

Additional file 4. PUT-165a-Gossypium_hirsutum-41616 for PIP1;1/PIP1;14. These data were provided as an example of PUT assembly contig analysis (See Table 3).

gi|164355644+
gi|84166930+
gi|164338038+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+

ATG
ATG
ATG
ATG
ATG
ATG
ATG

consensus

TTTATTGTCCT-TTTGGAATTTTGTACTAAAAAAGGAAGCAACGAAGAAGGGAAATG
M

gi|164355644+
gi|84166930+
gi|164338038+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+

. : . : . : . : . :
GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
TAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GAGGGCAGAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT

consensus

GAGGGCAAAGAAGAGGATGTTAGATTGGGAGCCAACAAGTTCACAGAGAGGCAACCAATT
E G K E E D V R L G A N K F T E R Q P I

gi|164355644+
gi|84166930+
gi|164338038+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+
gi|109825953+

. : . : . : . : . : . :
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCGGCTCCA
GGGACCGCGGCTCAGACTCAAGATGACGGGAAAGACTATTCTGAACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCACCGGCTCCA
GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCACCGGCTCCA
CACCACCGGCTCCA

consensus

GGGACCGCGGCTCAGAGTCAAGATGACGGGAAAGACTTACACTGAACCACCACCACCGGCTCCA
G T A A Q S Q D D G K D Y T E P P P A P

gi|164355644+
gi|84166930+
gi|164338038+
gi|84154773+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+
gi|109825953+

. : . : . : . : . : . :
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
GCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT

consensus

TTTTTCGAGCCTGGCGAGTTAACCTCATGGTCTTTTTATAGGGCTGGGATTGCCGAGTTT
F F E P G E L T S W S F Y R A G I A E F

gi|164355644+
gi|84166930+
gi|164338038+
gi|84154773+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+
gi|109825953+

GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA

consensus

GTGGCGACTTTCCTTTTCTTGTATATCACTATTTTGACTGTTATGGGAGTTGTTAAGGAA
V A T F L F L Y I T I L T V M G V V K E

gi|164355644+
gi|84166930+
gi|164338038+
gi|84154773+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+
gi|84144380+
gi|164293662+
gi|109825953+

AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
TGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
GCCCAACTGTGTGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT

consensus

AAAAC TAAGTGCCCAACTGT-TGGGATTCAAGGAATTGCTTGGGCTTTTGGTGGTATGAT
K T K C P T V G I Q G I A W A F G G M I

gi|164355644+
gi|84166930+
gi|164338038+
gi|84154773+
gi|84153628+
gi|31407051+
gi|164306946+
gi|78339752+
gi|118132685+
gi|164345171+
gi|84144380+
gi|164255409+
gi|164293662+
gi|109825953+

C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C

consensus

C T T T G C T C T T G T T T A C T G C A C T G C T G G C A T T T C A G G T G G C C A T A T C A A C C G G C A G T G A C
F A L V Y C T A G I S G G H I N P A V T

gi|164355644+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|84166930+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGA
gi|164338038+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|84154773+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|84153628+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|31407051+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|164306946+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|78339752+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|118132685+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|164345171+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|84144380+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|164255409+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|164293662+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
gi|109825953+ TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT

consensus TTTTGGACTGTTCTTGGCGAGGAAGCTGTCGTTGACGAGGGCAATTTTCTACATGGTGAT
F G L F L A R K L S L T R A I F Y M V M

gi|164355644+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGGTTTCATGGGGAAAGACTC
gi|164338038+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|84154773+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|84153628+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|31407051+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|164306946+ GCAGTGT ATGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGG-TTCATGGGGAA-GACTC
gi|78339752+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|118132685+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|82755095+ GCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|84144149+ GGGAA-GACTC
gi|164345171+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|84144380+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|164255409+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|164293662+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
gi|109825953+ GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC

consensus GCAGTGT TTTGGGCGCCATATGCGGGGCGGGTGTGGTGAAAGGTTTCATGGGGAA-GACTC
Q C L G A I C G A G V V K G F M G K T

gi|164355644+ AGTACGGTATGTTGGGCGCGGAGCTAA-CTCGGNGGCTCATGGCTACACAAAA-AGNG
gi|164338038+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|84154773+ AGTACGGTATGTTGGG-CGGCGGAGCTAAAGCTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|84153628+ AGTACGGTATGTNNGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|31407051+ AGTACGGTATGTTGGG-CGGCGGAAGCTAA-CTCGGTGGCTCATGGCTACACAAAAAGGTG
gi|164306946+ A-TACGGTATGTAGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-AGTG
gi|78339752+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|118132685+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|82755095+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|84144149+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|164345171+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|84144380+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|164255409+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAA-GGTG
gi|164293662+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|109825953+ AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|84174539- GGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
gi|109842739+ TGTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGATACGCGTA--GGTG

consensus AGTACGGTATGTTGGG-CGGCGGAGCTAA-CTCGGTGGCTCATGGCTACACAAAA-GGTG
Q Y G M L G G G A N S V A H G Y T K G

[illegible]

gi 164338038+	TTTGGGATTTCGGGGTGGGAACAATCCTTGGTGCAGCACTAACAGCACTCTTAC-AC
gi 78339752+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 118132685+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 82755095+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 84144149+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 164345171+	ATTGG-ATNTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 84144380+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACCAC
gi 164255409+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 164293662+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 109825953+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 84174539-	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC
gi 109842739+	ATTGG-ATTTTCTGGTGGGACCATTTCATTGGT-GCAGCACTAGC-AGCACTCTACC-AC

	.	:	.	:	.	:	.	:	.	:	.	:	.	:
gi 164338038+	GTTGGTGTGATCAGAGCCCTTCCTT-TCAAAACCAAG	TGA												
gi 78339752+	GTTG-TTGTGATCAGAGCCATTCCCTT-TCAAATC-AAG	TGA												
gi 118132685+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATCAAAG	TGA												
gi 82755095+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATCAAAG	TGA												
gi 84144149+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATCAAAG	TGA												
gi 164345171+	GTTG-TTGTGATCAGAGCCATTTCCTGNTTCAATC-AAG	TGA												
gi 84144380+	GTTG-TTGTGATCAGAGCCATTTCCTT													
gi 164255409+	GTTG-TTGTGATCAGAGCCATTCCTT-TCAACTCAACG	TGA												
gi 164293662+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATCAAAG	TGA												
gi 109825953+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATTNAAG	TGA												
gi 84174539-	GTTG-TTGTGATCAGAGCCATTTCCTT-TCAAATCAAAG	TGA												
gi 109842739+	GTTG-TTGTGATCAGAGCCATTTCCTT-TCGAATCAAAG	TGA												