

Table 1 Suppl. The comparison of theoretical Mr and pI, number of amino acids and their composition and recombinant protein solubility of *Lhcb1* among the different plants. Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura,; Q4 - Baguazhou Shuiqin.

Plants	Mr [Da]	pI	Number of amino acids	Basic amino acids [%]	Acid amino acids [%]	Aromatic amino acids [%]	Aliphatic amino acids [%]	Protein solubility [%]
Lhcb1-Q1	266	28581.6	5.18	9	11	11	20	79.9
Lhcb1-Q2	264	28035.0	5.16	9	10	11	20	79.4
Lhcb1-Q3	264	28005.9	5.30	9	10	11	20	84.0
Lhcb1-Q4	264	28034.0	5.30	9	10	11	20	82.2
<i>Apium graveolens</i>	264	28076.0	5.31	9	10	11	20	83.1
<i>Arabidopsis thaliana</i>	265	28053.9	5.28	9	10	11	19	82.0
<i>Glycine max</i>	264	28024.9	5.27	9	9	11	20	83.1
<i>Medicago truncatula</i>	265	28284.4	5.96	10	9	11	20	84.4
<i>Populus trichocarpa</i>	264	28052.9	5.29	9	10	11	19	84.0
<i>Vitis vinifera</i>	264	27940.8	4.99	8	10	11	20	80.5
<i>Selaginella moellendorffii</i>	272	28800.6	5.30	9	10	11	17	83.8
<i>Physcomitrella patens</i>	267	28337.0	5.13	9	10	10	19	74.5
<i>Oryza sativa japonica</i>	261	27552.5	5.29	9	10	11	20	78.1
<i>Zea mays</i>	264	28046.1	5.46	9	10	11	20	78.3

Table 2 Suppl. The accession numbers of selected amino acid sequences.

Plants	Lhcb1	Lhcb2	Lhcb3	Lhcb4	Lhcb5	Lhcb6
<i>Apium graveolens</i>	CAA99993					
<i>Arabidopsis thaliana</i>	18403546	15224471	15239602	15241005	15235029	15218330
<i>Glycine max</i>	356559635	358248712	359807444	356496185	356559931	356525886
<i>Medicago truncatula</i>	357496387	357495319	357452427	357461189		357445799
<i>Populus trichocarpa</i>	224114357	224131818	224122902	224090986	224127848	224054274
<i>Vitis vinifera</i>	225463432	225447576	359495476	225440674	225459918	225447745
<i>Selaginella moellendorffii</i>	302754140	302810440	302765457	302784975	302811558	302815978
<i>Physcomitrella patens</i>	168003213	168009155	168053612	168040609	168068372	168009894
<i>Oryza sativa</i>	Os01t0600900-02	Os03t0592500-01	115472785	Os07t0558400-01	Os11t0242800-01	Os04t0457000-01
<i>Zea mays</i>	226502967	226492741	226509520	219362417	162461137	226501904

Fig. 1 Suppl. The nucleotide acid sequences of *LHCP* from four plants: Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura, and Q4 - Baguazhou Shuiqin. The *dark* areas are identical sequences, the *pink* are 75 % identical sequences, and the *green* are 50 % identical sequences. The *arrows* imply the primer pair, the *blue* one is the forward primer, and the *red* one is the reverse complement primer.

Q1	ATGGCTGCTTCAACAATGGCCCTCTCTCCAGCTTTAGCAAGAAAGCAAGTCAAAGTAGCTTCATCTTCATCTTCATC	80
Q2	ATGGCTGCTTCAACAATGGCCCTCTCTCCAGCTTTAGCAAGAAAGCAAGTCAAAGTAGCTTCATCTTCATCT.....	75
Q3	ATGGCTGCTTCAACAATGGCCCTCTCTCCAGCTTTAGCAAGAAAGCAAGTCAAAGTAGCTTCATCTTCATCT.....	75
Q4	ATGGCTGCTTCAACAATGGCCCTCTCTCCAGCTTTAGCAAGAAAGCAAGTCAAAGTAGCTTCATCTTCATCA.....	75
Q1	TGAGCTCTTTGGAATCGAAGAGTTTCATGAGAATGCTGTCAAGGCACCGTGTCTACAGCCCGTGGTACGGTCCAG	160
Q2	TGAGCTCTTTGGAATCGAAGAGTTTCATGAGAATGCTGTCAAGGCACCGTGTCTACAGCCCGTGGTACGGTCCAG	154
Q3	TGAGCTCTTTGGAATCGAAGAGTTTCATGAGAATGCTGTCAAGGCACCGTGTCTACAGCCCGTGGTACGGTCCAG	154
Q4	TGAGCTCTTTGGAATCGAAGAGTTTCATGAGAATGCTGTCAAGGCACCGTGTCTACAGCCCGTGGTACGGTCCAG	154
Q1	ACCGTGTCAAGTACTTTGGGCCATTCTCTGTTAAGCCCCATCTTACTCTAGGTGATTCCTGGTGACTATGGCTGG	240
Q2	ACCGTGTCAAGTACTTTGGGCCATTCTCTGTTAAGCCCCATCTTACTCTAGGTGATTCCTGGTGACTATGGCTGG	234
Q3	ACCGTGTCAAGTACTTTGGGCCATTCTCTGTTAAGCCCCATCTTACTCTAGGTGATTCCTGGTGACTATGGCTGG	234
Q4	ACCGTGTCAAGTACTTTGGGCCATTCTCTGTTAAGCCCCATCTTACTCTAGGTGATTCCTGGTGACTATGGCTGG	234
Q1	GACACTGCTGGGCTTTTGGCTGATCCCGAGACTTTTCCAGAACCCTGAGCTTGAAGTCATCCACTCCAGATGGGCCAT	320
Q2	GACACTGCTGGGCTTTTGGCTGATCCCGAGACTTTTCCAGAACCCTGAGCTTGAAGTCATCCACTCCAGATGGGCCAT	314
Q3	GACACTGCTGGGCTTTTGGCTGATCCCGAGACTTTTCCAGAACCCTGAGCTTGAAGTCATCCACTCCAGATGGGCCAT	314
Q4	GACACTGCTGGGCTTTTGGCTGATCCCGAGACTTTTCCAGAACCCTGAGCTTGAAGTCATCCACTCCAGATGGGCCAT	314
Q1	GCTTGGAGCTCTTGGGTGTGCTTCCAGAGCTCTTGGCTCTAATGGTGTCAAGTTTGGCGAGGCGTGTGGTTCAAGG	400
Q2	GCTTGGAGCTCTTGGGTGTGCTTCCAGAGCTCTTGGCTCTAATGGTGTCAAGTTTGGCGAGGCGTGTGGTTCAAGG	394
Q3	GCTTGGAGCTCTTGGGTGTGCTTCCAGAGCTCTTGGCTCTAATGGTGTCAAGTTTGGCGAGGCGTGTGGTTCAAGG	394
Q4	GCTTGGAGCTCTTGGGTGTGCTTCCAGAGCTCTTGGCTCTAATGGTGTCAAGTTTGGCGAGGCGTGTGGTTCAAGG	394
Q1	CTGGATCACAATCTTTAGCGAGGGTGGACTTGACTATCTGGCAAAACCAAGCATGGTCTATGCTCAGAGCATCTTGTCC	480
Q2	CTGGATCACAATCTTTAGCGAGGGTGGACTTGACTATCTGGCAAAACCAAGCATGGTCTATGCTCAGAGCATCTTGTCC	474
Q3	CTGGATCACAATCTTTAGCGAGGGTGGACTTGACTATCTGGCAAAACCAAGCATGGTCTATGCTCAGAGCATCTTGTCC	474
Q4	CTGGATCACAATCTTTAGCGAGGGTGGACTTGACTATCTGGCAAAACCAAGCATGGTCTATGCTCAGAGCATCTTGTCC	474
Q1	ATCTGGGCTAACCAGATTATCTTGATGGGCGCAGTTGAGGGTTACCGTGTGCTGGTGGACCTCTGGGGAATTTGTGA	560
Q2	ATCTGGGCTAACCAGATTATCTTGATGGGCGCAGTTGAGGGTTACCGTGTGCTGGTGGACCTCTGGGGAATTTGTGA	554
Q3	ATCTGGGCTAACCAGATTATCTTGATGGGCGCAGTTGAGGGTTACCGTGTGCTGGTGGACCTCTGGGGAATTTGTGA	554
Q4	ATCTGGGCTAACCAGATTATCTTGATGGGCGCAGTTGAGGGTTACCGTGTGCTGGTGGACCTCTGGGGAATTTGTGA	554
GTTACCGTGTGCTGGTGA		
Q1	CCCACITTTACCGTGGAGGAAGCTTTGACCCATTAGGCCTTGCTGAAGACCCGAGGCATTTCGAGAGCTTAAGGTGAAGG	640
Q2	CCCACITTTACCGTGGAGGAAGCTTTGACCCATTAGGCCTTGCTGAAGACCCGAGGCATTTCGAGAGCTTAAGGTGAAGG	634
Q3	CCCACITTTACCGTGGAGGAAGCTTTGACCCATTAGGCCTTGCTGAAGACCCGAGGCATTTCGAGAGCTTAAGGTGAAGG	634
Q4	CCCACITTTACCGTGGAGGAAGCTTTGACCCATTAGGCCTTGCTGAAGACCCGAGGCATTTCGAGAGCTTAAGGTGAAGG	634
Q1	ACTGAAATAATGGAAGACTAGCCATGTTCTCCATGTTTGGTTCTTTGTACAAGCATATTGTACCGGAAAGGGTCCCTTA	720
Q2	ACTGAAATAATGGAAGACTAGCCATGTTCTCCATGTTTGGTTCTTTGTACAAGCATATTGTACCGGAAAGGGTCCCTTA	714
Q3	ACTGAAATAATGGAAGACTAGCCATGTTCTCCATGTTTGGTTCTTTGTACAAGCATATTGTACCGGAAAGGGTCCCTTA	714
Q4	ACTGAAATAATGGAAGACTAGCCATGTTCTCCATGTTTGGTTCTTTGTACAAGCATATTGTACCGGAAAGGGTCCCTTA	714
GAGAACAATGGCTAGTCTTCCA		
Q1	GAGCACTTGGCTGACCACCTTGCAGACCCTGTAAACAACAATGCTGGGCCCTTGCCACTAATTTGTTCCCGGAAAGTGA	801
Q2	GAGCACTTGGCTGACCACCTTGCAGACCCTGTAAACAACAATGCTGGGCCCTTGCCACTAATTTGTTCCCGGAAAGTGA	795
Q3	GAGCACTTGGCTGACCACCTTGCAGACCCTGTAAACAACAATGCTGGGCCCTTGCCACTAATTTGTTCCCGGAAAGTGA	795
Q4	GAGCACTTGGCTGACCACCTTGCAGACCCTGTAAACAACAATGCTGGGCCCTTGCCACTAATTTGTTCCCGGAAAGTGA	795

Fig. 2 Suppl. The alignment of amino acid sequences of Lhcb1 proteins without transit peptides from *A. graveolens*, *O. javanica* and other plants. Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura, and Q4 - Baguazhou Shuiqin. The *dark* areas are identical amino acid sequences, the *pink* are 75 % identical amino acid sequences, and the *green* are 50 % identical amino acid sequences. Putative Chl ligands are marked by the *asterisks*. The residues involved in trimer formation are marked by the *black* and *white inverted triangles*. The *box* indicates transit peptides.

Lhcb1- Q1	...MAASTMALSSPALARKE.VKVASSSSSELFGNRRVSMRMTVK...APVSNSEWYGFDRVKYLGEF	*****
Lhcb1- Q2	...MAASTMALSSPALAGKA.VKVAPSSS...ELFGNCRVSMRKTVK...APVSDSEWYGFDRVKYLGEF	
Lhcb1- Q3	...MAASTMALSSPALAGKA.VKVGPSSS...ELFGNCRVSMRKSVK...APVSNSEWYGFDRVKYLGEF	
Lhcb1- Q4	...MAASTMALSSPALAGKA.VKVAPSSS...ELFGNCRVSMRKTVK...APVSNSEWYGFDRVKYLGEF	
Apium graveolens	...MAASTMALSSPALAGKA.VKVAPSSS...ELFGNCRVSMRKTVK...APVSDSEWYGFDRVKYLGEF	
Arabidopsis thaliana	...MASSTMALSSPAFAGKA.VKPAASD...VLGSGRVTMRKTIVAK.PKGPSGSEWYGSDRVKYLGEF	
Glycine max	...MAASTMALSSSLAGQA.IKLAPSTP...QLGVGRVSMRKTAS...KTVISGSEWYGFDRVKYLGEF	
Medicago truncatula	...MAASTMALSSSMVGA.IKLSPTP...DLVMGRITMRKTTIK.KTAPSGSQWYGFDRVKYLGEF	
Populus trichocarpa	...MAAATMALSSPAFAGKA.VKLSPPSP...ELMGNCRVSMRKTIK...PVSSGSEWYGFDRVKYLGEF	
Vitis vinifera	...MAASTMALSSPSFAGTT.VKLGPNAS...DILGGERVSMRKTGF...KAPSGSEWYGFDRVLYLGFL	
Selaginella moellendorffii	MAAMTATATMAAPSTIFAGTGLKPFQNEFVARNASNGSRIAMRKITGKAS.SVGGGSEWYGFDRVKYLGEF	
Physcomitrella patens	MAAATAMHSTALAGQSLLKPVNELSRKGVNVEARVSMRKSAS...KSAASDSWYGADEFKYLGEF	
Oryza sativa japonica	...MAAATMALSSPALAG....KAAA....KVFGGGRITMRKSAAKPKPAASGSEWYGADEFVLYLGFL	
Zea mays	...MAAATMAITSRALVG....KPASTR..DVFGGGRITMRKTAGKPKPAASGSEWYGADEFVLYLGFL	
	* * ▽ ▽ * * *	
Lhcb1- Q1	SGKAPSYLIGFEFGDYGWDTAGLLADPETFSAKNRELEVIHSRWAMLGALGCVFEELLAHNCVKFGEAVW	
Lhcb1- Q2	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Lhcb1- Q3	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Lhcb1- Q4	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Apium graveolens	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Arabidopsis thaliana	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Glycine max	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Medicago truncatula	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Populus trichocarpa	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Vitis vinifera	SGDAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Selaginella moellendorffii	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Physcomitrella patens	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Oryza sativa japonica	SGEAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
Zea mays	SGQAPSYLIGFEFGDYGWDTAGLSADPETFSAKNRELEVIHSRWAMLGALGCVFEELLARNCVKFGCAVW	
	* * * * *	
Lhcb1- Q1	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Lhcb1- Q2	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Lhcb1- Q3	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Lhcb1- Q4	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Apium graveolens	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Arabidopsis thaliana	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Glycine max	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Medicago truncatula	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Populus trichocarpa	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Vitis vinifera	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Selaginella moellendorffii	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Physcomitrella patens	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Oryza sativa japonica	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
Zea mays	FKAGSQIFSEGGGLDYLGNPISLVEAQSILAIWATQVILMGAVEGYRVAG.GPLGEIVDFLYPGGSFDPLG	
	* * * *	
Lhcb1- Q1	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Lhcb1- Q2	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Lhcb1- Q3	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Lhcb1- Q4	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Apium graveolens	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Arabidopsis thaliana	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Glycine max	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Medicago truncatula	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Populus trichocarpa	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Vitis vinifera	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Selaginella moellendorffii	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Physcomitrella patens	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Oryza sativa japonica	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	
Zea mays	LADEEAEFAELKVKEIKNGRLAMFSMFGFFVQAIIVTGKGFLENLADH IADPVNNNAWAFATN FVPGK	

Fig. 3 Suppl. The prediction of transmembrane regions and orientation of Lhcb1. Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura, and Q4 - Baguazhou Shuiqin (i - inside (C-terminus), o - outside (N-terminus)).

