

**Supplementary information for the article:**

Spasovski, V.; Srzentić Dražilov, S.; Nikčević, G.; Baščarević, Z.; Stojiljković, M.; Pavlović, S.; Spasovski, D. Molecular Biomarkers in Perthes Disease: A Review. *Diagnostics* **2023**, *13* (3), 471. <https://doi.org/10.3390/diagnostics13030471>.



## Protein sequence alignment of COL2A1 protein. COL2A1 protein show high conservation of primary structure among distant

### Supplementary material

<https://www.uniprot.org>

CLUSTAL O(1.2.4) multiple sequence alignment

```
SP | P02458 | CO2A1_HUMAN  MIRLGAPQTLVLL---TLLVAAVLRCQG-QDVQEAGSCVQDGQRYNDKDVWKPEPCRICV 56
SP | P28481 | CO2A1_MOUSE  MIRLGAPQSLVLL---TLLIAAVLRCQG-QDAQEAGSCLQNGQRYKDKDVWKPSSCRICV 56
SP | Q91717 | CO2A1_XENLA  MFSFVDSRTLVLFAATQVILLAVVRCQDEEDVLDTGSCVQHGRYSKDVWKPEPCQICV 60
TR | Q2LDA1 | Q2LDA1_DANRE  MFRL LDSRTL LLLVATHSVLLSLVRCQEQEDDQEEFGGCVQDGQQYADRAVWKPEPCRVCV 60
      *:  :      :*:*:      : : :*:**  :*  : *.*:*.**:* * : *****. *::**

SP | P02458 | CO2A1_HUMAN  CDTGTVLCDDIICEDVKDCL SPEIPFGECCPICPTDLAT-ASGQPGPKGQKGE PGDIKDI 115
SP | P28481 | CO2A1_MOUSE  CDTGNVLCDDIICED- PDCLNPEIPFGECCPICPADLAT-ASGKLGPKGQKGE PGDIRDI 114
SP | Q91717 | CO2A1_XENLA  CDTGTVLCDDIICEESKDCPN AEIPFGECCPICPTEQSSTSSGQV LKGQKGE PGDIKDV 120
TR | Q2LDA1 | Q2LDA1_DANRE  CDSGTVLCDEVICEDLND CANPIISPGECCPICPADTDD-PIGSLGAKGQKGE PGDITDV 119
      **:*.*****:.**:  ** . * *****:;      * . ***** **

SP | P02458 | CO2A1_HUMAN  VGPKGPPGPQGPAGEQGP RGRDGDKGEK GAPGPRGRDGE PGTPGNP GPPGPPGPPGPPGL 175
SP | P28481 | CO2A1_MOUSE  IGPRGPPGPQGPAGEQGP RGRDGDKGEK GAPGPRGRDGE PGTPGNP GPAGPPGPPGPPGL 174
SP | Q91717 | CO2A1_XENLA  LGPRGPPGPQGPSGEQGS RGERGDKGEK GAPGPRGRDGE PGTPGNP GPV--GPPGPPGL 177
TR | Q2LDA1 | Q2LDA1_DANRE  VGPRGPAGPMGPPGEQGT RGERGAKGEKGS PGPRGRDGE PGTPGNP GPPGPPGPPGPPGL 179
      :*:* ** ** ** ** ** ** ** ** ** ** ** *:* ** ** ** ** *:* ** ** ** *:* ** ** ** *:* ** **

SP | P02458 | CO2A1_HUMAN  G-GNFAAQMAGGFDEKAGGAQLGVMQGPMGPMGPRGPPG PAGAPGPQGFQGNPGE PGEPG 234
SP | P28481 | CO2A1_MOUSE  SAGNFAAQMAGGYDEKAGGAQMGVMQGPMGPMGPRGPPG PAGAPGPQGFQGNPGE PGEPG 234
SP | Q91717 | CO2A1_XENLA  G-GNFAAQMAGGFDEKAGGAQMGVMQGPMGPMGPRGPPG PGTGAPGPQGFQGNPGE PGEPG 236
TR | Q2LDA1 | Q2LDA1_DANRE  G-GNFAAQMAGGFDEKAGGAQMGVMQGPMGPMGPRGPPG PPSGAPGPQGFQGNPGE TGE PG 238
      . *****.*:.******.******.******.******.******.******

SP | P02458 | CO2A1_HUMAN  VSGPMGPRGPPGPPGPKPGDDGEAGKPGKAGERGPPGPQ GARGFPGT PGLPGVKGHRGYPG 294
SP | P28481 | CO2A1_MOUSE  VSGPMGPRGPPGPPGPKPGDDGEAGKPGKSGERGLPGPQ GARGFPGT PGLPGVKGHRGYPG 294
SP | Q91717 | CO2A1_XENLA  AGGPMGPRGPPGPPSGKPGDDGEAGKPGKSGERGPPGPQ GARGFPGT PGLPGVKGHRGYPG 296
TR | Q2LDA1 | Q2LDA1_DANRE  PAGALGPRGPPGPPGPKPGSDGEAGKPGKAGERGPPGPQ GARGFPGT PGLPGIKGHRGHPG 298
```

. \* :\*\*\*\*\* \*\*\*\*.\*\*\*\*\*:\*\*\*\* \*\*\*\*\*:\*\*\*\*\*:\*\*\*\*:\*\*\*\*

|    |        |              |  |     |
|----|--------|--------------|--|-----|
| SP | P02458 | CO2A1_HUMAN  | LDGAKGEAGAPGVKGESGSPGENSGPMPGPRGLPGERGRTGPAGAAGARGNDGQPPGAG  | 354 |
| SP | P28481 | CO2A1_MOUSE  | LDGAKGEAGAPGVKGESGSPGENSGPMPGPRGLPGERGRTGPAGAAGARGNDGQPPGAG  | 354 |
| SP | Q91717 | CO2A1_XENLA  | LDGAKGEAGAAGAKGEGGATGEAGSPGMPGPRGLPGERGRPGSSGAAGARGNDGLPGPAG | 356 |
| TR | Q2LDA1 | Q2LDA1_DANRE | LDGAKGEAGAAGAKGEGSNGESGAPGMPGPRGLPGERGRPGATGAAGARGNDGLPGPAG  | 358 |

\*\*\*\*\* \* .\*\*\*. \*: \*\* \* :\*\*\*\*\* \* .\*\*\*\*\* \*\*\*\*\*

|    |        |              |  |     |
|----|--------|--------------|--|-----|
| SP | P02458 | CO2A1_HUMAN  | PPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQQPRGEPGTPGSPGPAGASGNPGTDG | 414 |
| SP | P28481 | CO2A1_MOUSE  | PPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQQSRGEPGNPGSPGPAGASGNPGTDG | 414 |
| SP | Q91717 | CO2A1_XENLA  | PPGPVGPAGAPGFPAGPSKGEAGPTGARGPEGAQQPRGESGTPGSPGPAGASGNPGTDG  | 416 |
| TR | Q2LDA1 | Q2LDA1_DANRE | PPGPVGPAGAPGFPSPGSKGEAGPTGARGPEGAQQPRGEAGTPGSPGPAGASGNPGTDG  | 418 |

\*\*\*\*\* .\*\*\*\*\*:\*\*\*:\*\*\*\*\* \*\*\*\*\* \* .\*\*\*\*\* \*\*\*\*\*

|    |        |              |  |     |
|----|--------|--------------|--|-----|
| SP | P02458 | CO2A1_HUMAN  | IPGAKGSAGAPGIAGAPGFPGRGPPGPQGATGPLGPKGQTGEPGIAGFKGEQGPKEGEPG | 474 |
| SP | P28481 | CO2A1_MOUSE  | IPGAKGSAGAPGIAGAPGFPGRGPPGPQGATGPLGPKGQAGEPGIAGFKGDQGPKEGEPG | 474 |
| SP | Q91717 | CO2A1_XENLA  | IPGAKGSSGGPGIAGAPGFPGRGPPGPQGATGPLGPKGQTGDPGVAGFKGEQGPKEIG   | 476 |
| TR | Q2LDA1 | Q2LDA1_DANRE | IPGAKGSAGASGIAGAPGFPGRGPPGPQGATGPLGPKGQSGDPIPGFKGEAGPKGERG   | 478 |

\*\*\*\*\*:\*. \*\*\*\*\*:\*\*\*:\*\*\*\*: \*\*\*\*\* \*

|    |        |              |   |     |
|----|--------|--------------|---|-----|
| SP | P02458 | CO2A1_HUMAN  | PAGPQGAPGPAGEEGKRGARGEPPGVPIGPPGERGAPGNRGFPQDGLAGPKGAPGERG  | 534 |
| SP | P28481 | CO2A1_MOUSE  | PAGPQGAPGPAGEEGKRGARGEPPGAGPIGPPGERGAPGNRGFPQDGLAGPKGAPGERG | 534 |
| SP | Q91717 | CO2A1_XENLA  | SAGPQGAPGPAGEEGKRGARGEPPAAGPNPPGERGAPGNRGFPQDGLAGPKGAPGERG  | 536 |
| TR | Q2LDA1 | Q2LDA1_DANRE | VLGPQGPSPGSEEGKRGPRGEPGSAGPLGPPGERGAPGNRGFPQDGLAGKAGAPDRG   | 538 |

\*\*\*\* \* :\*\*\*\*\* \*\*\*\*\* . \*\* \*\*\*\*\* \*\*\*\*\*:\*\*\*

|    |        |              |  |     |
|----|--------|--------------|--|-----|
| SP | P02458 | CO2A1_HUMAN  | PSGLAGPKGANGDPGRPGEPGLPGARGLTGRPGDAGPQKVGPSGAPGEDGRPGPPGPQ | 594 |
| SP | P28481 | CO2A1_MOUSE  | PSGLTGPKANGDPGRPGEPGLPGARGLTGRPGDAGPQKVGPSGAPGEDGRPGPPGPQ  | 594 |
| SP | Q91717 | CO2A1_XENLA  | VPGLGGPKGNGDPGRPGEPGLPGARGLTGRPGDAGPQKVGPSGASGEDGRPGPPGPQ  | 596 |
| TR | Q2LDA1 | Q2LDA1_DANRE | VPGLSGPKGGTDPGRPGEPGLPGARGLTGRPGDAGAQQKVGATGAPGEDGRPGPPPLG | 598 |

\*\* \*\*\*\* . \*\*\*\*\* \*\*\*\*\* :\*\* \*\*\*\*\* \*

|    |        |              |   |     |
|----|--------|--------------|---|-----|
| SP | P02458 | CO2A1_HUMAN  | ARGQPGVMGFPGPKGANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGERG  | 654 |
| SP | P28481 | CO2A1_MOUSE  | ARGQPGVMGFPGPKGANGEPGKAGEKGLAGAPGLRGLPGKDGETGAAGPPGPSGAPAGERG | 654 |
| SP | Q91717 | CO2A1_XENLA  | ARGQPGVMGFPGPKGANGEPGKAGEKGLVAPGLRGLPGKDGETGSQGNPAGPAGERG     | 656 |
| TR | Q2LDA1 | Q2LDA1_DANRE | ARGQPGVMGFPGPKGANGEPGKPEKGLVGRGTGLRGLPGKDGETGPSGPPGPVAVGERG   | 658 |

\*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\* \*\* \* .\*\*\*\*

|    |        |             |   |     |
|----|--------|-------------|---|-----|
| SP | P02458 | CO2A1_HUMAN | EQGAPGPSGFQGLPGPPGPPGEGGKPGDQGVPEAGAPLVGPRGERGFPGERGSPPAQ | 714 |
| SP | P28481 | CO2A1_MOUSE | EQGAPGPSGFQGLPGPPGPPGEGGKQGDQGIPEAGAPLVGPRGERGFPGERGSPPAQ | 714 |
| SP | Q91717 | CO2A1_XENLA | EQGPPGPSGFQGLPGPPGSPGEGGKPGDQGVPEAGAPLVGPRGERGFPGERSSGPQ  | 716 |

|    |        |              |   |      |
|----|--------|--------------|---|------|
| TR | Q2LDA1 | Q2LDA1_DANRE | EQQQPGPSGFQGLPGPTGAPGEPGKPGDQGVPGEGGAAGPTGPRGERGFPGERGGAGPQG<br>*** ***** * ** ** *****:***.*** * .*****. * **          | 718  |
| SP | P02458 | CO2A1_HUMAN  | LQGPRGLPGTPTDGPKGASGPAGPPGAQGGPGLQGMPPERGAAGIAGPKGDRGDVGEKG   | 774  |
| SP | P28481 | CO2A1_MOUSE  | LQGPRGLPGTPTDGPKGAAAGPDGPPGAQGGPGLQGMPPERGAAGIAGPKGDRGDVGEKG  | 774  |
| SP | Q91717 | CO2A1_XENLA  | LQGPRGLPGTPTDGPKGASGPSGPNQAQGGPGLQGMPPERGAAGISGPKGDRGDTGEKG   | 776  |
| TR | Q2LDA1 | Q2LDA1_DANRE | LQGPRGLPGTPTDGPKGAIGPAGAAGAQGGPGLQGMPPERGAAGISGAKGDRGDSGEKG<br>***** ** * *****.***: * ***** **                         | 778  |
| SP | P02458 | CO2A1_HUMAN  | PEGAPGKDGGRLTGP I GPPG PAGANGEKGEVGGPPG P AGSAGARGAPGERGETGPPG PAG  | 834  |
| SP | P28481 | CO2A1_MOUSE  | PEGAPGKDGGRLTGP I GPPG PAGANGEKGEVGGPPG P SGTGARGAPGERGETGPPG PAG   | 834  |
| SP | Q91717 | CO2A1_XENLA  | PEGASGKDGSRGLTGP I GPPG PAGPNGEKGESGSPG P PGI V GARGAPDRGENGPPG PAG   | 836  |
| TR | Q2LDA1 | Q2LDA1_DANRE | PEGAPGKDGSRGLTGP I GPPG P SGNKAKGETGPIGSIGAPGARGAPDRGEIGAPG PAG<br>**** * .*****. * ** ** * * * * *****.*** * ****      | 838  |
| SP | P02458 | CO2A1_HUMAN  | FAGPPGADGQPGAKGEQGEAGQKGDAGAPGPPGQPSGAPGPPGPTGVTGPKGARGAQGGP  | 894  |
| SP | P28481 | CO2A1_MOUSE  | FAGPPGADGQPGAKGDQGEAGQKGDAGAPGPPGQPSGAPGPPGPTGVTGPKGARGAQGGP  | 894  |
| SP | Q91717 | CO2A1_XENLA  | FAGPPGADGQSGLKGQGESGQKGDAGAPGPPGQPSGAPGPPGPTGVFGPKGARGAQGP  | 896  |
| TR | Q2LDA1 | Q2LDA1_DANRE | FAGPPGADGQPGNKGQGESGQKGDAGAPGPPGQPSGAPGPPGPTGVTGPKGARGAQGAP<br>***** * *:***:*****:***** ***** ***** *                  | 898  |
| SP | P02458 | CO2A1_HUMAN  | ATGFPGAAGRVP GPPG SNGNPGPPG P S GKD GPKGARGDSGPPGRAGEPGLQGPAGPPG  | 954  |
| SP | P28481 | CO2A1_MOUSE  | ATGFPGAAGRVP GPPG ANGNPGPAGPPG P KDGPKGVRGDSGPPGRAGDPGLQGPAGAPG   | 954  |
| SP | Q91717 | CO2A1_XENLA  | ATGFPGAAGRVP GPPG NPNPGPPG P SAGKEGPKGVRGDAGPPGRAGDPGLQGAAGAPG  | 956  |
| TR | Q2LDA1 | Q2LDA1_DANRE | ATGFPGAAGRVP GPPG NPNPGAAGPAGP S GKD GPKGVRGDAGPPGRAGDAGLRGPPGAPG<br>***** ** ***** ** * :**:* ** .***:*****: **:* * ** | 958  |
| SP | P02458 | CO2A1_HUMAN  | EKGEPGDDGPGSAGPPG P QGLAGQ R GIVGLPGQRGERGF PGLPGPSGEPGKQGAPGASG  | 1014 |
| SP | P28481 | CO2A1_MOUSE  | EKGEPGDDGPGSLDGPPG P QGLAGQ R GIVGLPGQRGERGF PGLPGPSGEPGKQGAPGASG   | 1014 |
| SP | Q91717 | CO2A1_XENLA  | EKGEPGEDGPGSPDGP P QGLSQRGIVGLPGQRGERGF PGLPGPSGEPGKQGGPSSG   | 1016 |
| TR | Q2LDA1 | Q2LDA1_DANRE | EKGAGEDGPPGPDG P S PAGLAGQ R GIVGLPGQRGERGF PGLPGPSGEPGKQGAPGSSG<br>**** *:*** * :** ** *:*****:*****.***.***           | 1018 |
| SP | P02458 | CO2A1_HUMAN  | DRGPPGVP GPPGLTG P AGE P GREGSPGADGPPGRDGAAGVKGDRGETGAVGAPGAPGPPG   | 1074 |
| SP | P28481 | CO2A1_MOUSE  | DRGPPGVP GPPGLTG P AGE P GREGSPGADGPPGRDGAAGVKGDRGETGALGAPGAPGPPG   | 1074 |
| SP | Q91717 | CO2A1_XENLA  | DRGPPGVP GPPGLTG P SGE P GREGNPGSDGPPGRDGTG I KGDRGETGPLGAPGAPGAPG  | 1076 |
| TR | Q2LDA1 | Q2LDA1_DANRE | DRGPPGVP GPPGLTG P AGE T GREGNPGSDGPPGRDGAAGVKGERGNTGPIGAPGAPGAPG<br>*****: ** *****.***:*****: **:* ** :***** **       | 1078 |
| SP | P02458 | CO2A1_HUMAN  | SPGPAGPTGKQDRGEAGA QG P M G P S P AGARGI Q G P Q P R G D K G E A G E P G E R G L K G H R G                              | 1134 |
| SP | P28481 | CO2A1_MOUSE  | SPGPAGPTGKQDRGEAGA QG P M G P S P AGARGI A G P Q P R G D K G E S G E Q G E R G L K G H R G                              | 1134 |



Protein sequence alignment of COL2A1 protein. COL2A1 proteins show high conservation of the primary structure among distant species. The figure shows CLUSTAL OMEGA (1.2.4) multiple sequence alignment of the regions of COL2A1 protein of different organisms: Human, *Mus musculus*, *Xenopus laevis*, and *Danio rerio* [Uniprot identifiers: P02458, P28481, Q91717, Q2LDA, respectively]. Highlighted amino acids show the mutations found in Perthes patients. c.638 G > A (p.Gly213Asp) (exon 9); c.2014 G > T (p.Gly672Cys) (exon 31) (NM\_001844.5) [50]; c.1888 G > A (p.Gly630Ser) (exon 29) (NM\_033150.3) [51]; c.3665 G > A (p.Gly1170Ser) (exon 50) (NM\_033150.3) [48,49].

48. Su, P.; Li, R.; Liu, S.; Zhou, Y.; Wang, X.; Patil, N.; Mow, C.S.; Mason, J.C.; Huang, D.; Wang, Y. Age at onset-dependent presentations of premature hip osteoarthritis, avascular necrosis of the femoral head, or Legg-Calvé-Perthes disease in a single family, consequent upon a p.Gly1170Ser mutation of COL2A1. *Arthritis Rheum.* **2008**, *58*, 1701–1706.

49. Miyamoto, Y.; Matsuda, T.; Kitoh, H.; Haga, N.; Ohashi, H.; Nishimura, G.; Ikegawa, S. A recurrent mutation in type II collagen gene causes Legg-Calvé-Perthes disease in a Japanese family. *Hum. Genet.* **2007**, *121*, 625–629. <https://doi.org/10.1007/s00439-007-0354-y>.

50. Kannu, P.; Irving, M.; Aftimos, S.; Savarirayan, R. Two Novel COL2A1 Mutations Associated with a Legg-Calvé-Perthes Disease-like Presentation. *Clin. Orthop. Relat. Res.* **2011**, *469*, 1785–1790. <https://doi.org/10.1007/s11999-011-1850-x>.

51. Li, N.; Yu, J.; Cao, X.; Wu, Q.-Y.; Li, W.-W.; Li, T.-F.; Zhang, C.; Cui, Y.-X.; Li, X.-J.; Yin, Z.-M.; et al. A novel p. Gly630Ser mutation of COL2A1 in a Chinese family with presentations of Legg-Calvé-Perthes disease or avascular necrosis of the femoral head. *PLoS ONE* **2014**, *9*, e100505.