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
Apium graveolens	264	28076.0	5.31	9	10	11	20	83.1
Arabidopsis thaliana	265	28053.9	5.28	9	10	11	19	82.0
Glycine max	264	28024.9	5.27	9	9	11	20	83.1
Medicago truncatula	265	28284.4	5.96	10	9	11	20	84.4
Populus trichocarpa	264	28052.9	5.29	9	10	11	19	84.0
Vitis vinifera	264	27940.8	4.99	8	10	11	20	80.5
Selaginella moellendorffii	272	28800.6	5.30	9	10	11	17	83.8
Physcomitrella patens	267	28337.0	5.13	9	10	10	19	74.5
Oryza sativa japonica	261	27552.5	5.29	9	10	11	20	78.1
Zea mays	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
Apium graveolens	CAA99993					
Arabidopsis thaliana	18403546	15224471	15239602	15241005	15235029	15218330
Glycine max	356559635	358248712	359807444	356496185	356559931	356525886
Medicago truncatula	357496387	357495319	357452427	357461189		357445799
Populus trichocarpa	224114357	224131818	224122902	224090986	224127848	224054274
Vitis vinifera	225463432	225447576	359495476	225440674	225459918	225447745
Selaginella moellendorffii	302754140	302810440	302765457	302784975	302811558	302815978
Physcomitrella patens	168003213	168009155	168053612	168040609	168068372	168009894
Oryza sativa	Os01t0600900-02	Os03t0592500-01	115472785	Os07t0558400-01	Os11t0242800-01	Os04t0457000-01
Zea mays	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.

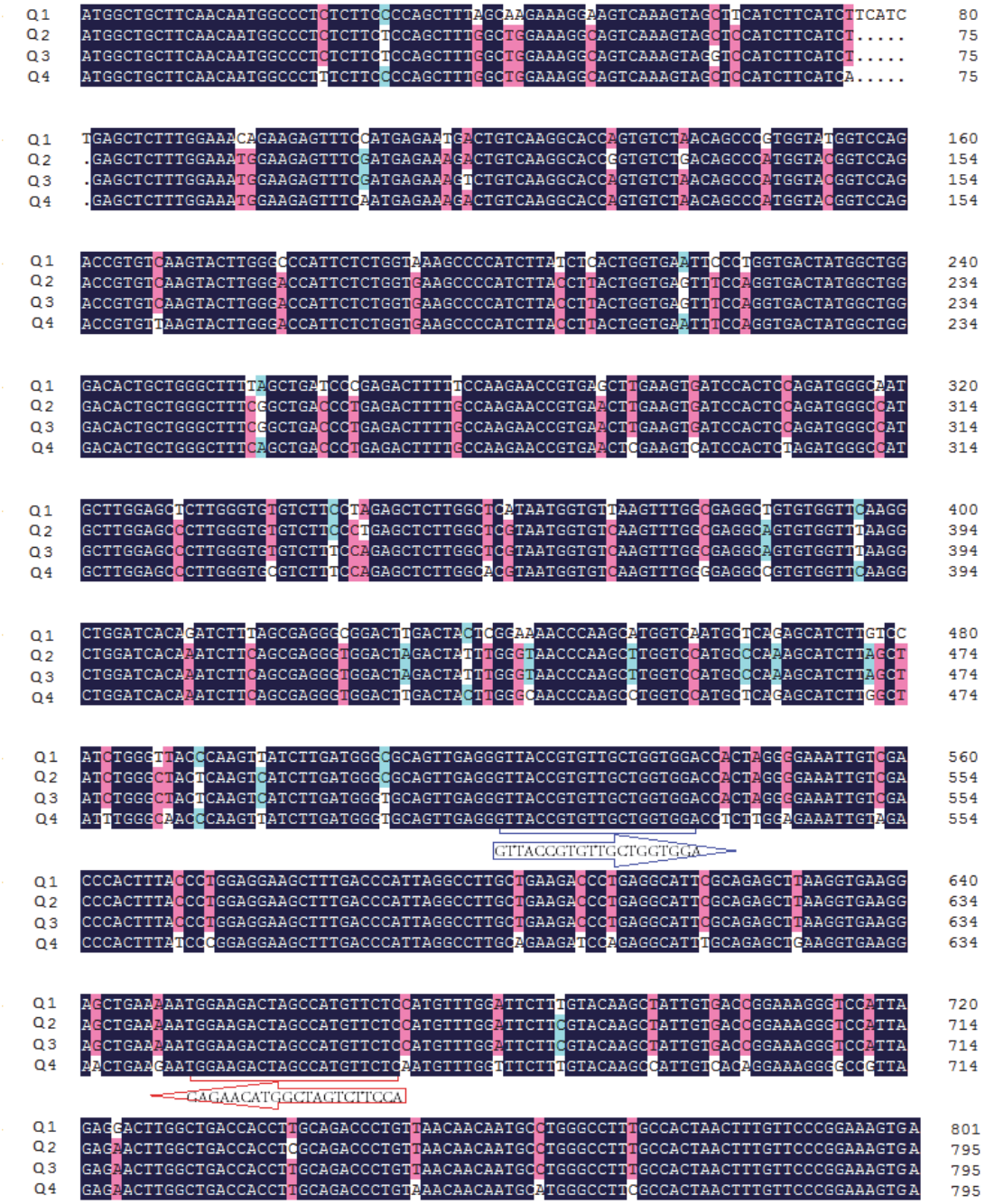

Fig. 2 Suppl. The alignment of amino acid sequences of Lhcb1 proteins without transit peptides from *A. graveolens*, *O. javanica* and other plants. Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura, and Q4 - Baguazhou Shuiqin. The *dark* areas are identical amino acid sequences, the *pink* are 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.

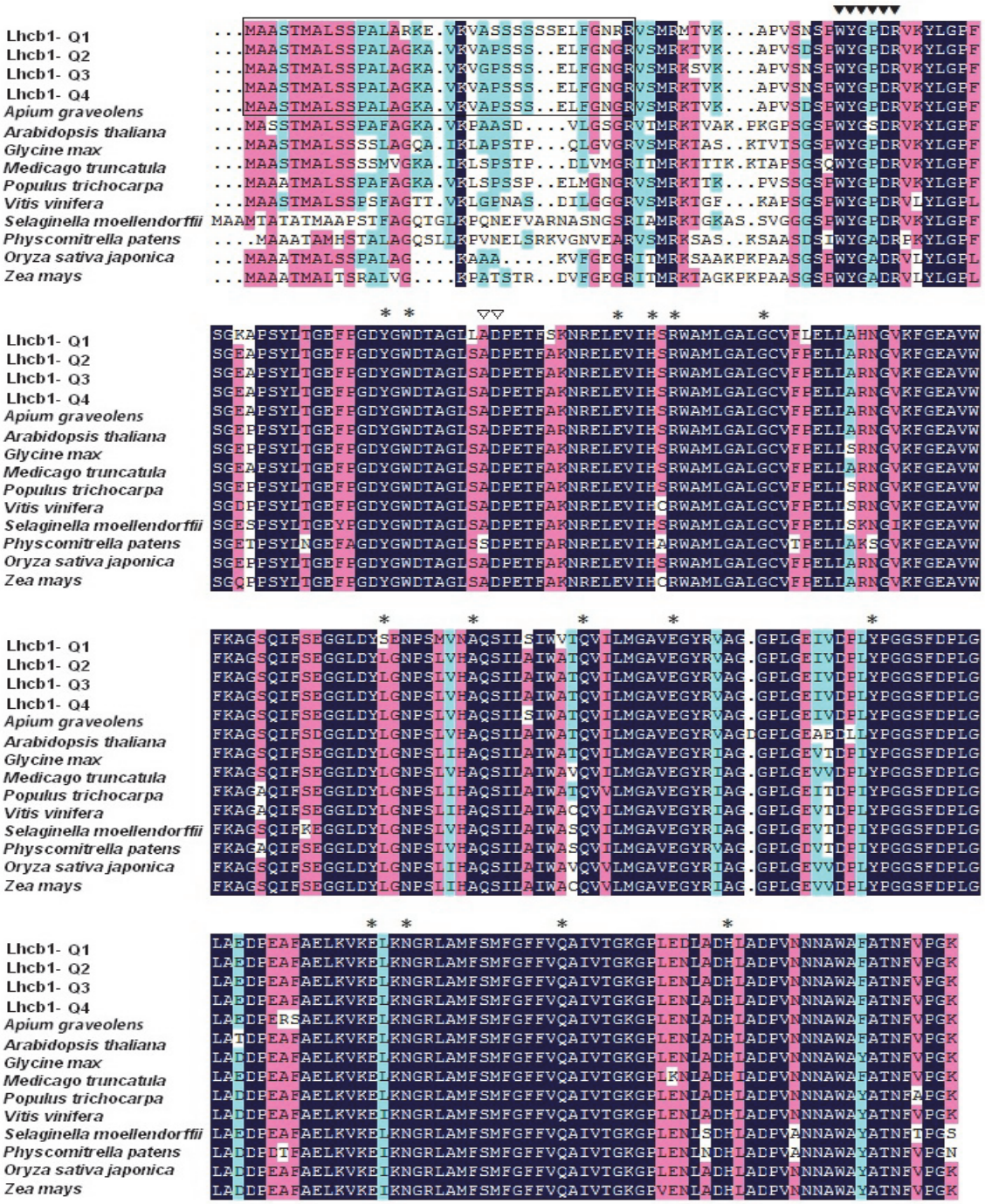

Fig. 3 Suppl. The prediction of transmembrane regions and orientation of Lhcb1. Q1 - Liuhe Huangxinqin, Q2 - Jinnan Shiqin, Q3 - Ventura, and Q4 - Baguazhou Shuiqin (i - inside (C-terminus), o - outside (N-terminus).