Analysis of COL1A1_1 gene (rs1107946) polymorphism as a risk factor for low birth weight

T. Ye. Shumna*E, T. O. Levchuk*AD, O. M. KamysnyiF

Zaporizhzhia State Medical University, Ukraine

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Objective. Identification of the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene and the pattern of allele distribution in low birth weight babies.

Materials and methods. A total of 168 babies were examined. The babies were divided into 3 groups depending on the birth weight: the 1st group comprised of 52 babies (birth weight was 1500–1999 g), the 2nd group – 76 babies (birth weight was 2000–2499 g) and the 3rd group – 40 babies (birth weight was more than 2500 g, that is they had normal birth weight). Polymerase chain reaction genotyping method was used.

Results. It was found that the frequency of the C allele detection was equal to 39.60 %, the A allele – 60.42 %, chi-square (df = 1) 29.17, P < 0.05. At the same time, the homozygous AA genotype was observed significantly more often and amounted to 52.98 % versus 32.14 % of the CC genotype cases. The heterozygous CA genotype was detected only in 14.9 % of children, significantly less than homozygous genotypes CC (df = 1) 13.92, P < 0.05 and AA (df = 1) 54.38, P < 0.05. The AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene was found among babies of the 1st and 2nd groups in 61.53 % and 52.63 %, CC – 23.08 % and 31.58 %, CA – 15.38 % and 15.79 % of cases, respectively. The CC genotype of the polymorphism was detected almost in half of babies from the 3rd group (47.5 %), while the AA genotype was detected only in 35.0 % and the CA genotype – in 17.5 %.

Conclusions. The molecular and genetic study of the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene showed that the determination of the A allele frequency was significantly higher than the C allele among the examined babies. Consequently, the homozygous AA genotype was significantly more common than the CC genotype. The results of the study indicated the prognostic value of the A allele for the risk of low birth weight – that is, the lower birth weight (1500–1999 g) was found in babies with homozygous AA genotype.

Key words: allelic genes, genotype, collagen, thinness, babies.
Birth weight is an important parameter that influences later health of a newborn. But on the other hand, the low birth weight is a consequence of premature birth or intrauterine growth retardation, which can lead to problems of adaptation to extrauterine life, severe postpartum complications and even death within the first 5 years of life. In case of the low birth weight, there is also a risk of developing chronic kidney disease, hypertension, obesity, neuropsychiatric symptoms, cognitive decline and autistic spectrum disorders at older ages [1–11].

Therefore, the topical issue of pediatrics is the study of low birth weight causes and the problem of “low weight” itself is polyethiological one [12,13].

Indeed, today it is known that all risk factors for a child’s low birth weight can be divided into intrauterine, placental, parental and genetic both from mother and father. That is, the weight and length of a fetus body and then of a newborn child are determined by genetic inheritance from both parents [14–16].

At the present stage, the data on the study of some polymorphisms affecting various risk factors for low birth weight, namely endocrine mechanisms [17], folic acid metabolism [18,19], vitamin D metabolism [20], metals exposure [21], cytokine profile changes [22], the influence of caffeine [23], etc., can most often be found in the literature. Although existing information about genetic factors caused by one or another gene polymorphism is not sufficient to study this problem.

However Y. V. Alegina and co-authors in their study demonstrated the association of COL1A1 gene polymorphism with miscarriage, but we did not meet the data on the incidence rate of this gene polymorphism in children born with low body weight, necessitating further our study [24]. After all, structural integrity of connective tissues is upheld by collagen, which is the most abundant protein in the human body, up to 25% of the total body protein. The word «collagen» means in Greek “glue”. At present, scientists know 28 types of collagen, but the most common in the human body is type 1 collagen. It is a part of many tissues: skin, ligaments, bones, cornea, placenta, arteries, liver, dentin, and the like. Type 1 collagen fibers have the highest mechanical strength among all collagen types. Collagen synthesis is regulated by about 40 genes. Therefore, virtually any gene mutation in collagen synthesis leads to loss or alteration of this protein function that, in turn, affects the properties of tissues and organs, including the “strength” of tissues. The gene COL1A1 encodes the component of type 1 collagen. The collagen molecule typically consists of three protein chains interlaced with each other. The main protein of the collagen includes two chains of collagen α1 and one chain of collagen α2 [25].

Objective
Identification of the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene and the pattern of allele distribution in low birth weight babies.

Materials and methods
To study the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene, babies with different birth weight were examined. A total of 168 children were divided into 3 groups according to the birth weight: the 1st group consisted of 52 babies (birth weight was 1500–1999 g), the 2nd group included 76 babies (birth weight was 2000–2499 g), the 3rd group – 40 babies (birth weight was more than 2500 g).

Polymerase chain reaction genotyping was performed with the Applied Biosystems (USA) using total DNA samples extracted from whole venous blood with a set of SNP-Screen reagents (manufacturer “Syntol”) on the amplifier CFX96TM Real-Time PCR Detection Systems (Bio-Radlaboratories, Inc., USA). This study was carried out in the Department of Molecular and Genetic Researches of the Educational Medical and Laboratory Center at the Microbiology Department of Zaporizhzhia State Medical University in Zaporizhzhia city (Head of the Microbiology Department, Head of the Department of Molecular and Genetic Researches of the Educational Medical and Laboratory Center of Zaporizhzhia State Medical University – MD, Professor O. M. Kamyshnyi). The work was carried out within the framework of the scientific and research work of the Children Diseases Department of Zaporizhzhia State Medical University.

The obtained results of the studied distribution of allele frequencies and genotypes were used to analyze the genetic structure of the population according to the Hardy-Weinberg equilibrium. To compare the allele frequencies and genotypes in different groups, the non-parametric statistical “2 × 2 Table” method, Chi-square (df = 1) test were used. Also, the odds ratio (OR) was calculated using a four-field table with a confidence interval (CI) construction by the Woolf method. To evaluate the diagnostic significance, the indicators such as sensitivity, specificity, accuracy and prognostic value of positive and negative results were determined. To
Results

A molecular and genetic study of 168 babies was performed to detect the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene, which showed that the frequency of C allele detection was equal to 39.6 %, A allele detection – 60.42 %. Chi-square (df = 1) 29.17, P < 0.05. At the same time, the homozygous AA genotype was observed significantly more often and amounted to 52.98 % versus 32.14 % of the CC genotype cases. The heterozygous CA genotype was detected only in 14.9 % of children, significantly less than homozygous genotypes CC (df = 1) 13.92, P < 0.05 and AA (df = 1) 54.38, P < 0.05.

The AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene was found among babies of the 1st group in 61.53 % of cases. Moreover, the AA genotype was significantly more common than the CA genotype (df = 1), 23.40, P < 0.05, and the CC genotype (df = 1), 15.76, P < 0.05. The CC genotype was found in 23.08 % of the examined babies in the 1st group, that was significantly less than the AA genotype (df = 1), 15.76, P < 0.05. Among the 1st group babies, the heterozygous CA genotype was detected in 15.38 %, although the CA genotype was less common than the CC genotype, no significant differences were found. These data are shown in Fig. 1.

Among the 2nd group babies, the AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene prevailed, which was recorded in 52.63 % of babies (Fig. 2).

However, among the 2nd group babies, the AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene was significantly more common than the CA genotype (df = 1), 23.40, P < 0.05, and the CC genotype (df = 1), 15.76, P < 0.05. Among the 3rd group babies, the heterozygous CA genotype was detected in 44 %, that was significantly less than the AA genotype (df = 1), 37 %, P < 0.05. Although the CA genotype was less common than the CC genotype, no significant differences were found. These data are shown in Fig. 3.

The AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene was significantly more common than the CA genotype (df = 1), 23.40, P < 0.05, and the CC genotype (df = 1), 15.76, P < 0.05. Among the 1st group babies, the heterozygous CA genotype was detected in 15.38 %, although the CA genotype was less common than the CC genotype, no significant differences were found. These data are shown in Fig. 1.

Among the 2nd group babies, the AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene prevailed, which was recorded in 52.63 % of babies (Fig. 2).

However, among the 2nd group babies, the AA genotype of the (rs1107946) polymorphism of COL1A1_1 gene was significantly more common than the CA genotype (df = 1), 23.40, P < 0.05, and the CC genotype (df = 1), 15.76, P < 0.05. Among the 3rd group babies, the heterozygous CA genotype was detected in 44 %, that was significantly less than the AA genotype (df = 1), 37 %, P < 0.05. Although the CA genotype was less common than the CC genotype, no significant differences were found. These data are shown in Fig. 3.
### Allele C, %

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele C, %</td>
<td>56.25</td>
<td>39.50</td>
<td>30.80</td>
</tr>
<tr>
<td>Allele A, %</td>
<td>43.75</td>
<td>60.50</td>
<td>69.20</td>
</tr>
</tbody>
</table>

### Global population
- Africa: 25.8, 69.3
- America: 30.0, 70.0
- East Asia: 30.4, 69.6
- South Asia: 23.6, 76.4
- Europe: 13.8, 86.2
- Children from our research: 60.4, 39.6

### Comparative characteristics of the allele frequency of the COL1A1_1 gene (rs1107946) polymorphism in the world population.

### Comparative characteristic of the AA, CA, CC genotypes frequency of the COL1A1_1 gene (rs1107946) polymorphism in the world population.

<table>
<thead>
<tr>
<th></th>
<th>AA, %</th>
<th>CA, %</th>
<th>CC, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global population</td>
<td>7.00</td>
<td>37.50</td>
<td>55.50</td>
</tr>
<tr>
<td>Africa</td>
<td>9.40</td>
<td>42.70</td>
<td>48.00</td>
</tr>
<tr>
<td>America</td>
<td>8.90</td>
<td>42.10</td>
<td>49.00</td>
</tr>
<tr>
<td>East Asia</td>
<td>9.90</td>
<td>40.90</td>
<td>49.20</td>
</tr>
<tr>
<td>South Asia</td>
<td>5.90</td>
<td>35.40</td>
<td>58.70</td>
</tr>
<tr>
<td>Europe</td>
<td>0.80</td>
<td>26.00</td>
<td>73.20</td>
</tr>
<tr>
<td>Children from our research</td>
<td>58.98</td>
<td>14.90</td>
<td>32.14</td>
</tr>
</tbody>
</table>
(df = 1), 6.91, P < 0.05, and CC genotype (df = 1), 22.92, P < 0.05. Thus, the CC genotype was found in 31.58 % of the examined babies in the 2nd group, that was significantly more common than the CA genotype (df = 1), 5.24, P < 0.05 and significantly less than the AA genotype (df = 1), 6.91, P < 0.05. The CA genotype was detected only in 15.79 % of the 2nd group babies, which was significantly less common than the AA genotype (df = 1), 5.24, P < 0.05, and CC genotype (df = 1), 22.92, P < 0.05. The CC genotype of the (rs1107946) polymorphism of COL1A1_1 gene frequency was significantly higher than the CA genotype (df = 1), 7.94, P < 0.05, in almost half of babies from the 3rd group (47.5 %). The CA genotype was detected in 17.5 % of babies, which was significantly less common than the CC genotype (df = 1), 7.94, P < 0.05. Among babies of the 3rd group, the AA genotype was revealed in 35 % of cases. The data are presented in Fig. 3. The following comparative analysis was done to detect the frequency of each genotype of (rs1107946) polymorphism of type 1 collagen (COL1A1_1) gene not only within each of the examined groups of babies, but also between the study groups, that is, we analyzed the frequency of each genotype detection depending on the birth weight. The CC genotype was significantly greater among the 2nd group babies (43.63 %), (df = 1), 5.95, P < 0.05, than among babies of the 1st group (22.82 %). There was no significant difference in comparison with the control group (Fig. 4 and Table 1).

Among the 1st group babies, the genotype CA was found in 29.63 %, among the 2nd group babies – in 44.44 %, among babies of the 3rd group – in 25.93 %; there were no significant differences between the three groups (Fig. 5 and Table 1).

The low percentage of the AA genotype was revealed among the 3rd group babies (16.28 %) as compared with the 1st group babies (df = 1), 9.56, P < 0.05 and the babies of the 2nd group (df = 1), 18.14, P < 0.05. Among the 1st group babies, the AA genotype was determined in 37.21 %, among the 2nd group – in 46.51 % (Fig. 6 and Table 1).

According to the Hardy-Weinberg equilibrium, the frequency of alleles at the CA (rs1107946) polymorphism of COL1A1_1 in the examined babies (Fig. 7) showed that the A allele detection was significantly higher in the 1st group babies (birth weight of 1500–1999 g) amounting to 69.2 % of cases (df = 1), 17.5, P < 0.05 and in the 2nd group (babies with the birth weight of 2000–2499 g) – 60.5 % (df = 1), 37.57, P < 0.05 in comparison to babies with birth weight of more than 2500 g (the 3rd group) – 43.75 %.

Also, the A allele frequency was significantly higher among the 1st group babies than among the 2nd group babies, (df = 1), 4.15, P < 0.05. According to the data of the multiplicative inheritance models, the distribution of the allele frequencies among babies also proved that the inheritance of a phenotypic feature such as a low birth weight was associated with the predominance of A allele of COL1A1_1 gene (rs1107946) polymorphism (Table 2).

But the homozygous AA genotype with a reliable odds ratio = 2.97, DI [1.26–7.00] was found only among the 1st group babies (birth weight of 1500–1999 g). Such pattern was not observed in case of the CA and CC genotypes.

### Table 1. Characterization of genotypes of the COL1A1_1 gene (rs1107946) polymorphism (abs./%)

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Genotypes n (abs/%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CC (55/100 %)</td>
</tr>
<tr>
<td>I</td>
<td>52</td>
<td>12/22.82 %</td>
</tr>
<tr>
<td>II</td>
<td>76</td>
<td>2/4.33 %</td>
</tr>
<tr>
<td>p (I – II)</td>
<td>&lt;0.05</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>III</td>
<td>40</td>
<td>12/30.56 %</td>
</tr>
<tr>
<td>p (I – III)</td>
<td>&gt;0.05</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>p (II – III)</td>
<td>&gt;0.05</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

### Table 2. Distribution of the allele frequencies among the babies according to multiplicative inheritance models

<table>
<thead>
<tr>
<th>Groups</th>
<th>Babies n</th>
<th>Alleles n</th>
<th>Alleles (abs/%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>I</td>
<td>52</td>
<td>104</td>
<td>72/69.23 %</td>
</tr>
<tr>
<td>II</td>
<td>76</td>
<td>152</td>
<td>92/66.53 %</td>
</tr>
<tr>
<td>p (I – II)</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>III</td>
<td>40</td>
<td>80</td>
<td>35/43.75 %</td>
</tr>
<tr>
<td>p (I – III)</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>p (II – III)</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&gt;0.05</td>
</tr>
</tbody>
</table>

### Discussion

We compared the results of our study with population-based studies. The frequency of the AA genotype detection in children enrolled in the study was significantly higher than in the world population [26]. The same pattern was observed in relation to the A allele. The opposite pattern was found regarding the CC and CA genotypes, as these genotypes were determined less frequently among the children in our study than in the population (Fig. 8, 9). This can be explained by the sample representativeness in our study, that is, children weighing less than 2500 g at birth. The study results showed the prognostic value of the A allelic gene for the risk of low birth weight, that is babies who were homozygous for the AA genotype had lower birth weight (1500–1999 g).

As a result of the study, it has been found that all low birth weight babies were carriers of the A allele. For instance, among babies carrying the A allele, the odds ratio was equal to 2.89, DI [1.58–5.31] in the 1st group and 1.97, DI [1.14–3.41] in the 2nd group.

Analyzing the study results, we calculated the prognostic value of the A allele detection among the low birth weight babies: sensitivity – 0.63, specificity – 0.55, positive prognostic value 0.82 and negative prognostic value – 0.69.

### Conclusions

1. The molecular and genetic study of the CA genotype of the (rs1107946) polymorphism of COL1A1_1 gene showed that the A allele frequency was significantly higher than the C allele and was equal to 60.42 % and 39.6 %, respectively. At the same time, the homozygous AA genotype was detected significantly more often and amounted to 52.98 % versus 32.14 % of the CC genotype cases, P < 0.05; the heterozygous CA genotype was determined only in 14.9 % of babies.

2. Depending on the birth weight, the genotypes had the following distribution: the homozygous CC genotype – 22.88 %, 43.63 %, 34.55 %, the homozygous AA geno-
type –37.21 %, 46.51 %, 16.28 % и the heterozygous CA genotype – 29.63 %, 44.44 %, 25.93 %. These data are not consistent with the population-based studies due to the representativeness of the birth weight sample. 3. It was found that the babies of the 1st and 2nd groups had the А allele more frequently with the odds ratio equal to 2,89, DI [1.58–5.31] and 1,97, DI [1.14–3.41], respectively. The results of the study showed the prognostic value of the А allele for the risk of low birth weight. Prospects for further studies. In the future, we are planning to study the effect of other collagen gene polymorphisms as well as their effect on the low birth weight.

Conflicts of interest: authors have no conflict of interest to declare.

Конфлікт інтересів: відсутній.

References


