

Supporting Online Material for

BmILF and i-motif structure are involved in transcriptional regulation of BmPOUM2 in *Bombyx mori*

Kangkang Niu¹, Xiaojuan Zhang¹, Huimin Deng¹, Feng Wu¹, Yandong Ren², Hui Xiang¹, Sichun Zheng¹, Lin Liu¹, Lihua Huang¹, Baojuan Zeng¹, Sheng Li¹, Qingyou Xia³, Qisheng Song⁴, Subba Reddy Palli⁵, Qili Feng^{1,*}

¹Guangzhou Key Laboratory of Insect Development Regulation and Application Research, School of Life Sciences, South China Normal University, Guangzhou 510631, China, ²State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China, ³State Key Laboratory of Silkworm Genome Biology, Southwest University, Chongqing 400716, China, ⁴Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA, and ⁵Department of Entomology, University of Kentucky, Lexington, KY 40546-0091, USA

*Correspondence to: Qili Feng, School of Life Sciences, South China Normal University, Guangzhou, China, 510631. Email: qlfeng@scnu.edu.cn; Tel & Fax.: +86-20-85215291.

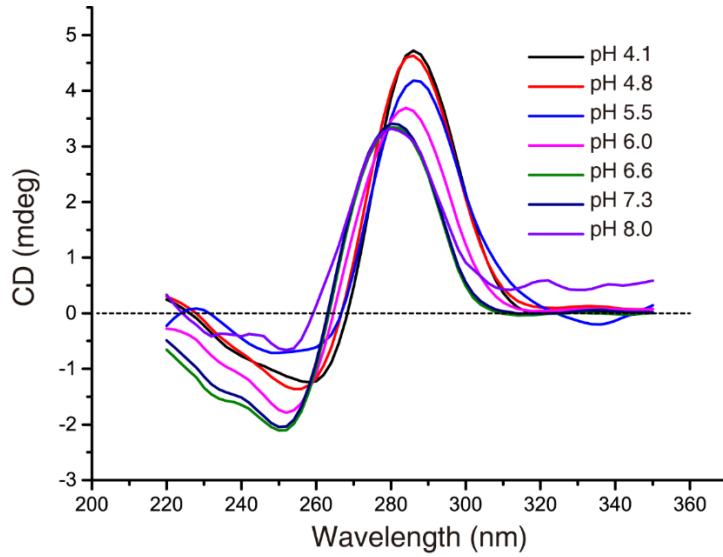
Supplementary Table 1. Candidate i-motif binding proteins. The proteins in red font were selected to test in EMSA.

Protein symbol	Protein introduction	UniProt database accession number
Uncharacterized protein	Uncharacterized protein	H9IWQ4
Uncharacterized protein	replication factor C subunit 3	H9JPK6
Polymerase delta interacting protein 3	polymerase delta interacting protein 3	Q0ZAL5
Uncharacterized protein	39S ribosomal protein L3, mitochondrial	H9J0W8
ILF	interleukin enhancer binding factor	Q1HQ53
Uncharacterized protein	Uncharacterized protein	H9JR48

Supplementary Table 2. The distribution of G4 and i-motif in the haploid silkworm genome.

	G4/i-motif	G4	i-motif	No. of corresponding Genes
Genome	6278	3174	3104	-
Genic region	2885	1193	1777	2086
Gene body	2220	868	1352	1497
CDS	1219	489	730	922
Intron	1040	403	637	809
Upstream 2kb region	357	147	210	332
Downstream 2kb region	519	230	289	462
Transposon element	3984	1895	2089	-
ncRNA (miRNA, rRNA, tRNA)	0	-	-	-

Supplementary Figure 1. CD analysis of the reverse ssDNA of the -88~-127 nts region of the *BmPOUM2* promoter at different pH values.



Supplementary Figure 2. Bioinformatics examination of the G4 and i-motif structures in the promoters of POUM2 orthologous genes.

<i>Bombyx mori</i>	-423	---	GCGGGT	-	CGGCGAGT	GC	CGG	-	CGAGT	GCGGCGAGT	GCGGCGGGC	ACGG	-	ACGCGAGACGT	-	ACAGAC
<i>Amyelois transitella</i>	-439	GACG	TCGAGT	TC	CGGTAC	GGGCGCT	TAC	CGCG	CAGAA	TTT	TCCGAA	CAGGG	CAGT	GAGGGG	TC	CACCGGC
<i>Danaus plexippus</i>	-274		ATG	TAGAGT	AGC	CGGCT	CGG	CGGA	TGA	TAGT	TAGC	GAGAG	CAG	CGGGG	TC	GGCGAG
<i>Papilio xuthus</i>	-423	-----	GCGT	AAGT	AATAC	TAAAAAA	TGTT	TGTA	AAATA	ATC	AGGT	GTTA	-	ATA	A	TAGGT
<i>Tribolium castaneum</i>	-409	-----	GGGG	TGAAT	CGA	AGGAG	AGT	CGT	GC	AGGA	---	GCG	AGCC	ACT	GGC	GAAGG
<i>Drosophila melanogaster</i>	-410	-----	CGCG	CGT	TCC	ACGAT	GCCT	CTT	ACT	CGGG	ATTT	GGT	GCG	TAGC	---	CGCG
<i>Apis mellifera</i>	-398	---	AAA	TAAAAAA	ATC	TAAAT	AAATA	ACC	AAAT	AAAA	CCG	AGT	TAG	CT	CA	ACGC
<i>Bombyx mori</i>	-361	GCG	-	AGAG	TTTT	CCAGT	G	GGGG	CGCG	-	AGGGG	CCG	AGG	-	GCGGG	CAAC
<i>Amyelois transitell</i>	-369	GCG	GGGGG	GC	ACT	CAC	GAG	CCCC	CGG	-	GGG	CAG	AGGG	GC	CGGG	CGCG
<i>Danaus plexippus</i>	-214	ACG	-	GGGGG	T	GAG	AGC	CGGG	TTG	-	GGGG	T	CGGG	CGGG	CGCG	CGC
<i>Papilio xuthus</i>	-369	GTA	-	GGA	GGG	CGCG	T	GGGG	CT	CGT	GT	GAG	CGCG	GGGG	CGCG	CGCG
<i>Tribolium castaneum</i>	-350	ACG	-	T	CAC	GC	ACT	T	GCG	AGT	GAG	-	AGG	T	CAC	CA
<i>Drosophila melanogaster</i>	-350	GCG	-	TT	GAT	CCG	AAA	AGT	TC	GGT	TG	-	AGA	T	CA	T
<i>Apis mellifera</i>	-331	ACG	T	GAA	ACCC	CAAG	CA	AA	T	GCC	G	-	GGA	TTTT	T	A
<i>Bombyx mori</i>	-297	AG	CGGG	-	CG	CC	CG	CT	G	GGG	CGG	CC	GG	CG	CG	CG
<i>Amyelois transitell</i>	-301	AG	CGGG	-	CGG	---	-	T	GGG	CGGG	CT	T	CCCC	---	C	T
<i>Danaus plexippus</i>	-152	AG	CGGG	-	CGG	C	-	G	T	GGG	CGGG	CC	CCCC	CT	---	C
<i>Papilio xuthus</i>	-300	AG	CGGG	-	CGG	CCG	-	G	T	GGG	CGGG	CC	CCCC	CT	---	C
<i>Tribolium castaneum</i>	-291	CG	GGCGG	-	GGA	T	CG	AA	T	GG	AGGG	GG	CT	G	AG	C
<i>Drosophila melanogaster</i>	-283	CG	GA	AG	CC	T	AC	G	A	C	G	A	C	T	C	G
<i>Apis mellifera</i>	-263	CG	CG	TT	CG	AG	GG	CC	AC	GG	CCCC	AAAA	AA	CA	AA	CG