PPP Project Annual Report 2018
The PPP-projects that have been established under the direction of the top sectors must submit an annual report on their technical and financial progress. This format is to be used for reporting the technical progress. A separate format ('PPP final report') is available for PPP-projects that have been completed in 2018.

The annual reports will be published in full on the websites of the TKIs/top sector, excluding the blocks 'Approval coordinator/consortium' and 'Planning and progress'. Please ensure that no confidential matters are left in the remaining blocks.

The PPP Project Annual Reports must be submitted to the TKI's before March 1st 2019. For Wageningen Research this will be coordinated via a central point.

General information	
PPP number	KV 1604-002
Title	*The Capsicum Genome Initiative*
Theme	Duurzame Plantaardige Productie
Executive knowledge institution(s)	Cluster Applied Bioinformatics, department of Bioscience, Wageningen University and Research
Research project leader (name + e-mail address)	Dr. ir. Sander Peters (sander.peters@wur.nl)
Coordinator (on behalf of private parties)	Dr. Aat Vogelaar (a.vogelaar@rijkzwaan.nl)
Government contact person	Annet Zweep
Total project size (k€)	854 keuro
Address project website	https://www.wur.nl/nl/Onderzoek-Resultaten/Onderzoeksprojecten-LNV/Expertisegebieden/kennisonline/KV-1604-002-Capsicum-Genome-Initiative.htm
Start date	31-03-2018
End date	31-03-2020

Approval coordinator/consortium *The annual report should be discussed with the coordinator/the consortium. The TKIs appreciate being informed of possible feedback on the annual report.*	
The coordinator has assessed the annual report on behalf of the consortium:	☐ approved ☐ rejected
Possible feedback on the annual report:	

Planning and progress (if there are changes to the project plan, please explain)	
Is the PPP going according to plan?	yes
Have there been changes in the consortium/project partners?	no
Is there a delay and/or deferred delivery date?	no
Are there any substantive bottlenecks? Provide a brief description	no

Are there any deviations from the projected budget?	no

Short content description/aim PPS
What is going on and how is this project involved?
What will be delivered by the project and what is the effect of this?

This TKI project started in March 2018 applying exiting new genomics technologies, including Genome Mapping, 10X Genomics, and state-of-the-art NGS. We aim to reconstruct and annotate new reference genomes of wild Capsicum species and mine the genetic diversity of gene bank collections such as from the Centre of Genetic Resources the Netherlands (CGN) and Katsetsart University Thailand. With the latest technical improvements, such as PacBio 'Sequel' technology and BioNano 'Saphyr', large insert sequencing for genomes such as *Capsicum* becomes economically feasible, while assessment of genetic diversity of a large panel of genomes can very well be supported by Illumina sequencing technology. The recently developed technology from 10X is an innovative technology that will be of further support for remapping to uncover inaccessible parts of the genome as well as megabased diploid assemblies preserving phasing details such as SNVs, indels, and structural rearrangements. In addition, genome wide physical profiling of reference genomes with BioNano technology has proven successful for extended scaffolding in our lab, solving many assembly problems for complexed sequences of crop genomes. The high quality *Capsicum* reference genome will serve as a golden standard to assess the genetic diversity of a large panel of pepper accessions from genebank collections such as from CGN (Centre for Genetic Resources the Netherlands) and Katsetsart University. The knowledge gained in the Capsicum Genome Initiative will allow breeders to have insight in genetic and genome structure diversity that is necessary to advance on precision breeding, reducing their R&D costs and better respond to economic, environmental, social and health issues inherent to agricultural crop production.

Results in 2018/ so far
Give a short description of the high-lights and project deliverable in 2018 / so far

The results have been presented and discussed in the CGI progress meeting in November 28, 2018 in Wageningen.

- Regeneration of 114 *Capsicum* accessions, including accessions from the Centre of Genetic Resources, The Netherlands (CGN) and Katsetsart University (Th) genebank collections
- Phenotypic characterization with ECPGR descriptors and photographic images
- Seed collection and distribution to the project partners under SMTA (22 accessions that did not set seed are regrown)
- Visual and laboratory sanity check for plant material and seed material
- DNA isolation and sequence library construction from 103 selected accessions including (3 accessions for de novo sequencing and 100 accessions for draft sequencing)
- QC check for DNA preps and library preps-> Go decision

- Chloroplast contamination ckeck is lower than 6% -> go decision
- IsoSeq library preparation from 4 tissues (leaves, roots, flowers, fruits) for 3 de novo accessions and QC check -> go decision
- Pacbio Sequel data production meeting the output requirement of 150Gb -> go decision
- Genome mapping for *C. chacoense* is 245Gb with an N50 molecule size of 360kb
- NGS, de novo genome map and hybrid assembly/scaffolding for *C. chacoense* with excellent statistics; ass. Size 3,08 Mb in 46 scaffolds with an N50 scaffold size 241Mb
- We have stablished a collaboration with Cornell University. Based on the assembly comparisons of HiC and BioNano scaffolds from *S. lycopersicum* cv. Heinz have that have been cross-checked with a third platform technology using 10X Genomics we have detected and analysed inconsistencies between HiC and BioNano scaffolds that are resolved in favour of BioNano genome scaffolds.
- Based on these findings we have proposed to use 10X + BioNano Genomics scaffolding for Capsicum *de novo* reference genome reconstruction -> favoured by all project partners.

Number of delivered products in 2018 *(in an appendix, please provide the titles and/or description of the products or a link to the products on public websites)*			
Academic articles	Reports	Articles in journals	Introductions/workshops
Titles/ description of the most important products in 2018 (5 at max) and their target group			

Appendix: Names of the products or a link to the products on a public website including the link to the project summary on Kennisonline

The link for the project at Kennis Online:

https://www.wur.nl/nl/Onderzoek-Resultaten/Onderzoeksprojecten-LNV/Expertisegebieden/kennisonline/KV-1604-002-Capsicum-Genome-Initiative.htm