

Acronyme	VizFaDa
Titre de la proposition	Visualisations of FAANG Data
Nom et prénom du responsable (coordinateur)	DEVAILLY Guillaume
Organisme du coordinateur	Identification de l'établissement : INRA GenPhySE (UMR 1388) Adresse : 24, chemin de Borde-Rouge - Auzeville Tolosane 31326 Castanet Tolosan
Choix d'une des catégories suivantes:	Outils – méthodes
Aide demandée	76 068 €
Coût complet :	183 644 €
Durée	24 mois
Mots clés de la proposition	transcriptomics ; epigenomics ; FAANG ; data visualisation

French scientific abstract

Les systèmes de production alimentaire devront devenir plus efficaces et plus durables pour relever les défis posés par la croissance démographique et la crise climatique. Les stratégies de sélection des animaux d'élevage ont permis des gains génétiques remarquables, mais les efforts doivent se poursuivre. La fourniture d'annotations de qualité pour les génomes animaux est nécessaire pour continuer l'amélioration de la valeur génétique des animaux d'élevages. L'initiative FAANG (annotation fonctionnelle des génomes animaux) vise à rassembler une communauté pour promouvoir les principes de données FAIR dans ce domaine¹⁻³. Un centre de coordination des données (DCC) de l'EMBL-EBI développe le portail de données FAANG pour permettre aux données FAANG d'être plus trouvables, accessibles, interopérables et favoriser leur réutilisation (principes FAIR)².

Cette proposition, VizFaDa, vise à produire des visualisations interactives des données FAANG à travers des applications web, et travaillera avec l'EMBL-EBI pour intégrer ces visualisations dans le portail existant. Tout d'abord, nous allons calculer les corrélations entre paires d'échantillons FAANG (expression des gènes et données épigénétiques), et afficher les résultats sous forme de heatmaps clusterisées de corrélation, interactives et filtrables. Les utilisateurs pourront téléverser leurs données traitées pour les comparer aux échantillons FAANG en quelques secondes. Ces heatmaps de corrélation fourniront une vue d'ensemble des données disponibles et de leurs similitudes.

Deuxièmement, nous intégrerons les données épigénétiques et transcriptomiques ensembles en produisant des profils épigénétiques empilés, près du début des gènes, de la fin des gènes et des exons intermédiaires, triés selon le niveau d'expression des gènes ou le taux d'inclusion des exons. Ces visualisations attrayantes exposeront les liens complexes existant entre les marques épigénétiques et la transcription, et ajouteront de la valeur au jeu de données FAANG.

Des efforts seront consacrés à rendre l'ajout de nouveaux échantillons de la base de données aussi automatique que possible, afin d'assurer la pertinence à long terme de nos outils. Le développement des applications web sera entièrement open source. Nous espérons que nos efforts renforceront l'attrait du portail de données FAANG pour les chercheurs et les sélectionneurs, et encourageront la réutilisation de ces données.

English scientific abstract

Food production systems need to be more efficient and sustainable to tackle the challenges posed by a growing population and a climate crisis. Breeding strategies have proven themselves to be essential in providing genetic gain for livestock, but efforts must go on. Providing quality annotations for animal genomes will be instrumental to further improve genetic gains. The FAANG (functional annotation of animal genomes) initiative aims at gathering a community to foster FAIR data principles in this field¹⁻³. A data coordination center (DCC) at the EMBL-EBI is developing the FAANG data portal to allow FAANG data to be more findable, accessible, interoperable and foster their reuse².

This project, VizFaDa, will produce interactive data visualisations of FAANG data through web applications, and we will work with the EMBL-EBI to integrate those visualisations with the existing portal. First we will compute pairwise correlations between FAANG samples (genes expression and epigenetic data), and render the results in the form of interactive, subset-able, clustered correlation heatmaps. Users will be able to upload their processed data to be compared to FAANG samples within seconds. Correlation heatmaps will provide an eagle-eye view of the data available and their similarities. Second, we will integrate epigenetic and transcriptomic data by producing stacked epigenetic profiles near gene starts, gene ends, and middle exons, sorted according to gene expression level or exon inclusion ratio. These attractive visualisations will expose the complex links existing between epigenetic marks and transcription, and will add value to the FAANG dataset.

Efforts will be devoted to make the addition of new samples of the dataset as automatic as possible, to ensure the long term stability of the proposal. Development of the web applications will be fully open source. Altogether, we hope our efforts will reinforce the FAANG data portal attractiveness for researchers and breeders, and will foster data reuse.

Table of content

French scientific abstract	1
English scientific abstract	1
Table of content	2
Proposal description	2
General context	3
State of the art	3
Previous work	3
Tasks	4

WP0: planning and dissemination	5
WP1: Visualisation of correlations across experiments	5
WP2: Epigenetic marks and transcription	5
WP3: extension to non FAANG data	5
Consortium	5
Proposal impact	6
Dissemination strategy and proposal sustainability	7
Self evaluation measures	7
Required budget	7
Planning	8
Bibliographie	8

1. Proposal description

General context

With a growing world population and an ongoing climate crisis, there is a pressure to improve food production systems for efficiency and sustainability. Animal breeding schemes have been successful so far in procuring significant genetic gain, including since the advent of genomic selection, but further improvement are nonetheless required.

The [FAANG](#) initiative¹, for Functional Annotation of Animal Genomes, aims at gathering forces at an international level to improve farm animal genome annotation and foster FAIR principles within the field. Any willing laboratory can submit data and metadata, if they follow the [FAANG data standards](#). Better farm animal genome annotations will be tremendously useful when deciphering the links between QTLs and phenotypes, and will ease the identification of the causal variants. Several european H2020-funded programs are part of the FAANG initiative, including GENE-SWitCH (led by colleagues from INRA), BovReg and AQUA-FAANG.

Data submitted to the FAANG initiative are gathered to a data portal, coordinated by the FAANG data coordination center (DCC) at the EMBL-EBI: data.faang.org. This portal allows simple and complex searches, selection of appropriate files and their subsequent download. It is currently funded through a BBSRC grant until 2020 and by the H2020-funded programs listed above. Future development will include improved API access, availability of uniformly processed data (funded by H2020 and USDA FAANG grants), and inclusion of genome tracks and a genome browser, to allows users to explore FAANG data at their regions of interest with only a web browser and a few clicks.

State of the art

The gold standard for sharing epigenomics data is currently the ENCODE⁴ [data portal](#), which allows powerful search of raw and uniformly processed data across a large selection of assays, species and tissues, and include quality analysis reports. It allows selection of files of interest through a “cart” for bulk downloads, as well as a powerful API access. It also links processed files into genome browsers (UCSC, QuickView and Ensembl).

Other projects have developed data exploration portal, such as the Roadmap Epigenomics⁵ [data portal](#), the BLUEPRINT⁶ [data portal](#), of the IHEC⁷ [data portal](#), with roughly similar features.

Individual laboratories have also developed powerful data portals, often offering additional tools and visualisations, and are a great source of inspiration for this project. The CODEX⁸ [database](#) gathered publicly available data, processed them uniformly, and rendered the results available through a website. It allows to search for transcription factors binding near a gene of interest specified by the user, and provides a few other tools: correlation analysis to compare multiple experiments, gene set enrichment tool, motif discovery, and cross-species comparison of tracks. [OC Genomics](#) is another initiative⁹ that is creating a data portal from publicly available data. It offers extensive quality control analysis, uniformly processed files, a powerful custom genome browser, and several tools, such as a comparator to correlate different experiments, a co-occupancy tools for transcription factors around gene annotation regions, and a temporal gene network tool for expression data.

Previous work

The FAANG community¹ has set the foundations to apply FAIR principles to farm animal genomic data. It has gathered an international community through international meetings, secured several sources of funding through both European and international grants. This initiative started establishing guidelines of best practices, and developing a data portal thanks to the effort of a data coordination center, managed by the EMBL-EBI. Notably, this data portal is [open source](#), and the community is able to support its development.

The coordinator of the VizFaDa project has past experiences that are highly relevant to this program. Amongst other studies¹⁰⁻¹², he developed the [Heat*seq](#) and [PEREpigenomics](#) web applications. Heat*seq¹³ allows fast visualisations of the comparison of hundreds of next generation sequencing experiments (RNA-seq, TF CHIP-seq, CAGE, gene lists) through clustered correlation heatmaps. Users can filter experiments through various fields, and get an overview of the similarities and dissimilarities between experiments. Users can also upload processed data and compare their experiments with hundreds of other publicly available ones in seconds (figure 1A). Heat*seq focused so far on human, mouse and drosophila data. PEREpigenomics explores the relationships between epigenetic marks and gene expression in the Roadmap Epigenomics (human) data by providing stacked profiles plots (figure 1B) for each assay of the dataset. These powerful visualisations allow for a fine comprehension of the relationships between marks and transcription, and are also useful to compare data qualities, tissues or species.

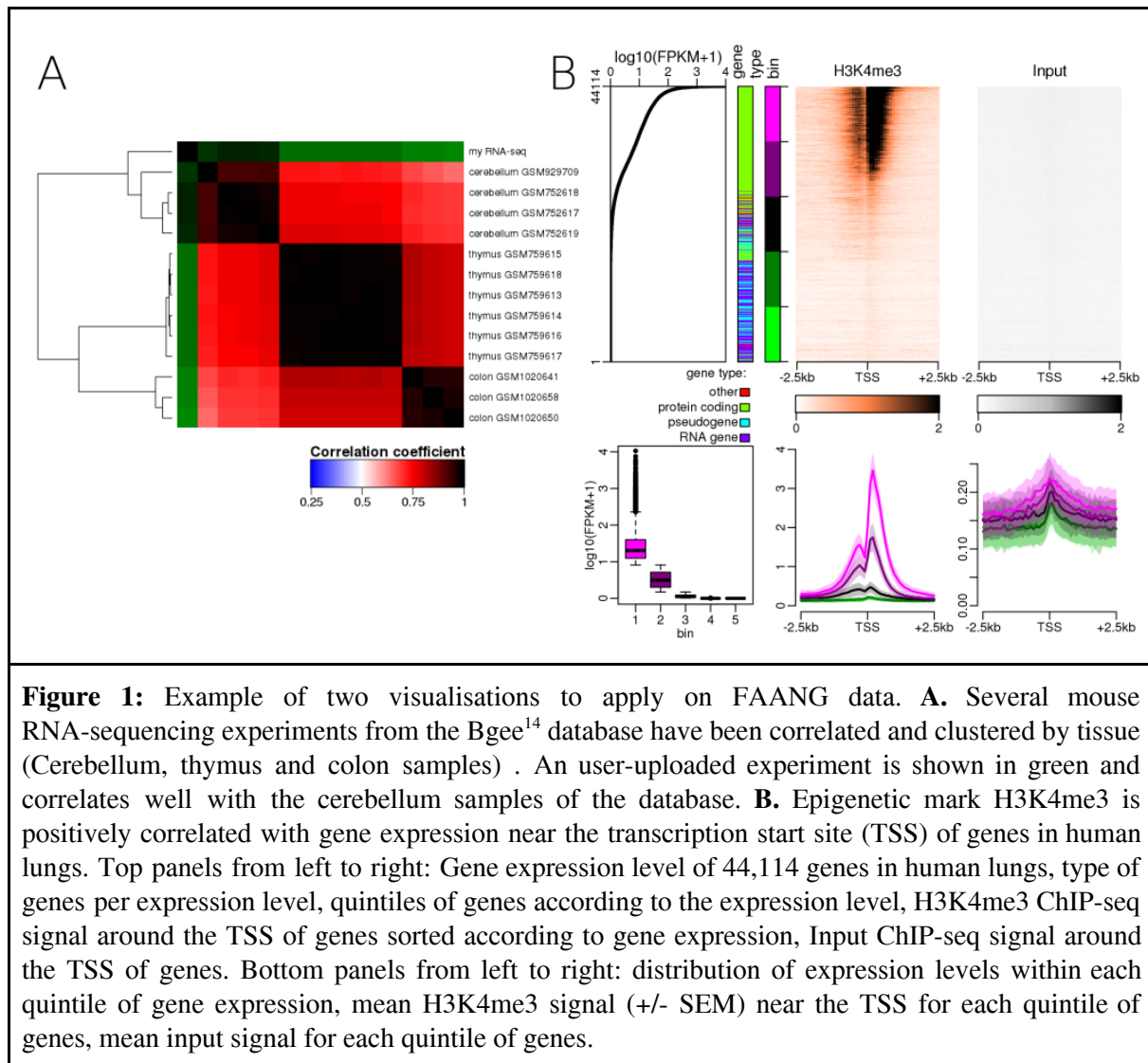


Figure 1: Example of two visualisations to apply on FAANG data. **A.** Several mouse RNA-sequencing experiments from the Bgee¹⁴ database have been correlated and clustered by tissue (Cerebellum, thymus and colon samples) . An user-uploaded experiment is shown in green and correlates well with the cerebellum samples of the database. **B.** Epigenetic mark H3K4me3 is positively correlated with gene expression near the transcription start site (TSS) of genes in human lungs. Top panels from left to right: Gene expression level of 44,114 genes in human lungs, type of genes per expression level, quintiles of genes according to the expression level, H3K4me3 ChIP-seq signal around the TSS of genes sorted according to gene expression, Input ChIP-seq signal around the TSS of genes. Bottom panels from left to right: distribution of expression levels within each quintile of gene expression, mean H3K4me3 signal (+/- SEM) near the TSS for each quintile of genes, mean input signal for each quintile of genes.

Tasks

This proposal, VizFaDa, aims at further reinforcing the attractiveness of the FAANG data portal by providing additional interactive visualisations and tools inspired by our previous work on Heat*seq and PEREpigenomics. The program is structured around 4 packages:

WPO: planning and dissemination

Planning of the work will be decided in close collaboration with the FAANG DCC at the EMBL-EBI, through two in-person meetings and regular video-conferences. The first meeting will happen shortly after the recruitment of a bioinformatics engineer. The second in person meeting will happen toward the middle of the project to serve as an intermediary progress report, to make any adequate changes for the second half of the project.

Dissemination efforts are described in section 4.

WP1: Correlations across experiments

This work package aims at adapting and improving Heat*seq for FAANG data, i.e. at displaying **correlations between experiments** in the form of clustered correlation heatmaps. It will focus on ChIP-seq, DNA methylation profiling, RNA-seq, and ATAC-seq assays in the various FAANG species (so far, bovine, chicken, pig, horse, sheep and goat). For this proposal, each species will be treated independently. Task 1.1 will consist in the development of a web application to display the correlations, and its close integration in the FAANG data portal. Users will be able to go from the heatmap to the assay pages, and *vice versa*, in one mouse click. Users will also be able to upload their own data to compare them with all other relevant data in seconds. Notably, this application will require pre-processing of the FAANG data (alignment, quantification, computation of the correlations). While *ad hoc* pre-processing will be used throughout task 1.1, task 1.2 will focus on automatic addition of new FAANG data into this web application, through the development of automatic routines. This should ensure the longevity of this project, as updating the application with new data will be as automated as possible. We estimate this package to be of low risk (as the coordinator has previously developed such projects) but will benefit greatly to the community.

WP2: Epigenetic marks and transcription

WP1 is focusing in one type of assay at a time. WP2 will thrive at providing **integrative visualisations of both transcriptomic and epigenomic data**. In the form of stacked epigenomic profiles at particular features (gene starts, gene ends, middle exons) sorted according to expression level and/or exon inclusion ratio (figure 1B). Such plots require to have gene expression and epigenetic marks data for the same tissues, which is more and more the case. Each compatible sample will have its links from the data portal to the visualisations (and *vice versa*). Task 2.1 will consist in developing the application, build on our work on the PEREpigenomics web application, and integrate its results in the FAANG data portal. Task 2.2 will consist in automating the addition of new samples in this web application, to ensure its future relevance. Again, we estimate this package to be of low risk (as the coordinator has previously developed such projects) but will benefit greatly to the community.

WP3: extension to non FAANG data

FAANG has a minimal set of requirements to accept new samples (it has nonetheless imported numerous [legacy datasets](#) deemed to be of use to the community). Non-FAANG data might still be very useful to the community. Work package 3 will aim at integrating many non-FAANG data available through databases such as ENA, ArrayExpress, SRA or GEO into web applications developed in WP1 and WP2. Indeed, adding relevant visualisations to already published data is a great way to advertise them and promote their re-use.

2. Consortium

Coordinator: **Guillaume Devailly**, INRA, GenPhySE, GenEpi. Guillaume Devailly has been working in epigenetics and bioinformatics. He is the main developer of web applications [Heat*seq](#)¹³ and [PEREpigenomics](#), that allow researchers to easily explore and visualize large dataset of transcriptomic and epigenomic data. He will supervise WP1 and 2, and do the work for WP3, by dedicating 40% of his time to the project for the 2 years.

A **bioinformatics engineer** will be hired for 18 months at INRA, GenPhySE to work on WP1 and WP2 with the help and guidance of the consortium members.

Philippe Bardou, INRA, GenPhySE, SIGENAE, has experience in web application developments, having notably developed [jvenn](#)¹⁵, [RumimiR](#)¹⁶ and the [ng6](#) database. He will provide help and expertise for the web application development and integration to the FAANG data portal, thanks to his competence in javascript. He will contribute to WP1.1 and 2.1 for an predicted time of 3 person-months.

Patrice Dehais, INRA, GenPhySE, SIGENAE, is part of the team administering the computing infrastructure GenoToul. He is administering several virtual machines, a [Galaxy instance](#), and has adapted Ensembl and biomaRt for local needs. He will provide help and expertise with system administration of the server hosting the developed web applications, contributing to WP1.1 and 2.1 for an predicted time of 3 person-months.

Sylvain Foissac has been deeply involved in the FAANG initiative from the very beginning¹³ and is a co-leader of the FrAgENCODE program¹⁷. He will dedicate 2 person-months to provide guidance in WPO, 1, 2 and 3.

Altogether, several people from the GenPhySE laboratory are involved in the FAANG initiative, and this project will benefit from having them in the working environment.

The local computing infrastructure [GenoToul](#), not part of this consortium, will be able to provide computing power and web server to process data for the web applications, and host and serve the web applications.

Also not part of the consortia, the EMBL-EBI team developing the FAANG DCC data portal (notably Peter Harrison and Guy Cochrane) will provide support to this project, as evidenced by the following letter of support:



Guy Cochrane, Team Leader and
Head of European Nucleotide Archive
E-mail: cochrane@ebi.ac.uk

25 May 2019

Dear Dr. Devailly,

I write to express my enthusiasm and support for your proposed VizFaDa project. I expect that, if funded, your work will provide substantial value to the user community around FAANG data.

As the operators of the FAANG Data Coordination Centre, we at EMBL-EBI are particularly keen to support close working between our teams and to see appropriate integration of your project's outputs into the data portal environment.

Specifically, we will host an engineer and the project coordinator for two visits to EMBL-EBI with two additional video calls; provide assistance with integration of visualisation outputs into the data portal; provide code review for suggested changes affecting the data portal; and integrate at the appropriate level the visualisations to the data portal as these reach sufficient quality and maturity.

I wish you luck with your proposal and look forward to continuing our collaboration.

Sincerely,

Guy Cochrane, PhD

3. Proposal impact

The FAANG initiative aims at gathering all researchers involved in farm animal genomics at **an international level**.

This proposal aims at providing additional data visualisations to the FAANG data portal. Indeed, without visualisations, users need to download heavy read files (fastq) and process them in order to assess their quality and relevance, a process that requires time, knowledge and computing power. Some users have complained that the data portal is for now under-developed, and effort to improve it are ongoing. Therefore, providing data visualisations directly on the portal will further valorize the data and foster their re-use. The visualisation will also provide an eagle-eye view of the relationships between different samples within the FAANG database. Our core target audience will be biologists, and our web application will not require any bioinformatics skills to be used.

The FAANG data portal has had 1001 unique visitors in 2018, corresponding to 4588 unique site visits. As several successfully funded European and international proposals are ongoing, with the three successful H2020 European grants only starting this summer, the number of scientists using the FAANG data portal is likely to continue to increase in the future. Those visitors will benefit from the new visualisations we will provide. Furthermore, any useful feature added to the FAANG data portal will contribute to its usefulness and attractiveness, resulting in a positive feedback loop attracting more and more visitors.

The visualisation scripts and web applications will be developed with the aim to render the addition of new data as automatic as possible, therefore ensuring both the transposability of this effort to new data and its longevity after the end of the project. Efforts produced here will directly benefit to the previous applications Heat*seq and PEREpigenomics, that focus on human and model organisms data.

Our web applications could also be mobilisable by other data repository, as in plant or microbial communities.

4. Dissemination strategy and proposal sustainability

Once presentable, the web applications will be advertised through social media and blog posts. As the development will occur publicly in a GitHub/GitLab repository, the community will be able to provide feedback and modifications since the early days of the development. At later stage of the program, progress will be exposed through poster and oral presentations in local, national, and at least two international events such as the ECCB or JOBIM conferences. As a key part of the FAANG data portal, the visualisations developed as part of this project will also be promoted by the FAANG DCC at international conferences and workshops, for example at PAG and ISAG conferences.

Finally, one or two open-access scientific paper will describe our work, with journal such as Oxford Bioinformatics, BMC Genomics or JOSS as potential target, either independently, or as part of an article presenting the FAANG data portal.

The FAANG data portal has secured funding for the years to come in various international grants, including three H2020 programs, and might decide to fund follow-up of this work if needed. In addition, most members of this consortium are in permanent positions and will be available to maintain the web applications for the years to come. Finally, if positively received by the community, additional grants might be sought to follow up and expand this work.

All web application, analysis scripts, and processed data files will be render publicly available as soon as created, or shortly after. Raw data is already publicly available at the FAANG data portal. The development will happen publicly on platform GitHub and/or GitLab.

5. Self evaluation measures

Several measures will be put in place to ensure high quality throughout the projects.

In person meetings with the team developing the FAANG-DCC will occur at the start and at the middle of the project to refine objectives, ensure good distributions of the tasks, analyse progress, provide feedback. Regular video-call will ensure the good coordination between the consortium and the FAANG DCC throughout the project.

Code reviewing for the applications will be put in place, both internally between the coordinator and the recruited bioinformatic engineer, and with the EMBL-EBI engineers for any code that will directly impact the FAANG data portal. This code reviewing will happen either through GitLab or GitHub.

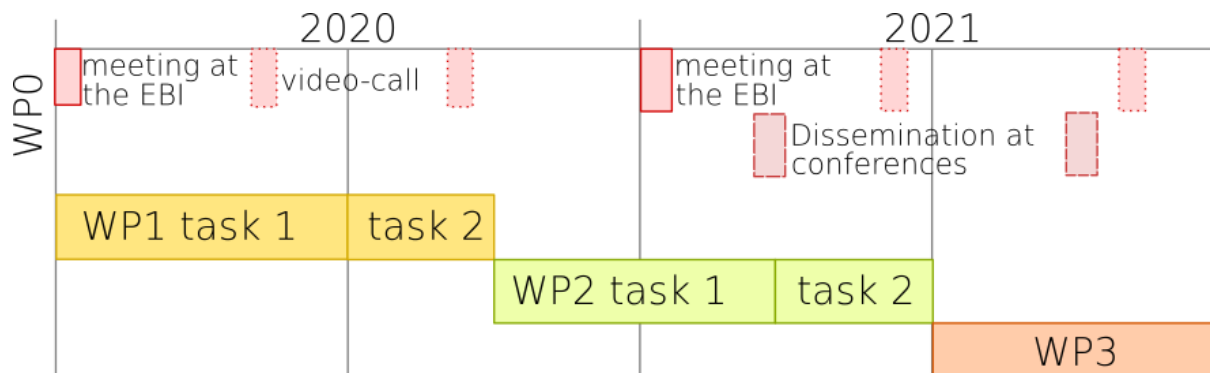
Once early version of the applications will become available, usage will be monitored to assess the effectiveness of our diffusion campaigns and the popularity of each tool, including metrics such as page view number, number of unique visitors, time spent on the applications, both in total and for each sub application, through an GDPR compliant tracker.

Applications will be developed for ease of maintenance and long term perennity in mind. Components (pre-processing pipelines, each application) will be separated to ensure robustness of the system.

6. Required budget

<p>The main spending item will be the recruitment of an engineer for 18 months of time to develop the web applications together with the project coordinator. Additional fundings are required for buying/renting computing and web-server power, travel , meetings and publishing expenses, and overheads.</p>	Budget item	Total amount
	Engineer (18 months)	56 934 €
	Traveling conferences	5 000 €
	Publication costs	500 €
	Computation & server	8 000 €
	Overhead	5 634 €
	TOTAL	76 068 €

7. Planning



A bioinformatics engineer will be recruited for 18 months to work on WP1 and WP2.

Bibliographie

- (1) The FAANG Consortium; Andersson, L.; Archibald, A. L.; Bottema, C. D.; Brauning, R.; Burgess, S. C.; Burt, D. W.; Casas, E.; Cheng, H. H.; Clarke, L.; et al. Coordinated International Action to Accelerate Genome-to-Phenome with FAANG, the Functional Annotation of Animal Genomes Project. *Genome Biol.* **2015**, *16* (1). <https://doi.org/10.1186/s13059-015-0622-4>.
- (2) Harrison, P. W.; Fan, J.; Richardson, D.; Clarke, L.; Zerbino, D.; Cochrane, G.; Archibald, A. L.; Schmidt, C. J.; Flicek, P. FAANG, Establishing Metadata Standards, Validation and Best Practices for the Farmed and Companion Animal Community. *Anim. Genet.* **2018**, *49* (6), 520–526. <https://doi.org/10.1111/age.12736>.
- (3) Giuffra, E.; Tuggle, C. K.; FAANG Consortium. Functional Annotation of Animal Genomes (FAANG): Current Achievements and Roadmap. *Annu. Rev. Anim. Biosci.* **2019**, *7* (1), 65–88. <https://doi.org/10.1146/annurev-animal-020518-114913>.
- (4) Davis, C. A.; Hitz, B. C.; Sloan, C. A.; Chan, E. T.; Davidson, J. M.; Gabdank, I.; Hilton, J. A.; Jain, K.; Baymuradov, U. K.; Narayanan, A. K.; et al. The Encyclopedia of DNA Elements (ENCODE): Data Portal Update. *Nucleic Acids Res.* **2018**, *46* (D1), D794–D801. <https://doi.org/10.1093/nar/gkx1081>.
- (5) Roadmap Epigenomics Consortium; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative Analysis of 111 Reference Human Epigenomes. *Nature* **2015**, *518* (7539), 317–330. <https://doi.org/10.1038/nature14248>.
- (6) Fernández, J. M.; de la Torre, V.; Richardson, D.; Royo, R.; Puiggròs, M.; Moncunill, V.; Fragkogianni, S.; Clarke, L.; Flicek, P.; Rico, D.; et al. The BLUEPRINT Data Analysis Portal. *Cell Syst.* **2016**, *3* (5), 491–495.e5. <https://doi.org/10.1016/j.cels.2016.10.021>.
- (7) Bujold, D.; Morais, D. A. de L.; Gauthier, C.; Côté, C.; Caron, M.; Kwan, T.; Chen, K. C.; Laperle, J.; Markovits, A. N.; Pastinen, T.; et al. The International Human Epigenome Consortium Data Portal. *Cell Syst.* **2016**, *3* (5), 496–499.e2. <https://doi.org/10.1016/j.cels.2016.10.019>.
- (8) Sánchez-Castillo, M.; Ruau, D.; Wilkinson, A. C.; Ng, F. S. L.; Hannah, R.; Diamanti, E.; Lombard, P.; Wilson, N. K.; Gottgens, B. CODEX: A next-Generation Sequencing Experiment Database for the Haematopoietic and Embryonic Stem Cell Communities. *Nucleic Acids Res.* **2015**, *43* (D1), D1117–D1123. <https://doi.org/10.1093/nar/gku895>.
- (9) Mendoza-Parra, M. A.; Saleem, M.-A. M.; Blum, M.; Cholley, P.-E.; Gronemeyer, H. NGS-QC Generator: A Quality Control System for ChIP-Seq and Related Deep Sequencing-Generated Datasets. In *Statistical Genomics*; Mathé, E., Davis, S., Eds.; Springer New York: New York, NY, 2016; Vol. 1418, pp 243–265. https://doi.org/10.1007/978-1-4939-3578-9_13.
- (10) Devailly, G.; Mantsoki, A.; Michoel, T.; Joshi, A. Variable Reproducibility in Genome-Scale Public Data: A Case Study Using ENCODE ChIP Sequencing Resource. *FEBS Lett.* **2015**, *589* (24 Pt B), 3866–3870. <https://doi.org/10.1016/j.febslet.2015.11.027>.
- (11) Devailly, G.; Joshi, A. Insights into Mammalian Transcription Control by Systematic Analysis of ChIP Sequencing Data. *BMC Bioinformatics* **2018**, *19* (Suppl 14), 409. <https://doi.org/10.1186/s12859-018-2377-x>.
- (12) Mantsoki, A.; Devailly, G.; Joshi, A. Dynamics of Promoter Bivalency and RNAP II Pausing in Mouse Stem and Differentiated Cells. *BMC Dev. Biol.* **2018**, *18* (1), 2. <https://doi.org/10.1186/s12861-018-0163-7>.
- (13) Devailly, G.; Mantsoki, A.; Joshi, A. Heat*seq: An Interactive Web Tool for High-Throughput Sequencing Experiment Comparison with Public Data. *Bioinforma. Oxf. Engl.* **2016**, *32* (21), 3354–3356. <https://doi.org/10.1093/bioinformatics/btw407>.
- (14) Bastian, F.; Parmentier, G.; Roux, J.; Moretti, S.; Laudet, V.; Robinson-Rechavi, M. Bgee: Integrating and Comparing Heterogeneous Transcriptome Data Among Species. In *Data Integration in the Life Sciences*; Bairoch, A., Cohen-Boulakia, S., Froidevaux, C., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; Vol. 5109, pp 124–131. https://doi.org/10.1007/978-3-540-69828-9_12.
- (15) Bardou, P.; Mariette, J.; Escudé, F.; Djemiel, C.; Klopp, C. Jvenn: An Interactive Venn Diagram Viewer. *BMC Bioinformatics* **2014**, *15* (1), 293. <https://doi.org/10.1186/1471-2105-15-293>.
- (16) Bourdon, C.; Bardou, P.; Aujean, E.; Le Guillou, S.; Tossier-Klopp, G.; LE PROVOST, F. RumimiR: A Detailed MicroRNA Database Focused on Ruminant Species. *bioRxiv* **2019**. <https://doi.org/10.1101/561720>.
- (17) Foissac, S.; Djebali, S.; Munyard, K.; Vialaneix, N.; Rau, A.; Muret, K.; Esquerré, D.; Zytnecki, M.; Derrien, T.; Bardou, P.; et al. Transcriptome and Chromatin Structure Annotation of Liver, CD4+ and CD8+ T Cells from Four Livestock Species. *bioRxiv* **2019**. <https://doi.org/10.1101/316091>.