RORC* (RAR-related orphan receptor C;  $r = 0.49$ ,  $p = 7.4 \times 10^{-11}$ ), and *ACE2* (angiotensin converting enzyme 2;  $r = 0.47$ ,  $p = 3.8 \times 10^{-10}$ ), all of which attract attention in the context of inflammatory bowel disease. *ACE2* is also known to enable cellular entry of the severe acute respiratory syndrome

coronavirus (SARS-CoV-2) [19]. Full results of correlation analysis are presented in Table S1 (supplement).

The negative correlation between *LCT* and *CD14* was confirmed in a dataset of 322 ileal biopsies from patients with inflammatory bowel diseases and healthy controls ( $r = -0.43$ ,  $p = 6.78 \times 10^{-16}$ ) by Haberman et al. [16] and in 78 ileal biopsies from patients with Crohn's disease and controls obtained by Vancamelbeke et al. ( $r = -0.55$ ,  $p = 1.83 \times 10^{-7}$ ) [17].

Principal results of gene ontology analysis of transcripts positively and negatively correlating with *LCT* is shown in Figure 4. Genes positively correlating with *LCT* were involved in nutrient transmembrane transport and metabolism. On the other hand, genes negatively associated with *LCT* related to symbiosis, inflammatory reactions, and protein trafficking.



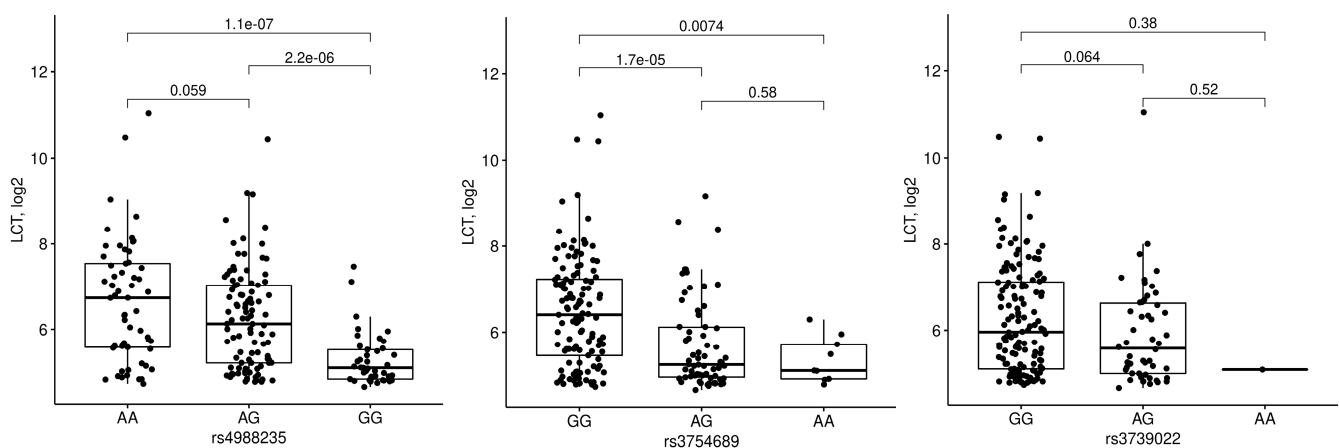
**Figure 4.** Gene set enrichment analysis of top 100 genes positively (**A**) and negatively (**B**) correlating with *LCT*. The k/K ratio indicates what fraction of all the genes in a given set overlapped with the investigated list of 100 genes (e.g., out of 353 genes in the Organic acid transport gene set, 13 were also found among genes most strongly correlating with *LCT*, yielding k/K = 0.037).

**Table 3.** Pearson’s correlations between mean *LCT* expression and other transcripts in the terminal ileum of healthy participants of CEDAR study.

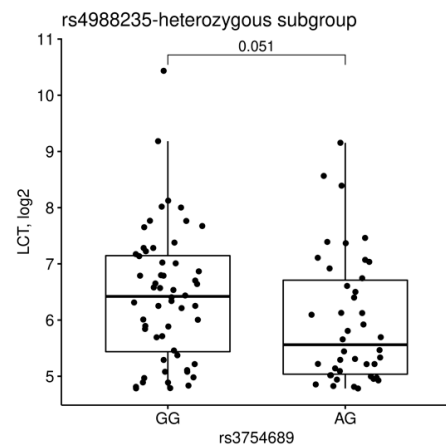
Gene	r	PfDR
Positively correlating with <i>LCT</i> expression		
<i>ALDH6A1</i> —Aldehyde Dehydrogenase 6 Family Member A1	0.56	$2.4 \times 10^{-14}$
<i>LOC441442</i>	0.56	$8.9 \times 10^{-14}$
<i>C18orf18</i> —Long Intergenic Non-Protein Coding RNA 526	0.55	$2.5 \times 10^{-13}$
<i>XYLB</i> —Xylulokinase	0.55	$2.8 \times 10^{-13}$
<i>CISD1</i> —CDGSH Iron Sulfur Domain 1	0.54	$6.3 \times 10^{-13}$
<i>BCL2L15</i> —BCL2 Like 15	0.54	$6.5 \times 10^{-13}$
<i>LOC100128907</i>	0.54	$8.3 \times 10^{-13}$
<i>CCDC25</i> —Coiled-Coil Domain Containing 25	0.54	$8.9 \times 10^{-13}$
<i>BEND7</i> —BEN Domain Containing 7	0.53	$2.3 \times 10^{-12}$
<i>APOM</i> —Apolipoprotein M	0.53	$3.2 \times 10^{-12}$
Negatively correlating with <i>LCT</i> expression		
<i>CD14</i> —Monocyte Differentiation Antigen CD14	−0.57	$1.1 \times 10^{-14}$
<i>RNASET2</i> —Ribonuclease T2	−0.50	$2.9 \times 10^{-11}$
<i>CDC37</i> —Cell Division Cycle 37, HSP90 Cochaperone	−0.50	$3.6 \times 10^{-11}$
<i>C7orf50</i> —Chromosome 7 Open Reading Frame 50	−0.50	$5.4 \times 10^{-11}$
<i>LOC730187</i>	−0.49	$6.3 \times 10^{-11}$
<i>UBE2I</i> —Ubiquitin Conjugating Enzyme E2 I	−0.49	$1.0 \times 10^{-10}$
<i>LOC650369</i>	−0.49	$1.0 \times 10^{-10}$
<i>RCC2</i> —Regulator of Chromosome Condensation 2	−0.49	$1.0 \times 10^{-10}$
<i>HK1</i> —Hexokinase 1	−0.49	$1.1 \times 10^{-10}$
<i>RPL27</i> —Ribosomal Protein L27	−0.48	$2.3 \times 10^{-10}$

### 3.3. Relations between *LCT* Genotype and Expression

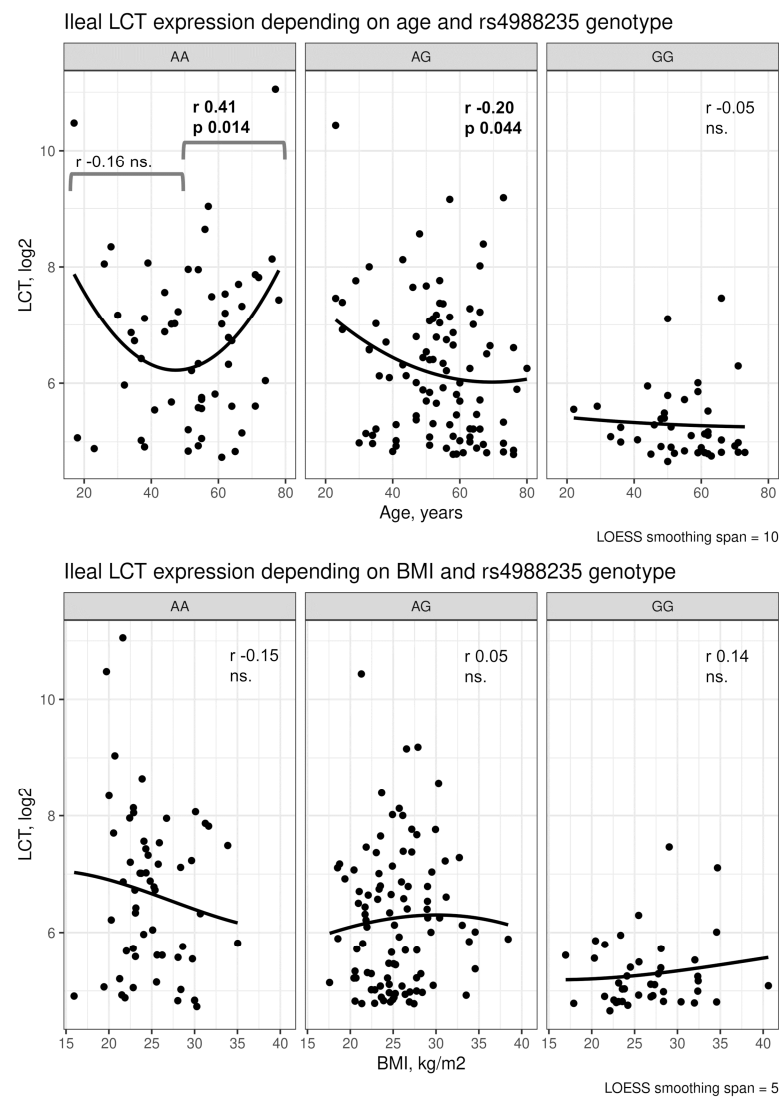
Expression of *LCT* was dependent on rs4988235 and rs3754689 genotypes (Figure 5). Comparisons of homozygous major (G) vs. heterozygous rs3739022 alleles revealed no statistically significant differences ( $p = 0.064$ ). Lower *LCT* expression was therefore associated only with G allele of rs4988235 and A allele of rs3754689. To verify if rs3754689 may modify the relationship between rs4988235 and *LCT*, analysis was performed in a subgroup heterozygous for rs4988235. The analysis did not reveal statistically significant differences, although a larger sample size would probably be required to determine whether such relationship exists (Figure 6).

**Figure 5.** Ileal *LCT* expression depending on *LCT* locus single nucleotide polymorphisms.

Further analyses revealed that rs4988235 might modify the relationship between *LCT* expression and age in a non-linear manner (Figure 7). Whereas in patients with GG and AG genotypes the *LCT* vs. age relationship seemed linear, in the AA group it appeared U-shaped, with greater *LCT* expression in early adulthood and in the old age. This effect, however, appeared to be driven by outliers and would require additional verification. No similar effects were found for BMI.



**Figure 6.** Ileal *LCT* expression depending on rs3754689 genotype in a subgroup of rs4988235-heterozygous (AG) CEDAR study participants.



**Figure 7.** Relationships of ileal *LCT* expression vs. age and BMI depending on the lactase persistence polymorphism rs4988235. Pearson’s r is shown in the case of AA genotype correlation analysis was conducted separately in patients younger and older than 50 years old due to the appearance of a U-shaped pattern.



## 4. Discussion

In this work, we described genomic and transcriptomic characteristics of lactase persistence status in the ileal mucosa of 196 controls from the CEDAR cohort [15]. Two representative SNPs related to ileal *LCT* gene expression: rs4988235 (major allele A) and rs3754689 (major allele G). No strong and consistent relationships were found for age, sex, BMI, or smoking status. A thought-stimulating negative correlation between *LCT* and *CD14* was found.

### 4.1. The Role of rs4988235 and rs3754689

The robust association between rs4988235 (C/T -13910) and lactase persistence in Europeans is well-known [8,12,20]. The A allele maintains high *LCT* activity, and thus enables cow's milk digestion in adulthood [20]. The A allele carriers have a greater mucosal expression of *LCT* in the intestine [21]. Our observations confirm this genetic regulation of *LCT* transcription.

In contrast, rs3754689 has been poorly described in the context of LP. It was investigated as one of the *LCT* haplotype markers, but was not strongly associated with lactose intolerance [22]. It related, however, to a lower risk of inhibitor development in patients with hemophilia A [23]. The variant rs3754689 requires further investigation to better understand its potential clinical value.

### 4.2. Ileal *LCT* Expression and Clinical Characteristics

Depending on the ethnicity, up to 70% of adults suffer from lactose intolerance as a result of a progressive decrease in intestinal lactase activity [11]. Genotyping rs4988235 (and the closely correlated rs182549) is considered to enable diagnosis of lactose intolerance in symptomatic patients older than six years [24] and may be considered to have accuracy increasing with age [3]. Nevertheless, the relationship with clinical status [6] or LP phenotype [13] is not strict. Epigenetic changes accumulate with age and may contribute to markedly different DNA regulatory landscapes [25,26]. These probably provide a variety in the programmed ability to digest lactose-rich dairy products. Our study did not investigate gene methylation, nor did it confirm the general (across genotypes) significance of age on *LCT* expression in the investigated group of European adults. However, high *LCT* expression constituted a common feature of rs4988235 AA-homozygotes in early adulthood and the old age. Further analyses are needed to check whether the U-shaped effect (which seemed mostly outlier-driven) can be confirmed in Europeans or found in other ethnic groups.

The ability to digest milk affects dietary habits in adulthood. Consumption of lactose-containing dairy products without risk of diarrhea or other symptoms may associate with a more varied diet, and thus distinct body composition. Genome-wide association studies indicated rs4988235 as associated with obesity [27–29], which can be explained by a lifetime exposure pattern to cow's milk and fat-rich dairy products [27–29]. However, this study did not confirm the relationship between ileal *LCT* expression and BMI. Neither smoking status nor sex-specific analyses provided any associations with the *LCT* variants, reinforcing the observation by Kettunen et al. of no gender impact on BMI in LP individuals [27].

### 4.3. Transcriptomic Correlates of Ileal *LCT* Expression Including *CD14*

Besides the ileal *LCT* genotype-expression relationships, we also explored its transcriptomic correlates. The top correlate (and a negative one) of *LCT* was the key monocyte marker *CD14*. Previously, expression of *LCT* (dependent on rs370868556), was related to the white blood cell pool [30]. It might be speculated that high *LCT* expression reduces distension of the bowel wall, contributing to less tissue stress that could attract *CD14+* cells. On the other hand, attracted *CD14+* cells could potentially damage villi and contribute to lactase deficiency. An association with *RNASET2* (ribonuclease T2) might indicate a reduced need to recognize pathogen RNAs, which could be related to antiviral activity of whey protein [31]. The lactose/monocyte axis was previously associated with the upregu-

lation of gastrointestinal antimicrobial peptides such as cathelicidin antimicrobial peptide (CAMP) and thus modulated mucosal immunity in the intestine [32]. Such relationships could intervene in the pathogenesis of small intestinal bacterial overgrowth. The correlation analyses described above indicate interesting new research themes at the interface between LP and immunity.

Other top negative correlates of *LCT* expression also included cyclin CDC37 (cell division cycle 37), which stabilizes LKB1 (liver kinase B1), the protein deficient in Peutz-Jeghers syndrome. There are, however, no data available on putative relationships between lactose digestion or milk consumption and outcomes in this genetic disease. The inverse association between hexokinase 1 and *LCT* expression may indicate shifts in mucosal carbohydrate processing.

*LCT* expression associated with metabolism-related *ALDH6A1* (aldehyde dehydrogenase 6 family member A1), which contributes to downregulation of the acetyl-CoA pathway. Such processes lead to adipose tissue changes in diabetic obese individuals [33]. A positive correlation between *LCT* and *ALDH6A1* may suggest a link between LP and adiposity, possibly mediated by microbiota [18], which could be further investigated. Another positive correlate of *LCT* expression was *XYLB*, encoding xylulokinase involved in energy metabolism, which we speculate might be related to small intestinal bacterial overgrowth. Interestingly, *CISD1* (CDGSH iron sulfur domain 1) also associated with *LCT*; *CISD1* is a mitochondrial protein reduced in cystic fibrosis, the deficiency of which may cause Wolfram syndrome with neurodegeneration [34]. Apparently, higher *LCT* expression correlates with enhanced energy metabolism and nutrient transport.

#### 4.4. Generalization and Limitations

Data obtained from the CEDAR study provide insight into the genetics of *LCT* locus in Europeans. They present two key advantages: a large cohort and gene expression profiling in intestinal biopsies. This foundation allowed us to conduct an ileum-focused analysis, determine *LCT* ileal expression, and identify novel *LCT* correlates. Dietary habits, especially daily milk intake, are lacking factors in our analyses. Moreover, no information on microbiota was available. It is necessary to reinstate that *LCT* gene expression is not synonymous with lactase enzyme activity. Data obtained from Europeans should not be generalized to the global population as not only the frequency of lactase persistence but also its epigenetic regulation may undergo population-specific fine-tuning. The results should also not be generalized to children. Finally, results from omics analyses may considerably change depending on the overall approach and methods.

## 5. Conclusions

To conclude, our study indicates a positive correlation between major A allele of rs4988235 as well as major G allele of rs3754689 and *LCT* gene expression in ileal mucosa. Ileal *LCT* expression does not seem related to age, sex, smoking, or BMI. Potential modification of the relationship between rs4988235 and LP by rs3754689 appears as an interesting research topic. The inverse correlation between *LCT* and *CD14* expression in the ileum is striking and requires further investigation, especially in the context of irritable bowel syndrome, small intestinal bacterial overgrowth, and cow's milk protein allergy.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/nu13041340/s1>, Table S1: Full results of *LCT* correlation analysis in CEDAR ileal biopsies from healthy controls.

**Author Contributions:** Conceptualization, J.K.N. and J.W.; methodology, J.K.N., E.D., M.D., and J.W.; software, J.K.N.; formal analysis, J.K.N.; investigation, J.K.N., E.D., M.D., N.T., V.K., and S.B.; resources, J.K.N., E.D., and J.W.; data curation, J.K.N.; writing—original draft preparation, J.K.N. and E.D.; writing—review and editing, M.D., N.T., V.K., S.B., and J.W.; visualization, J.K.N.; supervision, J.W.; project administration, J.K.N.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded from a grant awarded by the Polish National Science Center to J.W. (2017/25/B/NZ5/02783). V.K., S.B. (WKMOMU), J.K.N., M.D., and J.W. (PUMS) were supported by the Social Health Insurance Project, Republic of Kazakhstan (Contract No. SHIP-2.3/CS-02).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are publicly available in the Array Express (E-MTAB-6666 and E-MTAB-6667) and Gene Expression Omnibus (GSE57945 and GSE102133).

**Conflicts of Interest:** J.K.N. reports personal fees from Norsa Pharma within an EU-funded project, grant support from Biocodex, and non-financial support from Nutricia, outside the submitted work. J.W. reports personal fees and non-financial support from Biocodex, BGP Products, Chiesi, Hipp, Humana, Mead Johnson Nutrition, Merck Sharp & Dohme, Nestle, Norsa Pharma, Nutricia, Roche, Sequoia Pharmaceuticals, and Vitis Pharma, outside the submitted work, and also grants, personal fees, and non-financial support from Nutricia Research Foundation Poland, also outside the submitted work.

## References

1. Mađry, E.; Fidler, E.; Walkowiak, J. Lactose Intolerance—Current State of Knowledge. *Acta Sci. Pol. Technol. Aliment.* **2010**, *9*, 343–350.
2. Misselwitz, B.; Butter, M.; Verbeke, K.; Fox, M.R. Update on Lactose Malabsorption and Intolerance: Pathogenesis, Diagnosis and Clinical Management. *Gut* **2019**, *68*, 2080–2091. [[CrossRef](#)] [[PubMed](#)]
3. Couce, M.L.; Sánchez-Pintos, P.; González-Vioque, E.; Leis, R. Clinical Utility of LCT Genotyping in Children with Suspected Functional Gastrointestinal Disorder. *Nutrients* **2020**, *12*, 3017. [[CrossRef](#)]
4. Forsgård, R.A. Lactose Digestion in Humans: Intestinal Lactase Appears to Be Constitutive Whereas the Colonic Microbiome Is Adaptable. *Am. J. Clin. Nutr.* **2019**, *110*, 273–279. [[CrossRef](#)]
5. Rasinperä, H.; Kuokkanen, M.; Kolho, K.-L.; Lindahl, H.; Enattah, N.S.; Savilahti, E.; Orpana, A.; Järvelä, I. Transcriptional Downregulation of the Lactase (LCT) Gene during Childhood. *Gut* **2005**, *54*, 1660–1661. [[CrossRef](#)] [[PubMed](#)]
6. Mađry, E.; Lisowska, A.; Kwiecień, J.; Marciniak, R.; Korzon-Burakowska, A.; Drzymała-Czyż, S.; Mojs, E.; Walkowiak, J. Adult-Type Hypolactasia and Lactose Malabsorption in Poland. *Acta Biochim. Pol.* **2010**, *57*, 585–588. [[CrossRef](#)] [[PubMed](#)]
7. Tishkoff, S.A.; Reed, F.A.; Ranciaro, A.; Voight, B.F.; Babbitt, C.C.; Silverman, J.S.; Powell, K.; Mortensen, H.M.; Hirbo, J.B.; Osman, M.; et al. Convergent Adaptation of Human Lactase Persistence in Africa and Europe. *Nat. Genet.* **2007**, *39*, 31–40. [[CrossRef](#)] [[PubMed](#)]
8. Schirru, E.; Corona, V.; Usai-Satta, P.; Scarpa, M.; Oppia, F.; Loriga, F.; Cucca, F.; De Virgiliis, S.; Rossino, R.; Macis, M.D.; et al. Genetic Testing Improves the Diagnosis of Adult Type Hypolactasia in the Mediterranean Population of Sardinia. *Eur. J. Clin. Nutr.* **2007**, *61*, 1220–1225. [[CrossRef](#)] [[PubMed](#)]
9. Sibley, E.; Ahn, J.K.; Theodore, E. Woodward Award: Lactase Persistence SNPs in African Populations Regulate Promoter Activity in Intestinal Cell Culture. *Trans. Am. Clin. Clim. Assoc.* **2011**, *122*, 155–165.
10. Schlebusch, C.M.; Sjödin, P.; Skoglund, P.; Jakobsson, M. Stronger Signal of Recent Selection for Lactase Persistence in Maasai than in Europeans. *Eur. J. Hum. Genet.* **2013**, *21*, 550–553. [[CrossRef](#)] [[PubMed](#)]
11. Anguita-Ruiz, A.; Aguilera, C.M.; Gil, Á. Genetics of Lactose Intolerance: An Updated Review and Online Interactive World Maps of Phenotype and Genotype Frequencies. *Nutrients* **2020**, *12*, 2689. [[CrossRef](#)] [[PubMed](#)]
12. Enattah, N.S.; Sahi, T.; Savilahti, E.; Terwilliger, J.D.; Peltonen, L.; Järvelä, I. Identification of a Variant Associated with Adult-Type Hypolactasia. *Nat. Genet.* **2002**, *30*, 233–237. [[CrossRef](#)] [[PubMed](#)]
13. Leseva, M.N.; Grand, R.J.; Klett, H.; Boerries, M.; Busch, H.; Binder, A.M.; Michels, K.B. Differences in DNA Methylation and Functional Expression in Lactase Persistent and Non-Persistent Individuals. *Sci. Rep.* **2018**, *8*, 1–14. [[CrossRef](#)]
14. Olsen, W.A.; Li, B.U.K.; Lloyd, M.; Korsmo, H. Heterogeneity of Intestinal Lactase Activity in Children: Relationship to Lactase-Phlorizin Hydrolase Messenger RNA Abundance. *Pediatr. Res.* **1996**, *39*, 877–881. [[CrossRef](#)]
15. Momozawa, Y.; Dmitrieva, J.; Théâtre, E.; Deffontaine, V.; Rahmouni, S.; Charlotiaux, B.; Crins, F.; Docampo, E.; Elansary, M.; Gori, A.-S.; et al. IBD Risk Loci Are Enriched in Multigenic Regulatory Modules Encompassing Putative Causative Genes. *Nat. Commun.* **2018**, *9*, 2427. [[CrossRef](#)]
16. Haberman, Y.; Tickle, T.L.; Dexheimer, P.J.; Kim, M.-O.; Tang, D.; Karns, R.; Baldassano, R.N.; Noe, J.D.; Rosh, J.; Markowitz, J.; et al. Pediatric Crohn Disease Patients Exhibit Specific Ileal Transcriptome and Microbiome Signature. *J. Clin. Investig.* **2014**, *124*, 3617–3633. [[CrossRef](#)]
17. Vancamelbeke, M.; Vanuytsel, T.; Farré, R.; Verstockt, S.; Ferrante, M.; Van Assche, G.; Rutgeerts, P.; Schuit, F.; Vermeire, S.; Arijis, I.; et al. Genetic and Transcriptomic Bases of Intestinal Epithelial Barrier Dysfunction in Inflammatory Bowel Disease. *Inflamm. Bowel. Dis.* **2017**, *23*, 1718–1729. [[CrossRef](#)]
18. Heianza, Y.; Sun, D.; Ma, W.; Zheng, Y.; Champagne, C.M.; Bray, G.A.; Sacks, F.M.; Qi, L. Gut-microbiome-related LCT genotype and 2-year changes in body composition and fat distribution: The POUNDS Lost Trial. *Int. J. Obes.* **2018**, *42*, 1565–1573. [[CrossRef](#)]

19. Nowak, J.K.; Lindstrøm, J.C.; Kalla, R.; Ricanek, P.; Halfvarson, J.; Satsangi, J. Age, Inflammation, and Disease Location Are Critical Determinants of Intestinal Expression of SARS-CoV-2 Receptor ACE2 and TMPRSS2 in Inflammatory Bowel Disease. *Gastroenterology* **2020**, *159*, 1151–1154.e2. [[CrossRef](#)] [[PubMed](#)]
20. Fang, P.; Xiang, L.; Chen, W.; Li, S.; Huang, S.; Li, J.; Zhuge, L.; Jin, L.; Feng, W.; Chen, Y.; et al. LncRNA GAS5 Enhanced the Killing Effect of NK Cell on Liver Cancer through Regulating MiR-544/RUNX3. *Innate Immun.* **2019**, *25*, 99–109. [[CrossRef](#)] [[PubMed](#)]
21. Kuokkanen, M.; Enattah, N.S.; Oksanen, A.; Savilahti, E.; Orpana, A.; Järvelä, I. Transcriptional Regulation of the Lactase-Phlorizin Hydrolase Gene by Polymorphisms Associated with Adult-Type Hypolactasia. *Gut* **2003**, *52*, 647–652. [[CrossRef](#)] [[PubMed](#)]
22. Araujo, E.M.; Dos Santos, L.; Coutinho, R.; Assis, V.; Brandão, N.; Almeida, D.; Conceição, G.; Figueredo, C.; Fonseca, H.; Lima, M.L.; et al. Genetic and Oral Tests for the Diagnosis of Lactose Intolerance in Mixed-Ancestry Brazilians with Metabolic Syndrome. *Lifestyle Genom.* **2019**, *12*, 1–9. [[CrossRef](#)] [[PubMed](#)]
23. Gorski, M.M.; Blighe, K.; Lotta, L.A.; Pappalardo, E.; Garagiola, I.; Mancini, I.; Mancuso, M.E.; Fasulo, M.R.; Santagostino, E.; Peyvandi, F. Whole-Exome Sequencing to Identify Genetic Risk Variants Underlying Inhibitor Development in Severe Hemophilia A Patients. *Blood* **2016**, *127*, 2924–2933. [[CrossRef](#)] [[PubMed](#)]
24. Tomczonek-Moruś, J.; Wojtasik, A.; Zeman, K.; Smolarz, B.; Bąk-Romaniszyn, L. 13910C>T and 22018G>A LCT Gene Polymorphisms in Diagnosing Hypolactasia in Children. *United Eur. Gastroenterol. J.* **2019**, *7*, 210–216. [[CrossRef](#)] [[PubMed](#)]
25. Labrie, V.; Buske, O.J.; Oh, E.; Jeremian, R.; Ptak, C.; Gasiūnas, G.; Maleckas, A.; Petereit, R.; Žvirbliene, A.; Adamonis, K.; et al. Lactase Nonpersistence Is Directed by DNA-Variation-Dependent Epigenetic Aging. *Nat. Struct. Mol. Biol.* **2016**, *23*, 566–573. [[CrossRef](#)]
26. Oh, E.; Jeremian, R.; Oh, G.; Groot, D.; Susic, M.; Lee, K.; Foy, K.; Laird, P.W.; Petronis, A.; Labrie, V. Transcriptional Heterogeneity in the Lactase Gene within Cell-Type Is Linked to the Epigenome. *Sci. Rep.* **2017**, *7*, 41843. [[CrossRef](#)]
27. Kettunen, J.; Silander, K.; Saarela, O.; Amin, N.; Müller, M.; Timpson, N.; Surakka, I.; Ripatti, S.; Laitinen, J.; Hartikainen, A.-L.; et al. European Lactase Persistence Genotype Shows Evidence of Association with Increase in Body Mass Index. *Hum. Mol. Genet* **2010**, *19*, 1129–1136. [[CrossRef](#)] [[PubMed](#)]
28. Corella, D.; Arregui, M.; Coltell, O.; Portolés, O.; Guillem-Sáiz, P.; Carrasco, P.; Sorlí, J.V.; Ortega-Azorín, C.; González, J.I.; Ordovás, J.M. Association of the LCT-13910C>T Polymorphism with Obesity and Its Modulation by Dairy Products in a Mediterranean Population. *Obes. Silver Spring* **2011**, *19*, 1707–1714. [[CrossRef](#)]
29. Almon, R.; Álvarez-León, E.E.; Serra-Majem, L. Association of the European Lactase Persistence Variant (LCT-13910 C>T Polymorphism) with Obesity in the Canary Islands. *PLoS ONE* **2012**, *7*, e43978. [[CrossRef](#)]
30. Chen, M.-H.; Raffield, L.M.; Mousas, A.; Sakaue, S.; Huffman, J.E.; Moscati, A.; Trivedi, B.; Jiang, T.; Akbari, P.; Vuckovic, D.; et al. Trans-Ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. *Cell* **2020**, *182*, 1198–1213.e14. [[CrossRef](#)]
31. Ng, T.B.; Cheung, R.C.F.; Wong, J.H.; Wang, Y.; Ip, D.T.M.; Wan, D.C.C.; Xia, J. Antiviral Activities of Whey Proteins. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 6997–7008. [[CrossRef](#)] [[PubMed](#)]
32. Cederlund, A.; Kai-Larsen, Y.; Printz, G.; Yoshio, H.; Alvelius, G.; Lagercrantz, H.; Strömberg, R.; Jörnvall, H.; Gudmundsson, G.H.; Agerberth, B. Lactose in Human Breast Milk an Inducer of Innate Immunity with Implications for a Role in Intestinal Homeostasis. *PLoS ONE* **2013**, *8*, e53876. [[CrossRef](#)] [[PubMed](#)]
33. Dharuri, H.; 't Hoen, P.A.C.; van Klinken, J.B.; Henneman, P.; Laros, J.F.J.; Lips, M.A.; el Bouazzaoui, F.; van Ommen, G.-J.B.; Janssen, I.; van Ramshorst, B.; et al. Downregulation of the Acetyl-CoA Metabolic Network in Adipose Tissue of Obese Diabetic Individuals and Recovery after Weight Loss. *Diabetologia* **2014**, *57*, 2384–2392. [[CrossRef](#)] [[PubMed](#)]
34. Taminelli, G.L.; Sotomayor, V.; Valdivieso, Á.G.; Teiber, M.L.; Marín, M.C.; Santa-Coloma, T.A. CISD1 Codifies a Mitochondrial Protein Upregulated by the CFTR Channel. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 856–862. [[CrossRef](#)] [[PubMed](#)]