

Figure S2

Comparison between CpPGM1A and CpPGM1B at nucleotide and protein levels, showing that *CpPGM1A* and *CpPGM1B* genes are tandemly duplicated. They are highly conserved at both protein and nucleotide levels

Gene arrangement in the chromosome

cgd2_3260 (CpPGM1A)

cgd2_3270 (CpPGM1B)

DNA sequence alignment

cgd2_3260	ATGG	-----
cgd2_3270	ATGATGAAAAATTTTGAAACATCTCTTTTCATTAGTCTTGTGAAGTCAGTATGTCAATGCAAAAGAAAGTGTGGGAATTTGAGATATTTGCAGAAAAATAATCTGATCGAACCTACCGATGGTCAAGCAAAAGTAGTTACCTCCAGGATGTTCA	
cgd2_3260	-----	
cgd2_3270	AAATTAATCAGGGAAGGCCCTAATTAAGAAGAAAGTGCCTAGTAAACAAGAAATGTCATTGAAACCGAAACCACTGAGACTTATTATTATTCAGATAAAGGCTTAGATGCTTGGAGTCTAATGAAAAATGATAATCTATTAGATATGTTAGTTC	
cgd2_3260	AGCAAGTTCAAAGTAAGCCTTATCTTGACAGAGCCAGGAACAAGTGGTCTAAGAAAGAAAAACCGTGTATTTATGGAAGGAACCTTACTTGGCTAACTTTATTGAATCATACTTCCAATCATTCCCACCTGAAAAATTTTGAAGGTGCAACACTT	
cgd2_3270	AGAAGTTCAAATATCTCCTTATCTTGATCAAAAGCCAGGAACAAGTGGTCTAAGAAAGAAAAACCGTGTATTTATGGAAGGAACCTTACTTGGCTAACTTTATTGAATCATACTTCCAATCATTCCCACCTGAAAAATTTTGAAGGTGCAACACTT	
cgd2_3260	TTGTAGCTGGCGATGGTAGATTCTTCTACCAGAAGCAATACAAATATTTCGAAATTTGTCAGCACATAAAGTAAAGGATTTGGACTGGTGTAAAGCTTATGTTCAACACCAGCAGGAGTGTATCATAAGAGAAAGAGAAGCGG	
cgd2_3270	TTGTAGCTGGCGATGGTAGATTCTTCTACCAGAAGCAATACAAATATTTCGAAATTTGTCAGCACATAAAGTAAAGGATTTGGACTGGTGTAAAGCTTATGTTCAACACCAGCAGGAGTGTATCATAAGAGAAAGAGAAGCGG	
cgd2_3260	TATTCAGTCGGTGAATCTTATTAAACAGCTTCACACAACCCAGGTGGTATTGATGAGGACTTTGGAGTTAAATTTAAACGGAAGAATGGTGGCCAGCTCAAGATTCGGTACCAATACAAATTTTTGAAATACAAAAAATAACATCATATA	
cgd2_3270	TATTCAGTCGGTGAATCTTATTAAACAGCTTCACACAACCCAGGTGGTATTGATGAGGACTTTGGAGTTAAATTTAAACGGAAGAATGGTGGCCAGCTCAAGATTCGGTACCAATACAAATTTTTGAAATACAAAAAATAACATCATATA	
cgd2_3260	AGAAGATTTCACTTCTTAAATGATTAAAGTAAATTTGGCGTGAGGAATTAATACCCAATCAATTTACCGTCGAAGTAATAGATACATCTGAAGATTTGGCTAAAATTAATGAAAAAGATATTGATTTCGAAAGATTCAAAATCTACTTAAT	
cgd2_3270	AGAAGATTTCACTTCTTAAATGATTAAAGTAAATTTGGCGTGAGGAATTAATACCCAATCAATTTACCGTCGAAGTAATAGATACATCTGAAGATTTGGCTAAAATTAATGAAAAAGATATTGATTTCGAAAGATTCAAAATCTACTTAAT	
cgd2_3260	AGAAAAGATTTTAAATGGCTTTTGATTCTATGATCGTGGTGGCTGGACCATATGCAAGAAAACCTCTTTTGGACGAATTTGGATTGCCAGAAAGTTCAATGTTACATTTGGAATCAAAAGCAGATTTTGGAGGGCTTCATCCTGATCCAAATCT	
cgd2_3270	AGAAAAGATTTTAAATGGCTTTTGATTCTATGATCGTGGTGGCTGGACCATATGCAAGAAAACCTCTTTTGGACGAATTTGGATTGCCAGAAAGTTCAATGTTACATTTGGAATCAAAAGCAGATTTTGGAGGGCTTCATCCTGATCCAAATCT	
cgd2_3260	TACTTATGAAAAAGACTTAGTTGAATTAATGAAAGTAAATCTCCAGAAAAAGTTGATAAAACACAGCCAGATTTTGGTGCTGCTGGAGATGGTACTGTGATAGAAATATGATTTGGTAAAGGATCTTTGTTACCCCTTCAGATCTGTTG	
cgd2_3270	TACTTATGAAAAAGACTTAGTTGAATTAATGAAAGTAAATCTCCAGAAAAAGTTGATAAAACACAGCCAGATTTTGGTGCTGCTGGAGATGGTACTGTGATAGAAATATGATTTGGTAAAGGATCTTTGTTACCCCTTCAGATCTGTTG	
cgd2_3260	CAATTATTGTCATCATATGCAAAAGAAGCAATCCCATACTTCTCAAAGGATTTGGCTGGTGTCTCAAGATCTATGCCTACTAGTACTTCGCTCAATAATGTAGCAGAAAAGCTCGGAATCCCATGCTATGAGGTACCAACAGGATGGAATACTTT	
cgd2_3270	CAATTATTGTCATCATATGCAAAAGAAGCAATCCCATACTTCTCAAAGGATTTGGCTGGTGTCTCAAGATCTATGCCTACTAGTACTTCGCTCAATAATGTAGCAGAAAAGCTCGGAATCCCATGCTATGAGGTACCAACAGGATGGAATACTTT	
cgd2_3260	GGTAATTTAATGGATGCAAAAAATGATTGACATATGTGGTGAAGAATCTTTTGGACAGGAAGTGACCATATTCGTGAAAAGGATGGATTATGGCAGTGTGGCTTGGTTATCGATTCTAGCCACACAAGAAATCCAGACCCCAACAAAGCATTGGT	
cgd2_3270	GGTAATTTAATGGATGCAAAAAATGATTGACATATGTGGTGAAGAATCTTTTGGACAGGAAGTGACCATATTCGTGAAAAGGATGGATTATGGCAGTGTGGCTTGGTTATCGATTCTAGCCACACAAGAAATCCAGACCCCAACAAAGCATTGGT	
cgd2_3260	ATCTGTTGAGGATACCTAGGGAATTCGAAAAACATATGGAAGAACTACTACACTAGATTGATTACGAATCAATTTGAAACAGAAAAGCTGATCAATCTTTAAACATTTAAATCTTTAATGGAAGATCATCAGAAATTAAGAGAAATCA	
cgd2_3270	ATCTGTTGAGGATACCTAGGGAATTCGAAAAACATATGGAAGAACTACTACACTAGATTGATTACGAATCAATTTGAAACAGAAAAGCTGATCAATCTTTAAACATTTAAATCTTTAATGGAAGATTCAGAAATTAAGAGAAATCA	
cgd2_3260	TTAAACCATATGGTATGATATCAAAATTAAGTATAACTTACATATAATGATCCAGTAGATAAATCTGTCTGCAAAAAATCAGGGTTTAAAGATTATCTTTTCAAGGATAACTCCAGAGTTGTTTCCGCCATCATCAGGTACAGGTCAGTAGGAGCA	
cgd2_3270	TCGGTTCAAGTGGGTTGAAGGTTAATGATTAAGTATAACTTACATATAATGATCCAGTAGATAAATCTGTCTGCAAAAAATCAGGGTTTAAAGATTATCTTTTCAAGGATAACTCCAGAGTTGTTTCCGCCATCATCAGGTACAGGTCAGTAGGAGCA	
cgd2_3260	ACAATTCGAATTTACATGAAAGAGCTGTGAGTGATCAACAAAAGTTAATAGCATCAACGAAATCTTGGCGATTGATTGAGATAGTTGAGAAAAAGATTAATTAAGAAGATCGACAGGAAGATAGGCCAATGTTATAACTTAA	
cgd2_3270	ACTATTAGGATATATGAGAGAGCTGTAAAAACCTCAGGAATTTGAAAAACAACCAACAGCTCTAAATCACTTAATTGAGATAGTTGAGAAAAAGATTAATTAAGAAGATTACTGGTAGAAGCAAGCTACAGTAATTACTTAA	

Protein sequence alignment

CpPGM1A	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
CpPGM1B	MMKILNLFSLVLVSQYVNNANESVGNLRYLKINLIEPTDGGQKVVTSQDVQINGRALIKKKVSVTRNVIEETETETYYYSDDKGLDAWSPNENDNLLDMLVQEQISPYLDQPKGTSGLRKKTRVFMEGTYLANFIESYFQSFPPENF														
CpPGM1A	160	170	180	190	200	210	220	230	240	250	260	270	280	290	300
CpPGM1B	EGATLLVAGDGRFFLPEAIQIIEIAAAHKVKRIWTGVNGLCSTPAGSAIIREREGLIAGVGILLTASHNPGGIDEDFGVKFNEKNGGPAQDSVTNTIFEITKKLTSYKKISLPKIDLSKIGVQELIPNQFTVEVIDTSEDWLKMKKIF														
CpPGM1A	310	320	330	340	350	360	370	380	390	400	410	420	430	440	450
CpPGM1B	DFQKIQNLNRKDFKMAFDSMHGVAGPYARKLFIDEFGLPESSLLHLESKPDFGGLHDPNPLYAKDLVELMKVKSPEKVDKSTPDFGAAGDGDGRNMLGKGFVTPSDSVAIISYAKEAIPYFSGLAGVSRSMPTSTSLNNVAEK														
CpPGM1A	460	470	480	490	500	510	520	530	540	550	560	570	580	590	600
CpPGM1B	LGIPCYEPTGWKYFGNLDMAKMDIDGEEESFGTSDHIREKDLWAVLAWLSILAKNPDPKPLVSVDITREFWKTYGRNRYTRFDYESIETEKADQFFKHLNSLMEDHOKLREIIPKYPIDIKLTDNFYNDPVDKSVAKNGLRF														
CpPGM1A	610	620	630	640	650	660	670								
CpPGM1B	IFQDNRVVFRLSGTGSVGATIRIYIEKTVSDQTKVNSTSNEILGDLIEIVEKKIKLKESTGRDRPTVIT														

antigen site