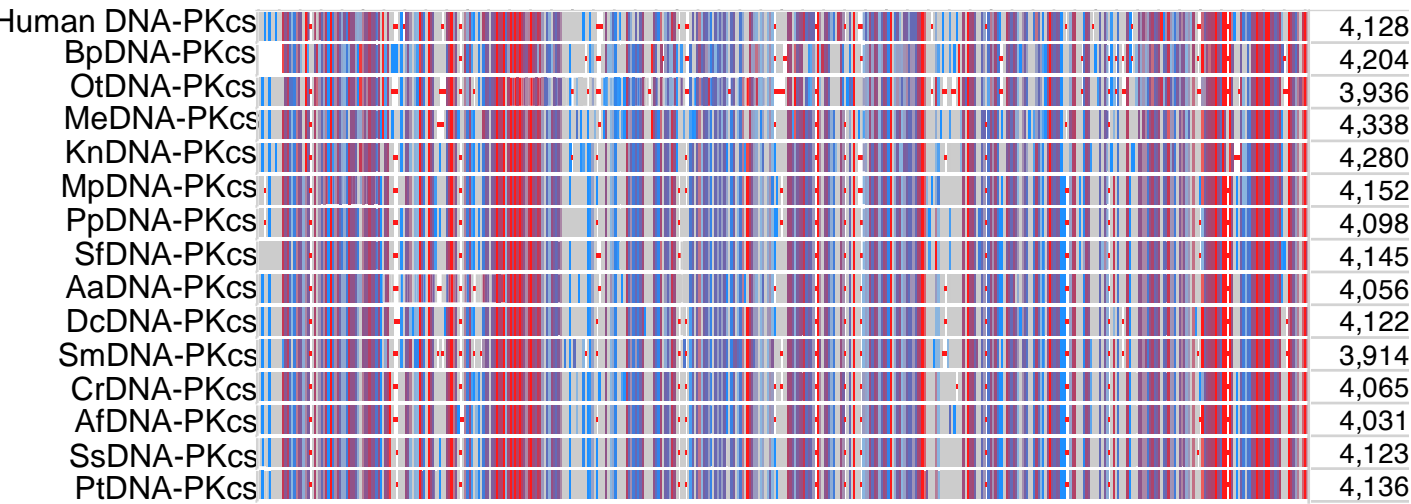


Supplementary figure 1.

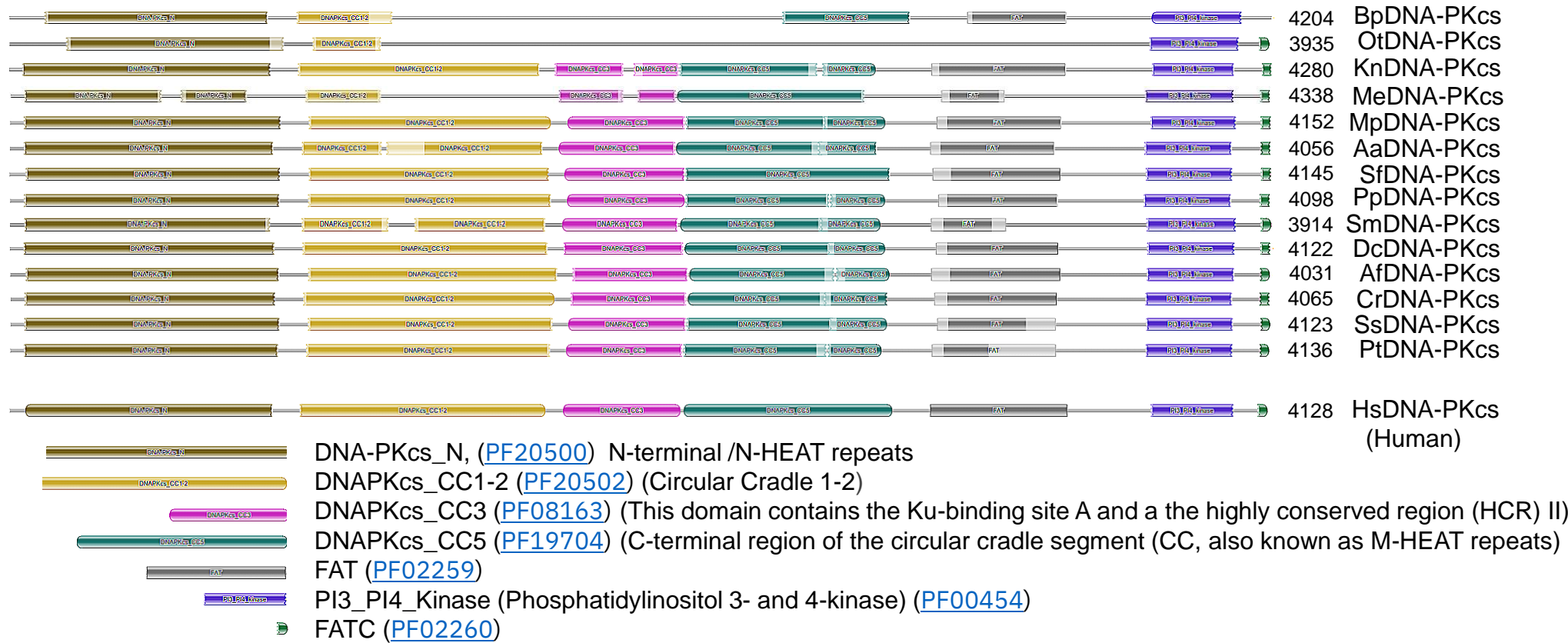


Supplementary figure 1.

Overall view of amino acid conservation in DNA-PKcs proteins of green plants. Putative DNA-PKcs proteins sequences from green plants and DNA-PKcs from human were aligned using COBALT, and regions of high amino acid similarity/conservation are represented by red, regions of moderate conservation in blue, and non-conserved amino acids in grey. The putative DNA-PKcs sequences are from Bp (*Bathycoccus prasinos*), Ot (*Ostreococcus tauri*), Me (*Mesotaenium endlicherianum*), Kn (*Klebsormidium nitens*), Mp (*Marchantia polymorpha*), Pp (*Physcomitrella patens*), Sf (*Sphagnum fallax*), Aa (*Anthoceros agrestis*), Dc (*Diphasiastrum complanatum*), Sm (*Selaginella moellendorffii*), Cr (*Ceratopteris richardii*), Af (*Azolla filiculoides*), Ss (*Sequoia sempervirens*), Pt (*Pinus taeda*), and Hs (*Homo sapiens*).

[illegible]

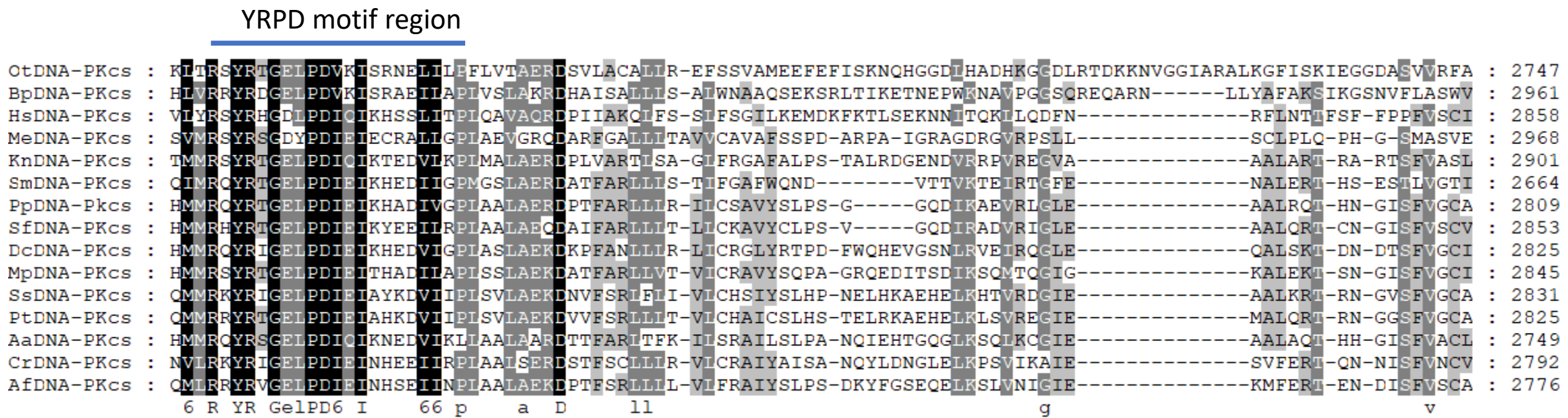
Supplementary figure 3



Supplementary figure 3

Detection of Pfam domains in human and plant DNA-PKcs. Protein sequence was searched against Pfam database using the program hmmscan. Conserved Pfam domains in human DNA-PKcs were found to be present in the majority of plant proteins, indicating possible conservation of structure and function. Sequence variability below the detection threshold could be the reason for absence of few domains in DNA-PKcs from chlorophytes. The putative DNA-PKcs sequences are from Bp (*Bathycoccus prasinus*), Ot (*Ostreococcus tauri*), Me (*Mesotaenium endlicherianum*), Kn (*Klebsormidium nitens*), Mp (*Marchantia polymorpha*), Pp (*Physcomitrella patens*), Sf (*Sphagnum fallax*), Aa (*Anthoceros agrestis*), Dc (*Diphasiastrum complanatum*), Sm (*Selaginella moellendorffii*), Cr (*Ceratopteris richardii*), Af (*Azolla filiculoides*), Ss (*Sequoia sempervirens*), Pt (*Pinus taeda*), and DNA-PKcs from Hs (*Homo sapiens*).

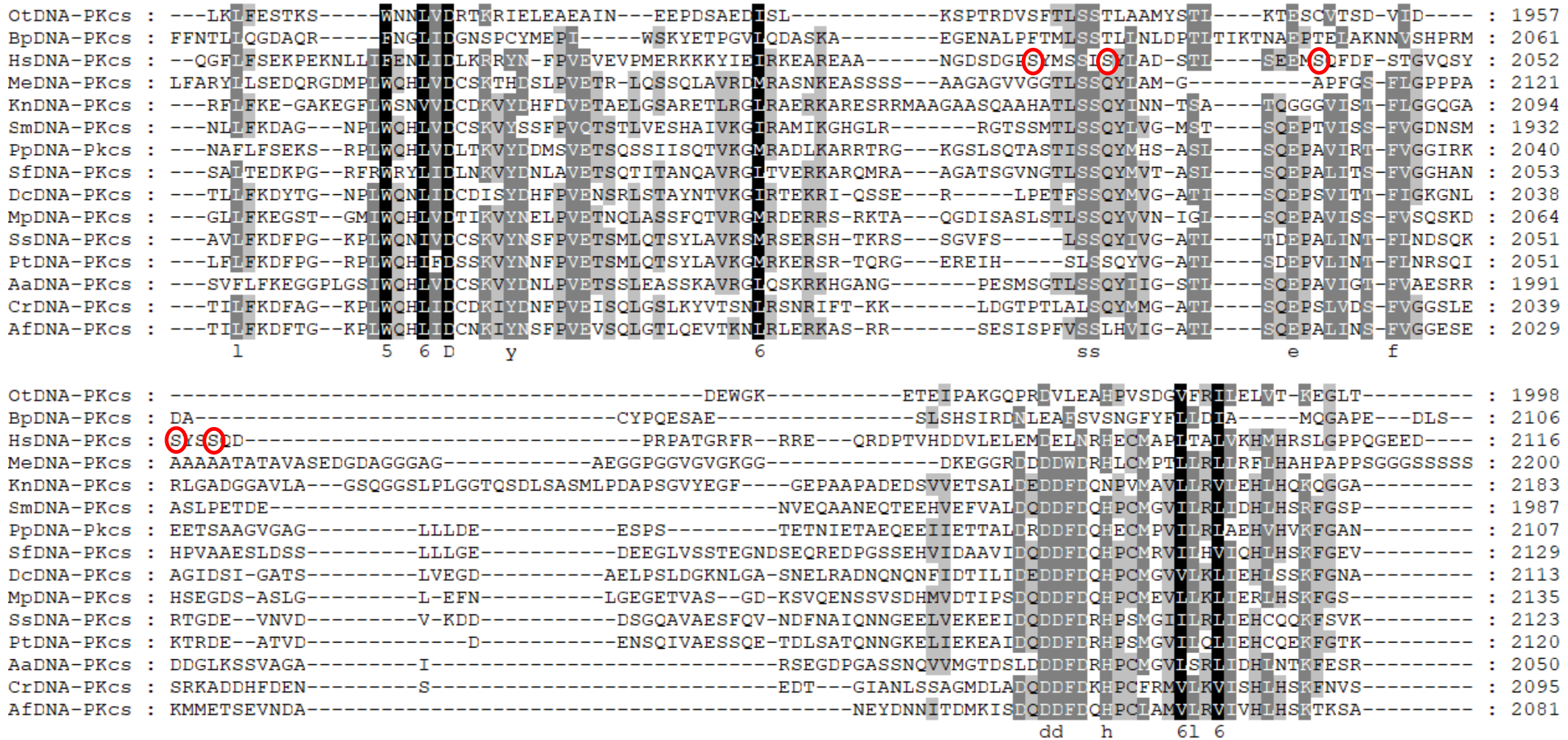
Supplementary Figure 4.



Supplementary Figure 4. Conservation of YRPD motif in green plants. Compared to metazoans, the YRPD motif is slightly modified in green plants (YR-x-G-(D/E)-(L/Y)-PD-(I/V)-x-I-x5-(I/V/L)-x-(P/L)-x-x), however the residues YRxxGxxPD are invariant like in other organisms.

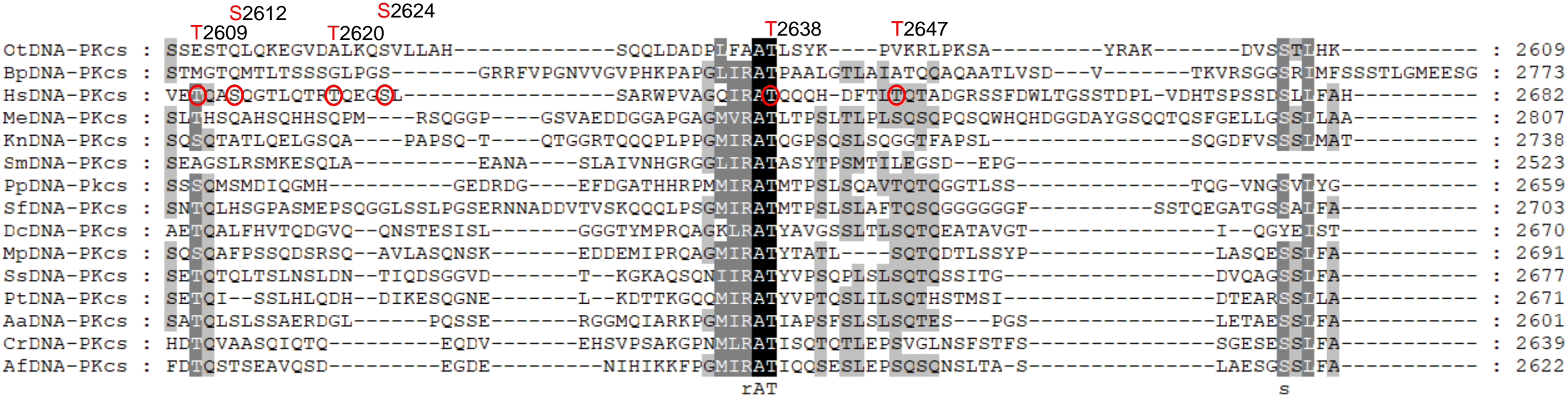
Bp- *Bathycoccus prasinos*; Ot- *Ostreococcus tauri*; Me- *Mesotaenium endlicherianum*; Kn- *Klebsormidium nitens*; Mp- *Marchantia polymorpha*; Pp- *Physcomitrella patens*; Sf- *Sphagnum fallax*; Aa- *Anthoceros agrestis*; Dc- *Diphasiastrum complanatum*; Sm- *Selaginella moellendorffii*; Cr- *Ceratopteris richardii*; Af- *Azolla filiculoides*; Ss- *Sequoia sempervirens*; Pt- *Pinus taeda*; Hs- *Homo sapiens*.

Supplementary Figure 5.



Supplementary Figure 5. PQR cluster of DNA-PKcs. Five serine residues of the PQR cluster of Human DNA-PKcs are marked with red circles. Lack of conservation of the PQR cluster residues in green plants can be observed.

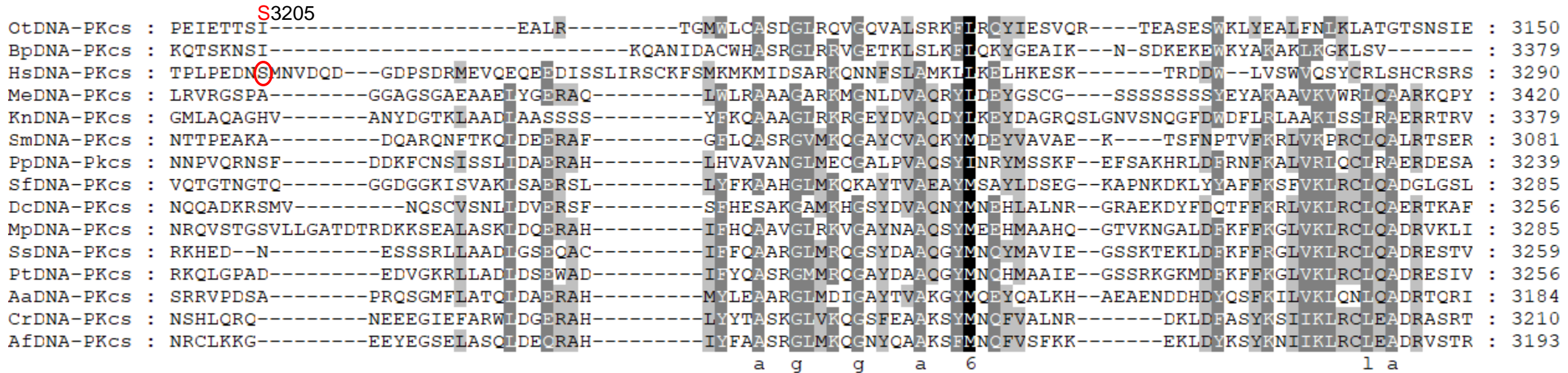
Supplementary Figure 6.



Supplementary Figure 6. ABCDE cluster of DNA-PKcs. The ABCDE cluster is a region of the human DNA-PKcs protein that contains six serine/threonine phosphorylation sites important for its activity. In this figure, these sites are highlighted in red circles and labeled. It can be observed that the T2638 site is invariant in all plants analyzed, indicating its importance for DNA-PKcs function across different species. Furthermore, T2609 site is conserved in most of the plant DNA-PKcs homologs examined, suggesting a possible functional role in these organisms.

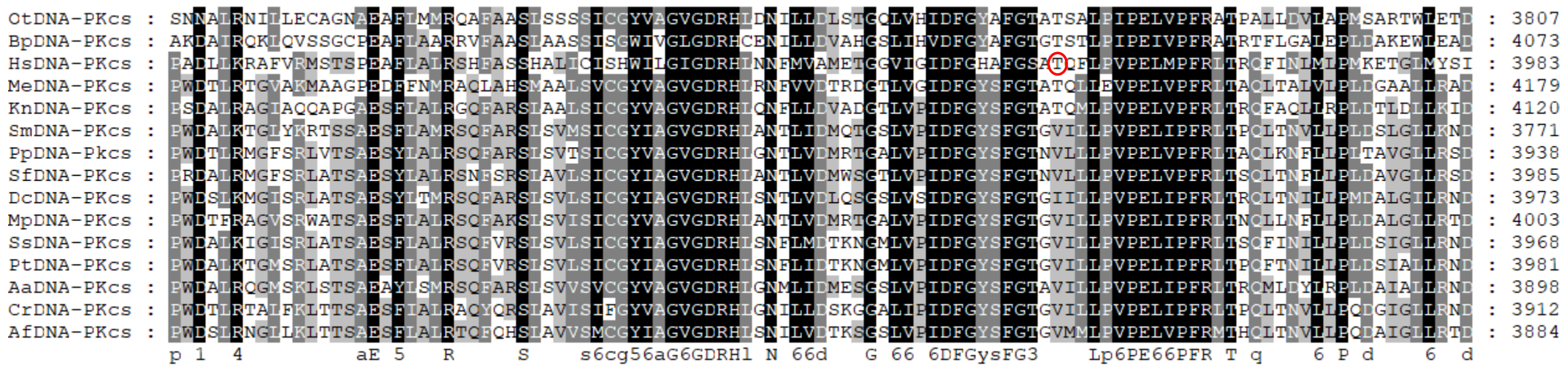
Bp- *Bathycoccus prasinos*; Ot- *Ostreococcus tauri*; Me- *Mesotaenium endlicherianum*; Kn- *Klebsormidium nitens*; Mp- *Marchantia polymorpha*; Pp- *Physcomitrella patens*; Sf- *Sphagnum fallax*; Aa- *Anthoceros agrestis*; Dc- *Diphasiastrum complanatum*; Sm- *Selaginella moellendorffii*; Cr- *Ceratopteris richardii*; Af- *Azolla filiculoides*; Ss- *Sequoia sempervirens*; Pt- *Pinus taeda*; Hs- *Homo sapiens*.

Supplementary Figure 7.



Supplementary Figure 7. Conservation of DNA-PKcs S3205 autophosphorylation site in plant lineages. Amino acid alignment of the FAT domain region of DNA-PKcs in selected plant species. The S3205 site in human DNA-PKcs is highlighted in red, and the lack of conservation in green plants can be observed.

Supplementary figure 8.



Supplementary Figure 8. Conservation of DNA-PKcs T3950 autophosphorylation site in plant lineages. Amino acid alignment of the kinase domain of DNA-PKcs from different plant lineages. The T3950 site is highlighted in red, and the substitutions by valine and isoleucine in some plant lineages can be observed.

Supplementary figure 9.

```
# -----
# LOCALIZER 1.0.4 Predictions
# -----
Identifier      Chloroplast      Mitochondria      Nucleus
BpDNA-PKcs      -                -                Y (KKRK,RASKRRI,RKYRALQKHSREQRK,KRRISSSGSDTKRRSQ,KRKDETATSLEARKYRA,KKNRVLFHTEVNHEKEKE)
OtDNA-PKcs      -                -                Y (KKRK,KRSIVIWGGGKKRKQL)
MeDNA-PKcs      -                -                Y (KRRS,RRKR,RRRR,RKYLREVLERSRRYKQ)
KnDNA-PKcs      -                -                Y (KKPR,KPRK,KRAR,RPRR,RRRR,SKRQKGR,RKFLREVLVASRRYKL,KRNTADMTRTQARRQRE,RRGEELREEYQFQKIYRK,RRPVREGVAAAL
MpDNA-PKcs      -                -                Y (KRQK,RKNEKDQSVYLDKIYKR)
PpDNA-PKcs      -                -                Y (KKRK,KRAR,KRQK,RKYLQEVVLASRRYKL,RRRGEDSQSVHFDDKIYKR,KKLQQLLVYIETERRKLSK)
SfDNA-PKcs      -                -                Y (KRAR,RKRK,RNKKRHR,RKPDMMNSDVQKRL)
AaDNA-PKcs      -                -                Y (KRKH,RRYLQEVVLASRRYKL,KRVQHSSVVRASAIRGKRN)
DcDNA-PKcs      -                -                Y (RRTASRNSEINIERLYRK,KRLAKRNIDVLQKERKDRL)
SmDNA-PKcs      -                -                Y (KRHR,RKRK,RKKTMDSKTLKLEKEKTK)
CrDNA-PKcs      -                -                Y (KRKL)
AfDNA-PKcs      -                -                Y (RRYLQEVLISSKRYKL,RKNMPWQYVGNVRKRHKG,RKHYYKILVYLYTERKMER)
SsDNA-PKcs      -                -                Y (LGEV,RKYLQEVVLASRRYKL,RRNITILLKNRKDRRM,RRKGQEHQSVHYEKIYRR)
PtDNA-PKcs      -                -                Y (KRKL,RRNISSLQENRKDRKM)

# Proteins analyzed: 14

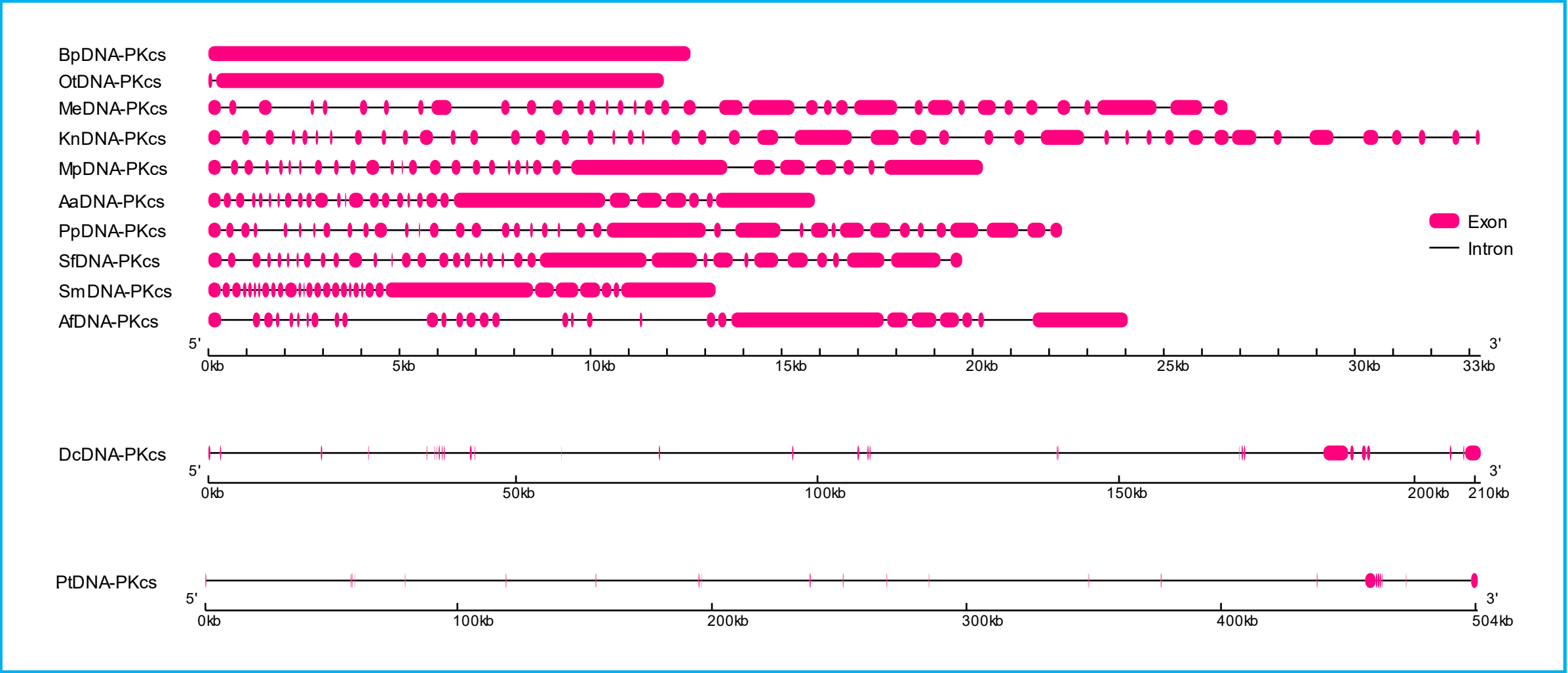
# Number of proteins with cTP: 0 (0.0%)
# Number of proteins with cTP & possible mTP: 0 (0.0%)
# Number of proteins with cTP & NLS: 0 (0.0%)
# Number of proteins with cTP & possible mTP & NLS: 0 (0.0%)
# Number of proteins with mTP: 0 (0.0%)
# Number of proteins with mTP & possible cTP: 0 (0.0%)
# Number of proteins with mTP & NLS: 0 (0.0%)
# Number of proteins with mTP & possible cTP & NLS: 0 (0.0%)
# Number of proteins with NLS and no transit peptides: 14 (100.0%)

# Summary statistics

# Number of proteins with chloroplast localization (cTP, cTP & possible mTP, cTP & NLS, cTP & possible mTP & NLS): 0 (0.0%)
# Number of proteins with mitochondrial localization (mTP, mTP & possible cTP, mTP & NLS, mTP & possible cTP & NLS): 0 (0.0%)
# Number of proteins with nuclear localization and no transit peptides: 14 (100.0%)
# Number of proteins with nuclear localization and with transit peptides: 0 (0.0%)
```

Supplementary figure 9. Subcellular localization prediction of putative plant DNA-PKcs proteins. LOCALIZER (<https://localizer.csiro.au/>) was used to predict the subcellular localization. All the 14 proteins used in the current study are predicted to be localized to nucleus.

Supplementary figure 10.



Supplementary Figure 10. Gene length and putative gene structure of DNA-PKcs from different plant lineages. Organisms with no complete gene sequence available, due to its representation in two different genome scaffolds or partially missing from genome scaffolds are excluded. Introns and exons are represented by dark lines and pink boxes, respectively.