AAPG2021	*Geisha*		PRC
Coordinator	**Abderrahman Khila**	48 Months	**578 224,64 €**
CES 13 - *Biologie cellulaire, biologie du développement et de l'évolution*			

Origin of evolutionary innovations:
from molecules and cells to organismal function

Summary table of persons involved in the project:

Partner (P)	Name	First name	Current position	Role & responsibilities	Involvement (person.month) throughout the project's total duration
***P1**: IGFL, UMR5242*	*KHILA*	*Abderrahman*	*DR2 CNRS*	***Leads Aim2:** Cellular and developmental genetics, Evolution*	*20*
***P1**: UMR 5242, Lyon*	*Viala*	*Séverine*	*IE CNRS*	*RNAi/Crispr-Cas9, in situ hybridization screen*	*15*
***P1**: UMR 5242, Lyon*	*Mendes*	*Juliette*	*Tech CNRS*	*Insect husbandry*	*7*
***P1**: UMR 5242, Lyon*	*Roux*	*Pascale*	*Tech INRAE*	*Dissection, RNAi and in situ hybridization screen*	*15*
***P1**: UMR 5242, Lyon*	*Bonneton*	*François*	*MC U. Lyon1*	*EvoDevo*	*7*
***P1**: IGFL, UMR 5242*	*Gillet*	*Benjamin*	*IR CNRS*	*scRNA-seq*	*5*
***P1**: IGFL, UMR 5242*	*Hughes*	*Sandrine*	*CR CNRS*	*scRNA-seq*	*5*
***P2** : UMR5239, Lyon*	*Francesconi*	*Mirko*	*CR Inserm*	***Leads Aim1:** Systems Biology, Modeling, Single Cell-sequencing analyses*	*10*
***P2** : UMR5239, Lyon*	*Modolo*	*Laurent*	*IR CNRS*	*NGS and statistics*	*10*
***P2** : UMR5239, Lyon*	*Janczarsky*	*Stephane*	*IE CNRS*	*NGS and statistics*	*5*
***P3** : UMR7261, Tours*	*Casas*	*Jérôme*	*Prof. U. Tours*	***Leads Aim3:** Applied physics and biomechanics*	*7*
***P3** : UMR7261, Tours*	*Pineirua*	*Miguel*	*MC U. Tours*	*Applied physics and biomechanics*	*7*
***P3** : UMR7261, Tours*	*Steinmann*	*Thomas*	*IR CNRS*	*Applied physics and biomechanics*	*15*

Any changes that have been made in the full proposal compared to the pre-proposal
No changes have been made in comparison with the pre-proposal.

Letter to panel members:

We would like to thank the panel for their assessment of the short version of this project and for their constructive criticism. Please note that this is a revised version of a full project that was declined in 2020. We have addressed concerns raised by reviewers of last years' version and now provide new data and specific responses about the consortium and aspects of fan 'origin' that the current panel members raised. Specifically:
***Panel critiques about the consortium**: We specify through each aim and on the title of the tasks which partner is leading and which partner provides support. We inserted a text box (in green, **pages 9, 11, 13**) at the end of each aim also explaining who is in charge. We added **Figure 9** (**page 14**) which also shows a schematic of how the project will be implemented. Finally, we write, for each task, the partner or the group of partners in charge in the Gantt chart in **page 16**.*
***Panel critiques about the origin of the fan**: This is an important question that can be viewed at multiple levels: molecular, cellular and developmental, as well as phylogenetic. We already know some aspects of the origin of the fan (see Santos et al. 2017). We have collected new data, now presented in **Figure 3** (**page 5**) showing that the fan originates from a conserved structure called the arolium. We have detailed, in **pages 4** and **5**, various aspects of fan origin and explained how this state of knowledge allowed us to streamline our hypotheses and expectations.*
We hope that these explanations meet the expectations of the panel members and the reviewers.

Origin of evolutionary innovations:
from molecules and cells to organismal function

I. Proposal's context, positioning and objective(s)

a. Objectives and research hypothesis

General objectives: The **'*Geisha'*** project will determine the cellular, genetic, developmental bases as well as the functional morphology of a striking evolutionary innovation: the **propelling fan** of water striders (Santos et al., 2017).

Synopsis: Evolutionary innovations are qualitatively new and beneficial phenotypes that allow the bearing lineages to access previously unexploited ecological opportunities (Liem, 1973; Losos, 2009; Wagner and Lynch, 2010). Studying these traits offers a unique opportunity to understand how novelty arises and evolves (Wagner, 2017; Wagner and Lynch, 2010). Although many iconic examples have been documented, it is still difficult to study the origin of evolutionary innovations for various reasons. For example, systems bearing striking evolutionary innovations may be intractable for experimentation, or the innovation itself may be too complex for manipulation. Furthermore, understanding innovations requires the consorted contribution of multiple disciplines to cover the various aspects of these traits, from molecular, developmental and cell biology to evolutionary ecology and functional morphology.

We have previously established the propelling fan, a novel trait found in all members of the water strider subfamily Rhagoveliinae, as a model for studying the origin and evolution of evolutionary innovations (Santos et al., 2017). The fan is not found outside the Rhagoveliinae and plays a critical role in the ability of this lineage of water striders to occupy fast flowing streams (Andersen, 1982; Santos et al., 2017). **The goal of this project** is therefore to uncover the cellular and genetic mechanisms underlying fan development and evolution, as well as fan function at the organismal level. To this end, we gathered a team of experts in developmental genetics and evolution, systems biology, machine learning and modeling, and applied physics and biomechanics. The project is organised in three complementary aims:

Aim-1 *will determine the cellular and molecular mechanisms underlying fan development.*
Aim-2 *will determine the developmental function of the genes that impart fan identity during development and evolution.*
Aim 3 *will determine the biomechanical properties of the fan and its contribution to locomotion performance.*

Expected results: A thorough understanding of the molecular, cellular and evolutionary mechanisms as well as functional morphology underlying the emergence of evolutionary innovations.

b. Position of the project as it relates to the state of the art

b.1. Background: Evolutionary innovations are defined as '***qualitatively new phenotypic traits that endow their bearers with new, often game-changing abilities***' (Wagner, 2011). Classic examples of evolutionary innovations include flowers, wings, feathers, eyes, or Cichlid pharyngeal jaws, each of which opened new ecological opportunities for the bearing organism (Losos, 2009; Wagner and Lynch, 2010). Evolutionary innovations evolve *de novo* and are absent from the immediate ancestor of their bearing lineage. Typically, there is a tight match between the morphology of the innovation and its functional properties, often revealing clues about the selective environment that led to its emergence. In addition, innovations are often associated with bursts of diversification and niche expansion allowing the acquisition of under-exploited ecological opportunities (Losos, 2009; Schluter, 2000).

There is a body of literature that outlines the important questions to be addressed if we are to achieve a systematic understanding of innovations (Losos, 2009; Schluter, 2000; Wagner, 2011; Wagner and Lynch, 2010). Chief among these is to understand how fitness is preserved while organisms explore novel adaptive phenotypes (Wagner, 2011). This requires that pre-existing developmental genetic functions are preserved while new mutations accumulate and introduce changes into the system. In this regard, our understanding of evolutionary innovations is best served when different levels of biological organization, from molecules to cells to organism, are included (Losos, 2009; Schluter, 2000; Wagner, 2011; Wagner and Lynch, 2010). Finally, because of its adaptive nature and strict match with the

selective environment, studies of innovation must address how selection shapes the phenotype. To do so, such studies would greatly benefit from the input of functional morphology experiments that can be powerful in understanding how the innovation imparted a selective advantage on its bearers within the environment where they specialise.

Studies of innovations encompass a wide range of plant and animal taxa. While our understanding of the role of innovations in driving diversification and speciation has been accumulating in the past years (Ronco et al., 2021), addressing the cellular and developmental genetic mechanisms underlying the emergence of innovations has been particularly difficult (Wagner, 2011; Wagner and Lynch, 2010). Furthermore, our understanding of the functional relevance and ecological forces driving the evolution of innovations and their connection to the mechanistic aspects of the trait remains limited (Losos, 2009; Wagner and Lynch, 2010). This difficulty can be explained by two primary observations. First, natural systems bearing striking evolutionary innovations usually lack tools and technologies otherwise available in standard model species. Second, various fields of expertise are required to cover the multiple aspects of evolutionary innovations, from molecular, developmental and cell biology to evolutionary ecology and functional morphology. Overcoming these hurdles requires the establishment of natural study species that bear striking innovations as laboratory models where state of the art tools of cellular and developmental genetics, along with systems biology and animal behaviour, can be applied (Francesconi and Lehner, 2014; Santos et al., 2017; Steinmann et al., 2021a)). **The generalization of new technologies, such as single cell sequencing or Crispr/Cas9 genome editing (Figure 1), combined with the established tools of cell and developmental genetics make water striders great models to study evolutionary innovations.**

Crispr/Cas9 genome editing in water striders

This is the newest addition to the already large technological arsenal in water striders. Because embryos are difficult to inject, we have recently optimized a genome editing technique consisting of targeting the Cas9/guide RNA complex to the ovaries of adult females (Chaverra-Rodriguez et al., 2018). To do this, we tagged a Cas9-GFP protein with a conserved water strider Vitellogenin ligand (Vtg) that specifically binds to the Vitellogenin Receptor (Roth et al., 2013), resulting in a Vtg-Cas9-GFP recombinant protein. Adult females were injected with this ribonucleoprotein complex and their F0 embryo progeny sorted based on GFP signal. These embryos were allowed to develop and later screened for mutants (**Figure 1)** *(Viala et al. in preparation*). Note that the *Gerris buenoi* Vtg tag used here is identical across all water striders, and successfully produced mutants in another species; *Limnogonus frnaciscanus* (**Figure 1 B and C**).

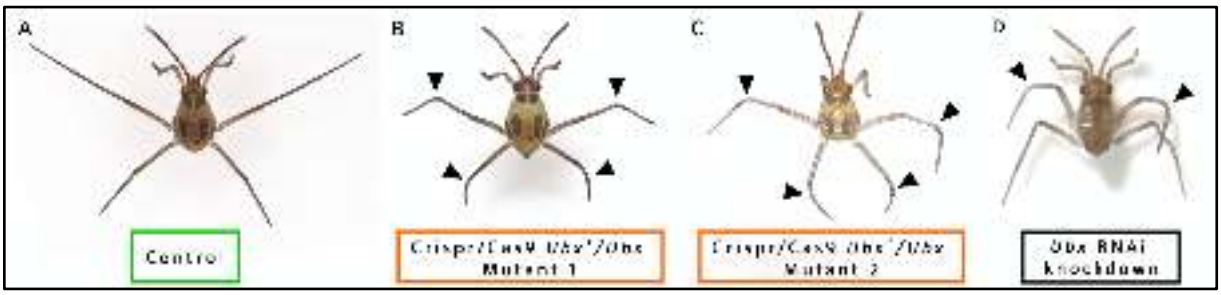

Figure 1: Crispr/Cas9 genome editing in water striders through targeted delivery to the ovary (Chaverra-Rodriguez et al., 2018). This method generates **heterozygous mutants** at F0 by editing the maternal chromosome. **(A)** Normal 1st instar nymph of the water strider *Limnogonus franciscanus*. **(B, C)** Two distinct heterozygous F0 *Ubx* mutants showing the typical shortening and curving of the legs (arrowheads). **(D)** Phenotype of *Ubx* RNAi knockdown in *Gerris buenoi* for comparison (Khila et al., 2009).

b.2. Study system: The ***Geisha*** project will study the multifaceted origin of evolutionary innovations using the propelling fan, found in *Rhagovelia* and its sister genus *Tetraripis* (some taxonomists consider this genus to be *Rhagovelia*) (Andersen, 1982; Santos et al., 2017), as a study system (**Figure 2**). The fan in *Rhagovelia* is composed of ~20 plume-like structures that can be deployed or retracted as the animal rows on the water (Santos et al., 2017). *Tetraripis* however have two pairs of fans, one in the mid-legs (**Figure 2C**) and another in the hind-legs (**Figure 2D**) (Andersen, 1982). Interestingly, *Tetraripis* fans exhibit a bush-like branching organization as opposed to the plumy fans of *Rhagovelia*

(**Figure 2B**). Because there is a single origin of the fan that maps to the Rhagoveliinae (see below), this indicates that the fan diversified subsequent to its emergence.

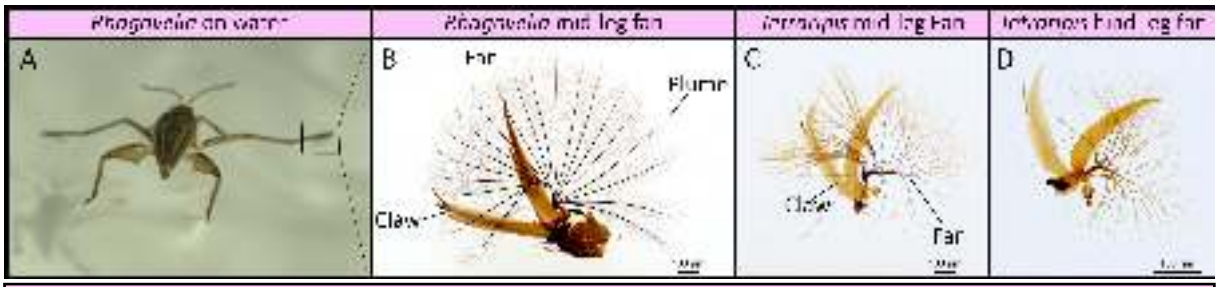

Figure 2: *Rhagovelia* possess a pair of plumy fans in the mid-legs (**A**, **B**) while *Tetraripis* possess two pairs of bushy fans in the mid-legs (**C**) and hind-legs (**D**).

b.3. State of our knowledge about the propelling fan: The fans of both *Rhagovelia* and *Tetraripis* have long been known to taxonomists as speed increasing structures (Andersen, 1982). However, through genetic manipulation and behavioural experiments in *Rhagovelia*, we have shown that the fan does not increase speed but rather allows the water strider to sustain permanent movement lifestyle in fast flowing streams. Therefore, the fan has evolved in association with the specialization of the Rhagoveliinae to fast flowing streams, an environment that is not accessible to fan-less water striders (Santos et al., 2017).

b.3.1 Phylogenetic origin of the fan: The family Veliidae contains 61 genera and about 900 species (Polhemus and Polhemus, 2008). Interestingly, the genus *Rhagovelia* alone contains almost half of the species count of the entire family (over 400), all of which have one pair of plumy fans in the mid-legs (**Figure 1B**) (Andersen, 1982; Andersen, 2000). This observation further supports the hypothesis that the fan contributed, at least partially, to the burst of speciation in this lineage. The fan is exclusive to the sub-family Rhagoveliinae that also includes *Tetraripis* (with two pairs of bushy fans), and is not found anywhere else outside this sub-family. This maps the evolutionary origin of the fan to the time when the Rhagoveliinae split from its closest known subfamily Veliinae, about 45 million years ago (*Armisen et al. in preparation*). There is a single origin of the fan, followed by diversification in terms of shapes and number of pairs. This is the rationale for our choice of three study species, which cover presence *vs.* absence, plumy *vs.* bushy, and one *vs.* two pairs of fans.

b.3.2 Developmental origin of the fan: We now know, based on the new results presented in **Figure 3**, that **the fan originates from the modification of the arolium** during development. The arolium, found in a wide range of insects, is a structure that develops between the two claws, in all three pairs of legs, and that plays an important role in adhesion to substrates (Frantsevich et al., 2008). The first and the third legs of *Rhagovelia* possess an arolium that is similar to that of fan-less species (**Figure 3A and D**), whereas this structure is replaced by a fan in the second leg of *Rhagovelia* (**Figure 3B**), and in the second and third leg of *Tetraripis* (**Figure 2C and D**). This indicates that the arolium is the ancestral structure and that the fan is the derived state. Inactivating the Hox gene *Sex-combs-reduced (Scr),* using RNAi*,* caused the arolium of the first leg to take the identity of a fan (**Figure 3D**). Similarly, inactivating the Hox gene *Ultrabithorax (Ubx)* also caused the arolium of the third leg to take the identity of a fan (**Figure 3F**). Please note that in the second legs, neither *Ubx* (**Figure 3E**) nor *Scr* (not shown) RNAi affected the fan, even though *Ubx* depletion modified the identity of one of the claws (compare green and black arrowheads in **Figure 3B and E**). These new results demonstrate that **(1) the fan originates during development from the transformation of the arolium**, and **(2) Hox genes repress fan** formation in the first and third legs, but do not regulate fan development in the second legs in *Rhagovelia.* These data have been helpful in streamlining the hypotheses about the cellular and developmental origin of the fan that we outline below.

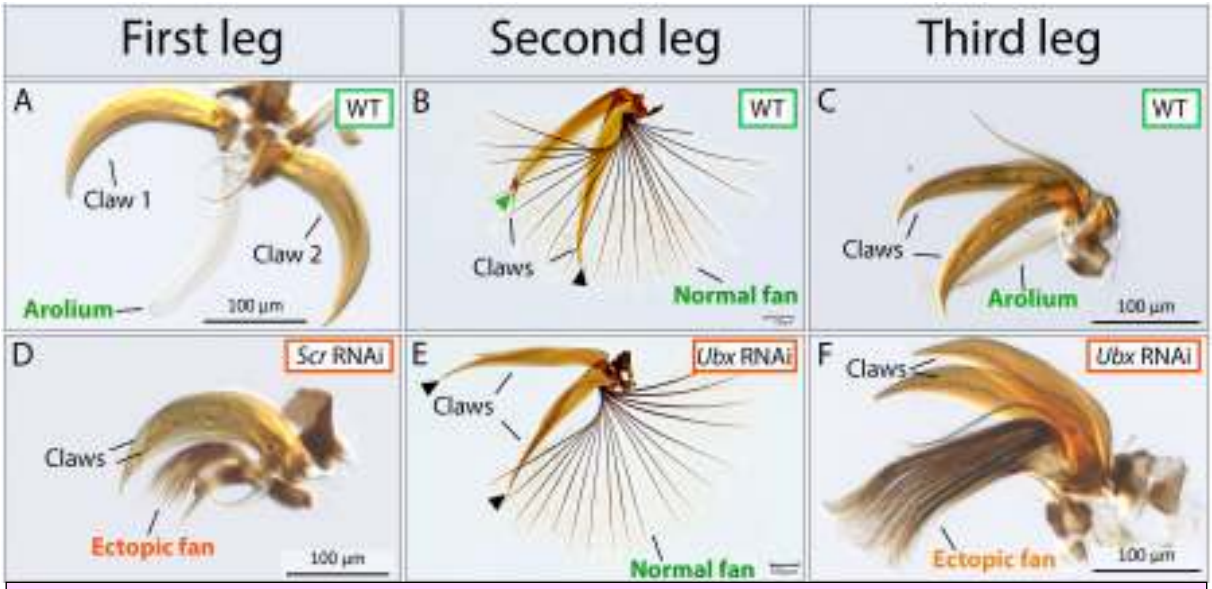

Figure 3: The fan originates from the arolium in *Rhagovelia*. (A) arolium in normal first leg. **(B)** the fan develops normally instead of arolium in second leg. **(C)** arolium in normal third legs. **(D)** *Sex-combs-reduced* RNAi transforms the arolium into a fan in the first leg. **(E)** *Ubx* RNAi does not affect the fan in the second leg. Note that the identity of one of the claws is transformed (compare green arrowhead in **B** with black arrowheads in **E**). **(F)** *Ubx* RNAi transforms the arolium into a fan in the third leg.

b.3.3. Molecular origin of the fan: The fan develops during embryogenesis and is replaced at each of the five nymphal molts (Santos et al., 2017). We now know that the fan in *Rhagovelia* is under the control of the two paralogous genes *geisha* (*gsha*) and *mother-of-geisha* (*mogsha*) (Santos et al., 2017). *gsha* is found exclusively in *Rhagovelia* (we do not know yet if this gene is found in *Tetraripis*), and both genes are expressed in the cell cluster that forms the fan but not those that form the arolium in *Rhagovelia* (Santos et al., 2017). In addition, these two genes, along with three additional markers of fan cells, are not expressed in fan-less species (Santos et al., 2017). This means that two key events led to the evolution of the fan in *Rhagovelia*: a gene duplication leading to the emergence of the *gsha* gene and a regulatory change that activated both paralogs, along with other genes, exclusively in the cluster of cells that form the fan at the tip of the mid-legs (Santos et al., 2017) (**Figure 4**). However, many important questions remain unanswered: what is the evolutionary sequence of these events? *gsha* and *mogsha* being transmembrane proteins, how do they become activated in fan cells and which genes do they regulate there? How is the activity of the components of fan network differs from that of the arolium? How does the development of the plumy fans in *Rhagovelia* compare to that of the bushy fans in *Tetraripis?* What does the fan contribute functionally to these organisms? And finally, is there any functional significance to having two pairs of fans rather than one?

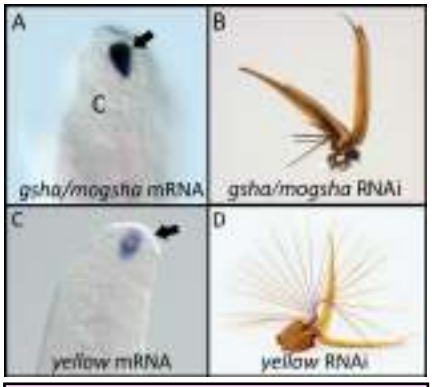

Figure 4: The fan originates from a cluster of cells in T2 tarsus. (A) Expression of the novel paralogous genes *geisha* and *mother-of-geisha* in fan cells found in the tarsus of the mid-leg. **(B)** *gsha/mogsha* RNAi depletes the fan specifically, and does not affect the claws. **(C)** Expression of the gene *yellow* early in fan cells. **(D)** *yellow* RNAi results in loss of fan pigmentation and rigidity.

b.4. Research hypotheses:

- **Cell and developmental genetics hypothesis:** The cells at the origin of the fan represent a modified version of the cells that normally give rise to the arolium and express a novel gene network including the *geisha* and *mother-of-geisha* genes (Santos et al., 2017).

- **Functional morphology and evolution hypothesis:** The plumy fans represent an improved version of the bushy fans that evolved under strong and persistent selection by the fast-flowing stream environment.

c. Methodology and risk management

Aims of the proposal: ***Geisha*** is organized in three aims covering the three important properties of evolutionary innovations: **(1)** Cellular, molecular, developmental and evolutionary origin; **(2)** Functional analyses and validation of the molecular mechanisms underlying fan development and evolution, and **(3)** functional relevance of the fan at the organismal level.

Aim-1: Cellular, molecular and developmental origin of the propelling fan

Leading partner: Partner 2 **Francesconi**
Supporting partner: Partner 1 **Khila** (Khila will provide the technical know-how and the already optimized infrastructure in relation with the insects to this aim)

Rationale: This aim will be primarily based on high throughput single cell RNA sequencing technology (scRNA-seq), combined with the expertise of the **Francesconi** lab in systems biology and machine learning, to address two important questions about the origin of the fan: **(1)** How arolium cells evolved into fan cells and **(2)** How gene activity directs the cells toward fan, rather than arolium development? Single cell RNA-seq data, combined with our pre-existing data on specific markers for fan cells and the neighbouring cells in the legs **(Figure 4 and Figure 5)**, constitutes a powerful method to characterize cell types, cell states, and possibly a putative network underlying the development and evolution of this innovation. The aim will use *Rhagovelia antilleana* (one pair of plumy fans; **Figure 2A-B**) as a focal species, in comparison with *Tetraripis zitteli* (two pairs of bushy fans; **Figure 2C-D**), and the closest known fan-less sister species *Oiovelia cunucunumana* (Santos et al., 2017). We subdivide **aim 1** into the following complementary tasks:

- ***Task 1- Insect breeding, dissection and leg tissue acquisition (to be conducted in Khila lab):***
The lab of the coordinator (**Khila**) has the necessary infrastructure and know-how for inbreeding many water striders including the focal species of this proposal (Santos et al., 2017; Vargas-Lowman et al., 2019). The *Rhagovelia* fan develops from a cluster of about 20 cells at the tip of the mid-legs. In an *in situ* hybridization screen examining dozens of transcripts that were enriched in the mid-legs, we have identified five genes, so far, that are expressed exclusively in this cell cluster: *geisha* and *mother-of-geisha* (**Figure 4A**), *yellow* (**Figure 4C**), *cuticular protein 19*, *Coiled-Coil Domain Containing 174* (**Figure 5A-B**). In addition, this screen identified a number of genes that are expressed in various cell types, other than fan cells, along the anterior-posterior and proximal-distal axes of the tarsus in *Rhagovelia* (**Figure 5**). This task will conduct single cell sequencing on the tarsi of the mid-legs (fans) and hind-legs (arolium) to reduce complexity and compare the various cell types in these tissues.

- ***Task 2- Target tissue and cell dissociation (to be conducted by the IGFL sequencing facility):***
We will focus on the tarsi for two main reasons. First, because in *Rhagovelia* the tip of the mid-leg is enriched with **fan cells**, and the tarsi of the hind-legs with **arolium cells** (Santos et al., 2017) (**Figure 4A, C; Figure 5A-B**). This will allow us to avoid the high complexity of whole animals and, therefore, enrich our high throughput sequencing datasets with the components of the targeted gene networks. Second, the results of our *in situ* hybridization screen (Santos et al., 2017) showed that the tarsus contains multiples cell types based on the markers they express (**Figure 5C-F**). This diversity of genetic marker-expressing cells in the tarsus will allow us to cluster the single cell RNA-sequencing data into distinct groups of cell types. This strategy has the significant advantage of reducing costs, focusing the analysis, and enriching the samples with the targeted cell types without losing the comparative power between various cell types (claws, integument, fans, arolium etc.). Protocols for tissue dissociation into single cells are available in multiple lineages from flies to

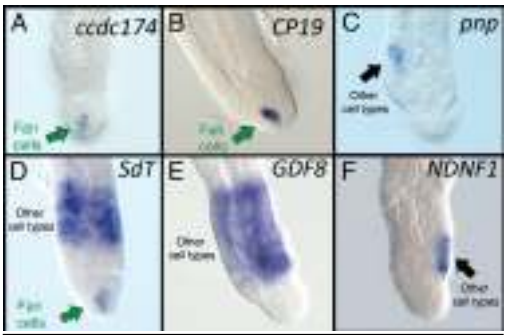

Figure 5: Tarsi of mid-legs contain a diversity of cell types. (A, B) two additional markers exclusively expressed in fan cells. **(C-F)** markers of other cell types across various axes of mid-leg tarsus. Note that some (*e.g.* SDT) may be expressed in fan cells and other cell types. Ccdc: Coil-coil domain containing protein; CP: Cuticular Protein; pnp: purine nucleoside phosphorylase; SdT: Stardust; GDF: Growth differentiation factor; NDNF: Neuron derived neurotrotrophic factor.

planarians (Alles et al., 2017; Plass et al., 2018). We will adapt a dissociation protocol for our samples of dissected tarsi of the second and third legs of *Rhagovelia*, *Tetraripis* and *Oiovelia* (see sampling below). **A first pilot test** will be first conducted by our sequencing facility to determine the amount of starting material and optimise the number of dissociated cells. Then the dissociated cells of the primary sample will be used to prepare libraries for sequencing.

- ***Task 3- Generation of scRNA-seq time series during leg development (to be conducted in Khila lab):*** This task will include *Rhagovelia* (one pair of plumy fans)*, Tetraripis* (two pairs of bushy fans) and the fan less *Oiovelia cunucunumana*. First, we will stage embryos, see (Vargas-Lowman et al., 2019), and conduct time series on the focal species *Rhagovelia antilleana*. To do this, we will build single-cell sequencing datasets on a total of eight developmental stages. We have shown that the first signs of fan development appear by about 35% of embryogenesis (Santos et al., 2017). Therefore, our embryonic time series will include two stages before and four stages after fan markers become expressed (**Figure 4 and 5**). These embryonic time series will be separated by about 5% of embryonic developmental time. We also know that the fan grows and gets replaced at each nymphal moult (Santos et al., 2017). For post-embryonic samples, we chose the first and the second nymphal instars because they are the closest to the embryonic time series. This brings the total of samples in *Rhagovelia* to 16 (2 legs x 8 developmental stages).
  For *Tetraripis* and *Oiovelia*, we will limit our sampling to **two homologous developmental stages**; the Katatrepsis stage and the first nymphal instar. Katatrepsis is a stage around 45% of embryogenesis in water striders, where the embryo makes a 180° turn to occupy the ventral side of the yolk (Vargas-Lowman et al., 2019). These samples will allow us to perform binary comparisons between legs with and without fans, between plumy and bushy fans, and between two pairs and one pair, requiring an additional set of eight samples (2 species x 2 legs x 2 developmental stages). Each sample will contain a pool of 50 legs (or otherwise informed by the pilot test) to ensure that a sufficient number of cells is included. We will target about 1000 to 2000 dissociated cells per sample, which largely covers the cell diversity of the tarsus, and will use 10X genomics and Illumina sequencing.

- ***Task 4- Data analysis and gene regulatory network reconstruction (to be conducted in the Francesconi lab):***
  Task 4- a- Reconstruction of developmental trajectories and cell subpopulation dynamics during development (to be conducted in the Francesconi lab): We will use a recent pipeline (Melsted et al., 2021) to pseudo-align reads to **reference transcriptomes** (already available), assign them to cell of origin and collapse them by unique molecule identifiers (UMIs) to produce both spliced and un-spliced transcript counts, which will also allow us to compute RNA velocity, see below. We will then combine, normalize and batch correct the different libraries (**Figure 6A**) using a canonical correlation analysis (CCA) based procedure in the 'Seurat' toolkit (Butler et al., 2018).

  We will then reconstruct putative cell developmental trajectories during leg development from similarities of single cell transcriptomes. Many algorithms are available for trajectory reconstruction that can detect multiple branching developmental trajectories, such as diffusion maps (Coifman et al., 2005) and diffusion pseudotime (successfully employed by Partner 2 **Francesconi** to reconstruct trans-differentiation and reprogramming trajectories from time series data (Francesconi et al., 2019)) or uniform manifold approximation and projection (UMAP) (Becht et al., 2018) (recently employed in single-cell reconstruction of developmental trajectories during *C. elegans* embryogenesis (Packer et al., 2019) and mouse organogenesis (Cao et al., 2017) and available in the 'Monocle' toolkit (Cao et al., 2017). These methods exploit both intrinsic variability across individuals within each time point and variation across time points to reconstruct continuous developmental trajectories and reorder cells along them (Haghverdi et al., 2016; Trapnell et al., 2014). As no prior time-series data are available in water striders, it is possible that the sampling frequency we initially chose will not be enough to capture fast dynamics and/or to reconstruct fully continuous developmental trajectories for all the cell populations. In this case we will profile the additional time points needed to fill the gaps (**Figure 6B-i**). Furthermore, we will use recently developed algorithms that are able to reconstruct not only static developmental trajectories but also the cell dynamics along those trajectories, such as algorithms that exploit the ratio between spliced and non-spliced RNA to predict the direction and the speed of gene expression change (**RNA velocity**) of each cell along the developmental trajectory (Bergen et al., 2019; La Manno et al.,

2018; Melsted, 2019). In this way we can further distinguish sub-populations of cells close in the developmental trajectory but with very different velocities (**Figure 6B-i**) and we can follow distinct cell subgroups across development. Further, RNA velocity is an additional source of information useful for reconstructing regulatory networks (see Task 4b below). We will then perform differential gene expression analysis, gene clustering (**Figure 6B-ii**) and gene set enrichment analysis (**Figure 6B-iii**) to identify key genes and processes characterizing each cell subpopulation.

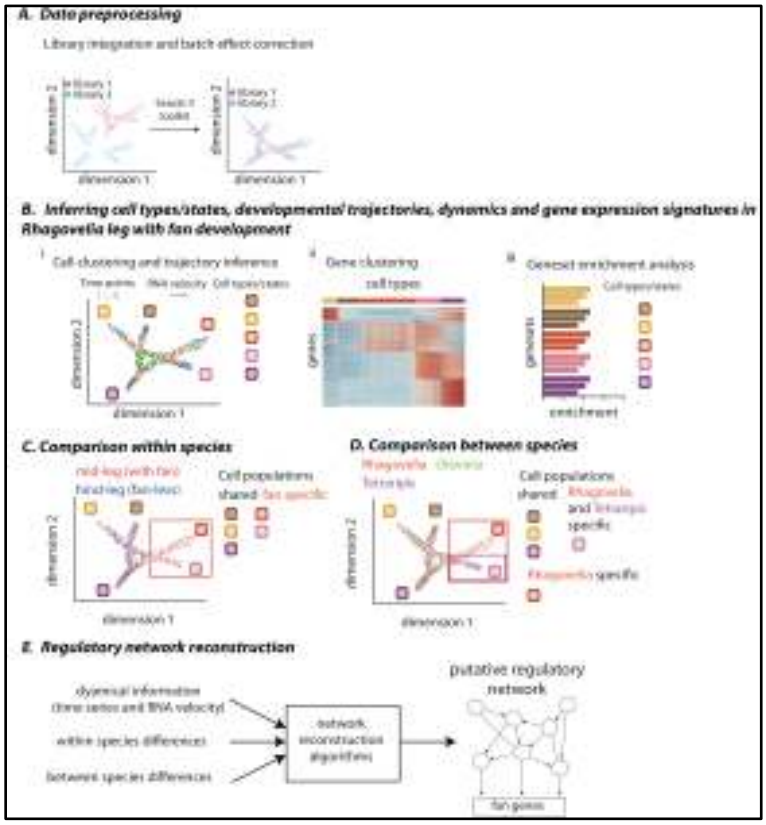

Figure 6: Data preprocessing and data analysis. A) Libraries will be integrated and batch corrected before further analyses. **B)** We will: (i) reconstruct leg with fan developmental trajectories, infer cells RNA velocity and identify cell types from our single cell time series data; (ii) Identify cell specific gene cluster and (iii) characterize them by gene set enrichment analysis. **C)** We will compare leg with fan development to leg without fan in *Rhagovelia*, to identify fan specific cell types and trajectories. **D)** we will compare leg with fan development in *Rhagovelia* to leg development in sister species to identify *Rhagovelia* specific cell types and developmental trajectories. **E)** We will combine this information to reconstruct a putative gene regulatory network.

Task 4- b- Intra- and inter-specific comparisons of cell transcriptional profiles (to be conducted in the Francesconi lab): This task will benefit from existing transcriptomic datasets in the lab of Partner 1 covering both whole and leg transcriptomes (*Armisen et al. In preparations*). We will again use Seurat toolkit (Butler et al., 2018) to compare the leg with the fan (mid-leg) and the leg without fan (hind-leg) within *Rhagovelia* (***Figure 6C***), and legs with a fan across species (***Figure 6D***). The first aim here is to extract cell trajectories and dynamics that are specific for the fan-developing tarsi in *Rhagovelia* accounting for dynamics shared by both types of leg development (***Figure 6C***). The second aim is to compare development from mid-leg tarsi between *Rhagovelia* (plumy fan), *Tetraripis* (bushy fan) and *Oiovelia* (fan-less), in order to identify both shared and species-specific developmental trajectories and dynamics (***Figure 6D***). Using this strategy, we aim at identifying key processes and potential regulators of fan development and understanding how they could have emerged during evolution.

Task 4-c- Gene regulatory network (GRN) reconstruction (to be conducted in the Francesconi lab): An important advantage of adopting single cell transcriptomics approach is that it provides thousands of independent measurements for thousands of genes. This gives enough power to enable causal inference and large-scale gene regulatory network reconstruction (**GRN**), where traditionally samples were far fewer than the number of genes hampering such a task (Marbach et al., 2010). However, despite being highly promising, GRN reconstruction from single cell data remains challenging (*i.e.,* due to the noise and the sparsity of single cell data) and it is currently an area of active research. We will nonetheless try to push the boundaries and attempt to reconstruct, at least partially, the putative regulatory network responsible for the formation of the fan during development (***Figure***

6E). Many algorithms have been recently developed (or adapted) to exploit single cell expression data to reconstruct gene regulatory networks even without additional protein binding data such as PIDC (Chan et al., 2017), Scribe (Qiu et al., 2020), dynGenie3 (Huynh-Thu and Geurts, 2018), GRISLI (Aubin-Frankowski and Vert, 2018) or Bingo (Aalto et al., 2020). Many algorithms exploit dynamic information from pseudo-time, time series design and RNA velocity to improve the accuracy of the network (dynGenie3 (Huynh-Thu and Geurts, 2018) Scribe (Qiu et al., 2020) GRISLI (Aubin-Frankowski and Vert, 2018) Tenet (Kim et al., 2021) or Scode (Matsumoto et al., 2017). Moreover, recent algorithms can exploit phenotypic differences across single cell subpopulations to recover phenotype- or condition-specific gene regulatory networks such as SimiC (Peng et al., 2021). Therefore, we can also jointly analyze our samples and exploit genetic differences across species and developmental differences between the hind and mid legs to recover fan-specific regulatory networks. Networks obtained with different algorithms will be compared and merged to obtain a consensus network which will include genes and regulatory interactions predicted by multiple independent methods as this has been showed to improve network accuracy (Marbach et al., 2012).

What do we learn from aim-1? Identifying the cell state that distinguish fan cell from arolium and other cell types in the tarsus, and building a putative gene network underlying the emergence of the fan during development and in evolution. Also, this aim will help us choose candidate genes and putative genetic interactions to test experimentally in **aim 2**.

Who performs the aim? Partner 2 (**Francesconi**) will take the lead on this aim because his lab has the systems biology, modeling and machine learning expertise required to conduct these experiments. Partner 1 (**Khila**) will provide the necessary help regarding the know-how of rearing and handling and manipulating the insects. Both partners are located at the same campus of ENS de Lyon (IGFL and LBMC), thus making the collaboration highly accessible. A **postdoctoral fellow** with expertise in **bioinformatics and systems biology** will be hired for two years to work within the two teams. This fellow will have support from a lab manager, **Severine Viala**, who has extensive experience in the study and manipulation of water striders to help with dissection and molecular biology. The aim will also benefit from a CNRS animal technician, **Juliette Mendes**, who manages our insect facility. This aim will also greatly benefit from the expertise and the logistics of the newly established Spatial-Cell-ID ***Equipex+ in*** Lyon. This initiative provides integrated single-cell spatio-temporal analysis in strong relation with the national network of genomics, imaging, and bioinformatics infrastructures “France Génomique”, “France BioImaging”, and IFB cloud.

Aim-2: Function of the components of the gene network underlying fan development and evolution:
Leading partner: Partner 1 **Khila**
Supporting partner: Partner 2 **Francesconi** (Francesconi will continue to provide input about gene expression profiles and prioritization of target genes and putative genetic interactions to be experimentally tested).

Rationale: This aim will determine the molecular functions and genetic interactions of a subset of fan network components, and compare these functions across species. The **Khila** lab has strong expertise in *in situ* hybridization and RNAi screens, where dozens of genes can be processed per experience manipulator (Armisen et al., 2015; Toubiana et al., 2020; Vargas-Lowman et al., 2019). Here, we hypothesize that the differences in gene network composition between fan and arolium cells across legs and species will explain the differences in the presence/absence of fans, number of fans, and morphology of the fans (plumy or bushy). These experiments will be executed using a combination of Crispr/Cas9 genome editing (for non-essential genes such as *gsha*) and RNAi knockdown (for essential genes such as *Ubx*) (Crumiere and Khila, 2019; Khila et al., 2012; Santos et al., 2015; Santos et al., 2017; Vargas-Lowman et al., 2019). The aim is subdivided into four complementary tasks that will be conducted primarily in the lab of the coordinator (**Khila**).

- ***Task 5- In situ hybridization screen:*** We will conduct a high throughput *in situ* hybridization screen to validate the components of the gene networks and cell type markers as established across species in **aim 1**. The same approach was conducted in (Armisen et al. 2015; Santos et al. 2017 and Toubiana et al. **PLoS Biology In Press**) and led to the discovery of the *gsha/mogsha* genes and the other markers of fan development in *Rhagovelia* (Santos et al., 2017). The screen will cover transcripts that match the *in-silico* results from **aim 1** and that show expression in fan cells but not arolium cells, or *vice versa*

(as shown in Figures 3 and 4). We have the capability to test dozens to hundreds of genes for the duration of the grant, which will help us understand how differences in gene expression establishes differences in fan states. In addition, our access to Spatial-Cell-ID through Equipex will allow us to perform a combinatorial labeling approach to associate unique barcodes with individual RNA species. Barcodes are then read through a series of sequential hybridization (smFISH) and imaging, and visualize hundreds of genes in the same specimen. This method, though powerful, requires optimisation and, if proved difficult, we will continue to rely on classical *in situ* hybridization.

- ***Task 6- Functional screen of validated genes:*** Here we will use RNA interference, which works on the fans (**Figure 4**) and across water striders (Khila et al., 2012; Santos et al., 2015; Santos et al., 2017; Vargas-Lowman et al., 2019). Transcripts that were validated by in situ hybridization and that showed **specific expression in fan cells, and those known to play regulatory functions, will be given priority in this functional screen.** RNAi knock down and screening for fan phenotypes is also well established in the coordinator's lab (Armisen et al., 2015; Khila et al., 2012; Santos et al., 2015; Santos et al., 2017; Toubiana et al., 2020). Special attention will be given to genes whose knockdowns deplete the fan as seen in ***Figure 3B***, or modify its morphology or composition (e. g. from plumy to bushy). However, all genes that affect the development and evolution of the fan will be analyzed and their effect interpreted. The number of genes to be tested will depend on the results from sc-RNA-seq and in situ hybridization. The **Khila** lab has the capability of performing RNAi screens on dozens of genes per project, see (Santos et al., 2017; Toubiana et al., 2020; Vargas-Lowman et al., 2019)

- ***Task 7- Test of genetic interactions:***

Task 7- a- RNAi/Crispr-Cas9 combined with scRNA-seq: To test the genetic interactions between components of the fan network, we will conduct another single cell sequencing experiment, but this time on tarsi dissected from **Katatrepsis stage** embryos of Crispr-Cas9 mutants or RNAi-treated individuals. The genes *gsha/mogsha* (which activate the fan in T2) will be used to uncover some the downstream genetic interactions underlying the function of these genes during fan development. Because neither *gsha* nor *mogsha* are essential for survival (Santos et al., 2017), we will establish mutant lines where these two genes are knocked out, separately. This is an improved experiment compared to RNAi, which could not discriminate between the two paralogs due to high sequence similarity. To discriminate between *gsha* and *mogsha* with genome editing, we detected single nucleotide polymorphisms (SNPs) in over 10 PAM (Protospacer Adjacent Motif, often NGG) sequences that diverged between *gsha* and *mogsha*. These SNPs in PAMs will allow us to design guide RNAs that are biased toward either one of the two genes for targeting and generate *gsha* and *mogsha* mutants, separately. This experiment will be done both in *Rhagovelia* and *Tetraripis* (for both genes if *Tetraripis* also has *gsha*, or for *mogsha* alone if this gene is exclusive to *Rhagovelia*). Another important objective for this experiment is to uncover possible **upstream regulators of *gsha* and *mogsha***. To achieve this, we will give special attention to **transcription factors that are expressed exclusively in fan cells**. If a protein is responsible for activating *gsha/mogsha* in fan cells, we would expect the loss of expression of these genes when that regulator is inactivated by RNAi, as shown in **Figure 7**. The Crispr/Cas9 mutants or RNAi knockdowns will be used to harvest tarsi of the T2 and T3 legs for single cell RNA sequencing. The data from this experiment will be compared to those obtained from tarsi of normal animals. The results will strengthen our capability of network inference and help identify targets and interactors of *gsha/mogsha* in regulating fan development. This experiment will require 8 additional libraries: 2 genes x 2 legs x 2 species.

Figure 7: Example of test of genetic interaction. (A) The gene gilt is expressed in the mid-leg tarsus of water striders where second legs are longer than third legs (Armisen et al., 2015). **(B)** When we depleted *Ubx* using RNAi, gilt become de-repressed in the third leg (red arrow), indicating that *Ubx* represses *gilt*.

Task 7-b- RNAi/Crispr-Cas9 combined with *in situ* hybridization: In a second approach, we will test the impact of gene knockdown or knockout, using RNAi or Crispr-Cas9, on their putative targets using in situ hybridization (see example in **Figure 7**). This approach consists of inactivating the gene of interest and staining the resulting embryos with a set of probes directed against potential targets. This method works quite well in water striders and allowed us to show, for example, a newly discovered interaction between Ubx and its target *gilt* (**Figure 7B**) (Armisen et al., 2015).

What do we learn from aim-2? These experiments will allow us to assign a function to the genes that are exclusively expressed in fan cells. It will also allow us to validate interactions between genes that were predicted to compose the fan network in each species. Finally, this aim will provide a good understanding of how differences in network composition and gene function result in the transformation of the arolium into fan, and how differences in fan shape, fan number or fan presence/absence are encoded by differences in gene activity.

Who performs the aim? The coordinator **(Khila)** will lead this aim and will hire a **PhD student** to conduct the experiments, in collaboration with the postdoctoral fellow. The student and the Postdoc will also benefit from the support of a lab technician (**Pascale Roux**) who has expertise in molecular biology, immune-histochemistry and RNAi in water striders, and *Maître de Conferences* (**François Bonneton**) who is expert in developmental genetics and EvoDevo.

Aim 3: Biomechanical properties of the fan and its contribution to locomotion performance

Leading partner: Partner 3 **Casas**
Supporting partner: Partner 1 **Khila** (**Khila** will generate animals with genetically manipulated fans to be tested for functional morphology through biomechanics experiments by Casas).

Rationale: Here we hypothesize that fans, with different shapes and number of pairs, implies differences in the way legs are moved and impulses transferred into the water body. We expect to find a clear and positive relationship between the forces harvested from the leg movement, to be computed from the surface waves and water body vortices, and the number and hierarchical levels of the fan and kinematics of movement (Steinmann et al., 2018; Steinmann et al., 2021b).

- ***Task 8- Quantifying and reproducing leg and fan kinematics in 3D***

Task 8- a- Quantifying leg kinematics and fan spreading during propulsive phase: Our first aim is to characterize the typical trajectory and deformation of *Rhagovelia* median propulsive leg (the one with the fan) (**Figure 8**). We will pay special attention to the interaction between the water surface and the fan that generally moves just below the interface. Given the small size and high speed of the movement, we will optimize and apply several techniques available in the lab, such as the quantitative synthetic Schlieren technique (Moisy et al., 2009; Wildeman, 2018) and the recently developed 3D particle tracking technique (3D-PTV) (Steinmann et al., 2021b). This last one is especially adapted to the coupled measurement of free surface topography together and simultaneously with the flow velocity beneath the surface (***Figure 7***).

Task 8- b- Quantifying leg morphology and wettability: High precision tomographic scanning (Micro-CT and Nano-CT scans) will be done on different types of fan states produced through phenotypic manipulation by the coordinator (**Khila**): normal fans, surgically removed fans (but keeping the claws), depleted fans through *gsha/mogsha* RNAi (the phenotype is highly specific to the fan and does not affect claws, see **Figure 4B**) or flexible fans (*yellow* RNAi, **Figure 4D**) (Santos et al., 2017). The wettability of the legs will be tested using high-precision tensiometry techniques (Gao and Jiang, 2004; Shi et al., 2007).

Task 8-c- 3D printing of scaled-up mechanical models of the fan

Printing a 3D model will allow us to study the phenomena at higher magnification using scaled-up 3Dprinted legs (**Figure 8**). The 3D printing will be carried out on the basis of the scans. We will do this in collaboration with Prof. G. Krijnen (NL) who is a specialist in these aspects, and with whom Partner 3 has collaborated in the past on similar questions (Jaffar-Bandjee et al., 2018)

Task 8-d- Experiments with natural and 3D printed legs

Our aim is to reproduce accurately the movement of the propulsive leg of the insect with a mechanical leg. The mechanical simulation of this complex and rapid movement requires making the artificial leg move at high velocity in three dimensions.

- **Task 9- Quantifying the momentum transfer between leg and water**

Task 9-a- Visualization of surface waves and body water movements under moving legs: During their propulsion, water striders are able to reach high horizontal velocities, but the movement has also a non-negligible vertical component (**Figure 8**). We want to study the dynamics of the relative importance of the surface waves versus vortices, as we know that the momentum initially imparted on the waves during the animal's stroke should be afterwards transferred to the fluid during wave relaxation. The presence and extent of completeness of a fan (plumy, bushy, reduced, flexible, etc.) will indeed change the forces imparted on the fluid. We will use horizontal and vertical PIV measurements, a well-mastered technique in the **Casas** lab, see (Steinmann and Casas, 2014), the Schlieren technique and the adapted Tomo-PIV (both of which also well mastered, (Steinmann et al., 2018; Steinmann et al., 2021b)) to quantify precisely the water surface deformation and the vortices generated underneath.

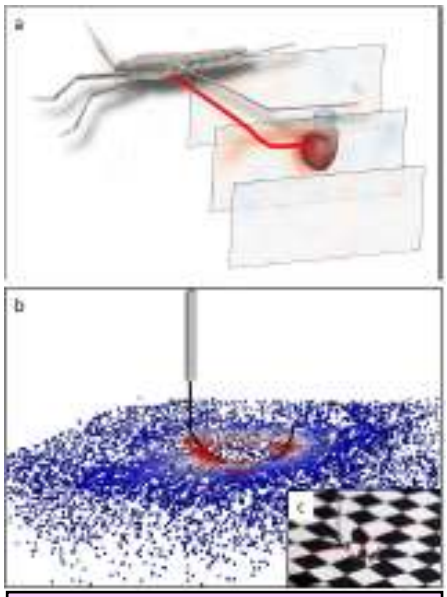

Figure 8. Visualisation by Tomographic velocimetry and three-dimensional tracking of the vortex created in the water bulk by a fan-less *Gerris* **(a)** and by a model leg moved with identical kinematics on a platform **(b)**.

Task 9-b- Experimental estimation of leakiness and drag of artificial fans: Leakiness is the amount of fluid going through a structure rather than around it. We will run experiments, in water and oil (in order to match the Reynolds numbers observed in nature), with artificial 3D printed fans of varying porosity and extract these two parameters. This task will be done with the fans of both *Rhagovelia* and *Tetraripis*. We expect a structural optimality in terms of "porosity" of the structure enabling the displacement of large amounts of water with a minimum of material. The **Casas** lab already demonstrated expertise in this field (Jaffar-Bandjee et al., 2020; Steinmann et al., 2021b).

Task 9-c- Numerical simulations of the hydrodynamics of propulsion and estimation of the forces at play: We plan to compute the interfacial flow, resulting from the leg movement, using simulations of finite elements. We will simulate the coupled equations describing the interface position and the flow field. The air-water interface will be modeled as a diffuse interface, which is advected by the flow field. We will first compute the interfacial flow during movement of *Oiovelia* (fan-less). These results will be then compared to the simulations of the flow around of *Rhagovelia* median propulsive legs. The numerical simulations on *Rhagovelia* will be used to estimate the local forces acting on the moving legs and to reproduce the complex interactions between leg-cum-dimple geometry and vorticity generation in each case, we will determine pressure and viscous drag acting on the fan structure and compare them to the forces acting on *Rhagovelia*. The quantification of the different sources of drag forces will have a major importance in our comprehension of the function of the fan. This framework has already been successfully adapted to the study of the propulsion of the water strider *Gerris paludum (*Steinmann et al 2021 JFM).

- **Task 10- Ecomechanics:** here we will perform a comparative estimation of locomotory performance according to fan morphology. We will estimate the role of the different leg architectures to the locomotory performance by identifying the relationship between structural features of the fan (porosity or leakiness, number of branches and density of hairs, plume-like or fractal-like branches) and the efficiency of propulsion *i.e.,* maximizing the momentum exchange with the fluid for a minimal loss of energy imparted to the fluid and to the waves. The effect of ontogeny on the water deformation and

propulsion is very strong. We will therefore make identical comparisons over the ontogeny of these water striders.

What do we learn from aim-3? A detailed understanding of the fluid dynamic characteristics underlying the interaction between the fan and the water surface, and therefore how the fan contributes to locomotion performance.

Who performs the aim? The aim will be led by partner 3 **(Casas)** who will hire a **Postdoctoral fellow** (1 year) with expertise in experimental fluid dynamics. The postdoc can collaborate with the PhD student and lab manager from the lab of the coordinator (**Khila**) to generate animals with modified fans (e. g. through *geisha* or *yellow* RNAi (which reduced fan stiffness) or through surgical removal of the fan). These animals will be transported to Tours where the biomechanics experiments will be conducted in the lab of partner 3.

Innovative nature of the project

A major difficulty in understanding the cellular and genetic mechanisms underlying the emergence of evolutionary innovations is to obtain a systematic, comprehensive and unbiased characterization of all cell types and transcriptional states in the system under study. Furthermore, studies that link the mechanistic aspects of evolutionary innovations (developmental genetics and cell biology) to the ultimate ecological forces driving their evolution are difficult to achieve. This project takes advantage of a highly tractable innovation, the propelling fan, to identify the cell types and states underlying the emergence of the fan from its ancestral structure: the arolium. Furthermore, the project is poised to uncover the differences in gene networks and gene function underlying the change from arolium to fan, and the mechanisms associated with the evolution of various fan shapes and number of pairs. Finally, results from this project will provide a comprehensive understanding of the functional contribution of the fan to the lineages where it evolved through uncovering the relationship between the morphological properties of the fan and it function as a swimming organ. These advances are only possible thanks to the ease of manipulating the model, state of the art technology such as single cell sequencing and manipulation of gene function, and the combined expertise of the consortium covering developmental genetics and evolution, systems biology, modeling and machine learning, and a critical contribution from applied physics and biomechanics. Therefore, this project will greatly advance our understanding of the cellular, molecular, evolutionary origin of this innovation, as well as its impact on the biology of the organism.

Scientific and technical barriers to be lifted

Model organisms: The coordinator (**Khila**) has established water striders, including the focal species, as models for evolutionary developmental genetics (Crumiere and Khila, 2019; Khila et al., 2014; Santos et al., 2015; Santos et al., 2017; Vargas-Lowman et al., 2019), and we do not anticipate any barriers.

ScRNA-seq: This experiment has never been done in water striders and we are aware of the difficulties that might arise. However, both the coordinator and partner 2 are within a scientific environment with strong expertise in scRNA-seq and modeling. Within IGFL (**partner 1**), we will receive continuous advice from **Michalis Averof, Jonathan Enriquez** and **Yad Ghavi-Helm**, all of whom performed scRNA-seq on crustaceans or flies. The project will benefit from the expertise and equipment of the recent ANR funded *Spatial-Cell-ID Equipex+* initiative, which was granted to ENS-Lyon to set a national facility of spatial transcriptomics. Spatial-Cell-ID aims to determine the transcriptome of each cell in tissues, either through data acquisition *in situ* at high resolution, or through algorithmic repositioning of the cell into the tissue. Within LBMC (**partner 2**), we will receive support from an already established systems biology and modeling network where some of us, **L. Modolo**, are already active (see https://anr-singlestatomics.pages.math.cnrs.fr/#people). Finally, **Benjamin Gillet** and **Sandrine Hughes**, the co-heads of the IGFL sequencing facility, have established scRNA-seq protocols for various organisms including flies and the crustacean Parhyale. The facility will perform the optimization of cell dissociation, sorting, library construction and sequencing of a pilot sample. This will allow us to establish the protocol for water striders and chose the optimal cell number and sequencing depth. Budget for these initial pilot tests is included in our global budget.

Backup plan: In the highly unlikely event that scRNA-seq proves difficult to perform, the project can still rely on classical comparative transcriptomics of the tarsi from the three species listed. This method is routine in the lab of the coordinator and was used for *Rhagovelia* here (Santos et al. Science 2017).

Gene network inference: The success of the project does not depend on having a reconstructed gene network of the fan, and will make ground breaking discoveries by simply identifying the set of genes responsible for fan development and evolution. However, we will push the analyses as far as possible and attempt to reach at least a partial network, as protocols and pipelines are available and **Partner 2** is familiar with them.

d. Ability of the project to address the research issues covered by the chosen research theme

The project focuses on developmental biology, cell biology, and genetics with an important added value of biomechanics and applied physics. These topics are at the heart of the CES 13 – Biologie cellulaire, biologie du développement et de l'évolution.

II. Organisation and implementation of the project

The proposal combines expertise from developmental genetics, ecology, cell biology, machine learning and systems biology, and physics of fluids that none of the teams possesses independently. This will allow a comprehensive understanding of how evolutionary innovations **develop**, **function**, and **evolve**. The coordinator (**Khila**) with partner 2 (**Francesconi**) will focus on how various cell types and gene networks underlying evolutionary innovations emerge and operate. Partner 3 (**Casas**) will focus on the physics of fan function and how it contributes to the biology of the animal. These topics represent a challenge in the study of species adaptation and diversification and are highly difficult to combine in a single project. The consortium spans a unique spectrum of questions and scales, from molecules to cells to whole organismal function. This is rarely met in other examples of evolutionary innovations, and therefore offers a unique opportunity for impactful contributions to our understanding of how these traits emerge and evolve (**Figure 9**).

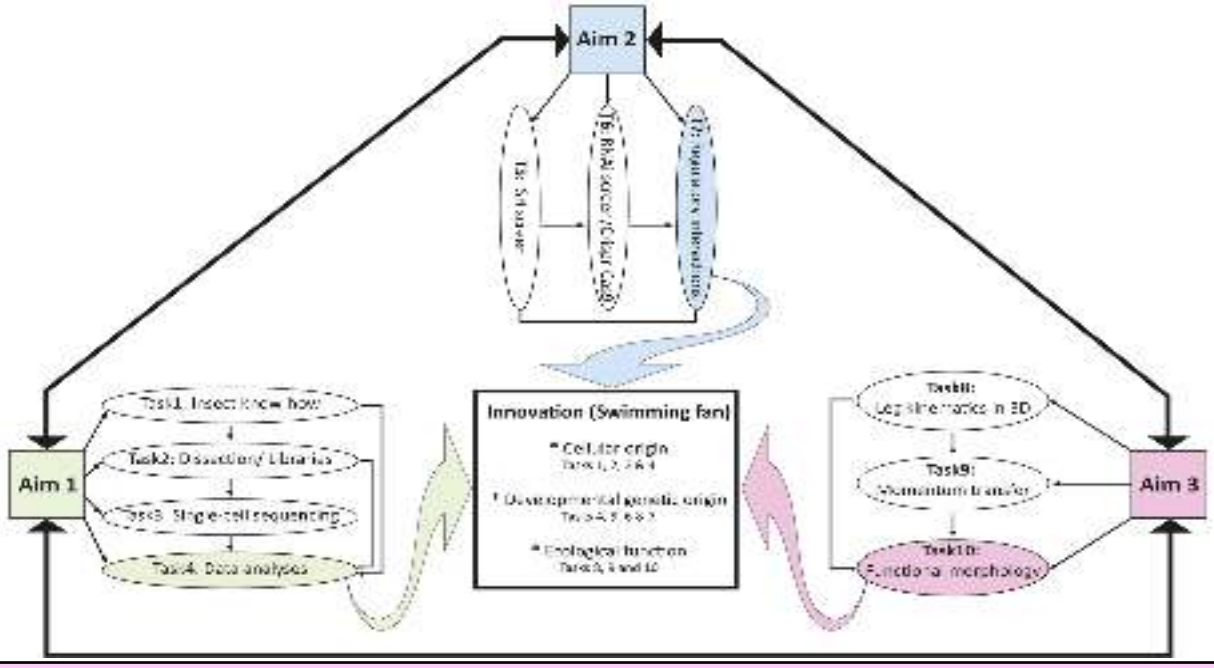

Figure 9: Three complementary aims, each led by an expert partner with contribution from the other partners. **Francesconi** leads Aim1, which requires expertise in modeling, systems biology and scRNA-seq. He will have support from **Khila** who has the insect know-how. **Khila** leads Aim2, which require EvoDevo, cell biology and developmental genetics. **Casas** leads Aim3, which requires applied physics and biomechanics expertise. Combined, the three aims cover the molecular, developmental, evolutionary, and functional aspect of the fan as an innovation.

a. Scientific coordinator and its consortium / its team

This is a collaborative project built around the complementary expertise of three teams; two are located at the Lyon Gerland campus (**Khila** and **Francesconi**) and the third in Tours (**Casas**). Our consortium covers developmental genetics, evolution, behaviour, fieldwork, and a deep knowledge of the model system (**Khila**), combined with single cell sequencing, modeling and systems biology (**Francesconi**), and physics and biomechanics of locomotion, from flying to walking to skimming the water surface (**Casas**). **Francesconi** will lead aim1 **(Figure 9**), which requires expertise in systems biology and machine learning. The **Khila** lab will contribute to this aim by providing the know-how of the model that is necessary to conduct the insect dissections and manipulation. **Khila** will lead aim2 (**Figure 9**), based on the results from aim1, which requires expertise in developmental genetics and gene expression and function analyses. **Casas** will lead aim3 (**Figure 9**), which requires expertise in applied physics and biomechanics. **Casas** will receive support from **Khila** to generate animals with genetically or surgically manipulated fans. Our consortium will also receive an invaluable expertise in scRNA-seq and next generation sequencing and bioinformatics from the IGFL sequencing facility (**PSI**) headed by **Sandrine Hughes** and **Benjamin Gillet.** The facility has expertise in optimizing single cell sorting, library preparation and sequencing using 10X technology. This is critical for the efficient implementation of the various packages and the supervision of students and postdoctoral fellows.

*** Partner 1, the coordinator: Abderrahman Khila** is *Directeur de Recherche* at CNRS and former ATIP-Avenir and ERC investigator, heading the *Laboratory for Developmental Genomics and Evolution at* IGFL, ENS de Lyon. He specializes in the study of the developmental genetic mechanisms underlying the evolution of adaptive traits in water striders, and has been developing state of the art tools, sequence resources and protocols for breeding water striders including the focal species (Santos et al., 2015). He will dedicate 40% of his time to the project and be in charge of coordinating the progress and ensuring regular meetings among the members of the consortium. He has experience with non-model insects in general, regularly conducts fieldwork in the Amazonas, which hold impressive insect diversity. **Khila** is a former Special Visiting Scientist at Oswaldo Cruz Institute in Rio de Janeiro, Brazil, and currently holds a Guest Professor status at Uppsala University, Sweden. He holds an Academic Editor role at '*PLoS Biology*' and an Associate Editor role at '*Evolution Letters'* and '*EvoDevo*' journals.

**** Partner 2:*** **Mirko Francesconi,** tenured researcher at Inserm (CRCN), leads the recently established **Quantitative Regulatory Genomics** team at **LBMC**, ENS Lyon. During his research career, he has always worked at the interface between experimental and computational biology. He has led several studies which made **important contributions** in the dissection of **both genetic** (Francesconi et al., 2011; Francesconi and Lehner, 2014) and **non-genetic** (Francesconi et al., 2019; Perez et al., 2017) **sources of phenotypic variation**, using both genome-wide **computational and experimental systems biology approaches** in model organisms. He has extensive experience in data integration and modelling and especially focuses on the analysis of the **dynamics of gene expression** both to extract hidden phenotypic information and to formulate hypotheses about the molecular mechanisms involved in the generation of phenotypic variation (Francesconi et al., 2019; Francesconi and Lehner, 2014, 2015; Perez et al., 2017). He also recently became interested in the origin and the consequences of cell-to-cell heterogeneity on cell plasticity and cell fate determination during development, differentiation and reprogramming, using genome-wide single-cell transcriptomic data (Francesconi et al., 2019).

**** Partner 3:*** **Jérôme Casas,** Professor "classe exceptionelle" and former director of IRBI from 2001 to 2008. IRBI, a joint Université de Tours-CNRS research unit on insect biology, is the largest such unit in France and one of the largest in Europe. **Casas** is an IUF senior member and was an IUF junior member earlier, a rare occurrence. He did hold a visiting chair position at Tucson (USA) and ESPCI (Paris) and is currently holding a visiting chair in bioinspired microtechnology at the LETI/CEA (Grenoble). His work specializes in physical ecology of insects; the physiology, behaviour and population dynamics of consumer-resource interactions; the sensory ecology of mimicry; and biologically- inspired technology, particularly biomimetic flow sensing. One notable feature of his approach is the blending of natural history with both state-of-the-art technology and modeling. His group is composed of applied mathematicians, engineers and biologists (Krijnen et al., 2019; Steinmann et al., 2018; Steinmann and Casas, 2014; Steinmann et al., 2019; Voise and Casas, 2010).

Implication of the scientific coordinator and partner's scientific leader in on-going project(s)

Name of the researcher	Person.month	Call, funding agency, grant allocated	Project's acronym	Name of the scientific coordinator	Start – End
A Khila	24	ANR	SocioDev	A Khila	2017-2021
J. Casas	3	Région Centre	PHEROAERO	J. Casas	2019-2021
M. Francesconi	36	ANR	InterPhero	M. Francesconi	2019-2024

b. Implemented and requested resources to reach the objectives

	Month 1-6	Month 7-12	Month 13-18	Month 19-24	Month 25-30	Month 31-36	Month 37-42	Month 43-48
Task1: Insect husbandry	P1	P1						
Task 2: Optimizing cell dissociation	P2+PSI	P2+PSI						
Task3: scRNA-seq	P2+PSI	PSI						
Task4: Data analyses		P2	P2					
Task5: In situ hybridization screen			P1+P2	P1+P2	P1+P2			
Task6: RNAi screen and Crispr/Cas9			P1+P2	P1+P2	P1+P2	P1+P2	P1+P2	
Task7: Tests of genetic interactions			P1+P2	P1+P2	P1+P2	P1+P2	P1+P2	
Task8: Leg kinematics in 3D	P3	P3	P3					
Task9: Tests of momentum transfer		P3	P3	P3				
Task10: Fan functional morphology			P3+P1	P3+P1	P3+P1			
Article drafting and publications			P1+P2+ P3	P1+P2+ P3	P1+P2+ P3	P1+P2+ P3	P1+P2+ P3	P1+P2+ P3

Gantt chart describing the timeline for the execution of each task throughout the 48 months of the project. **P**: Partner. **PSI**: IGFL sequencing facility. The leading partner for each task is written first.

Scientific and technical justification of the funds requested for each partner

The consortium is requesting **578224.64€** to execute the project. The cost is distributed between the coordinator (304 012.80€), Partner 2 (137 132.80€) and Partner 3 (137 079.04€) as follows:

Partner 1: Abderrahman Khila (IGFL)

- **Staff expenses: 172 440**

This includes a PhD salary for 36 months (117 000€) and a **Bioinformatics Postdoctoral fellow** for 12 months (55 440€). The PhD student will formally be co-supervised by the coordinator (**Khila**) and Partner 2 (**Francesconi**) who are affiliated with two neighbouring institutes in the same campus. The postdoctoral fellow will also be co-supervised by **Khila** and **Francesconi**. This fellow will be hired for two years in total and the salary of the second year is part of the budget of partner 2.

- **Instruments and material costs: Computers 4 000**€.
- **Outsourcing / subcontracting:**

Sequencing 35 000: These are costs for 16 single-cell-sequencing samples covering 8 comparative samples from *Tetraripis* and *Oiovelia*, and 8 samples from mutants/RNAi-treated embryos. This includes costs for library preparation (1500€/library) and Illumina sequencing. The cost also takes into account any optimazation steps. The **remaining 16 samples** are budgeted by Partner 2 below.

- **Publications 10 000**€: These are both publication and page charges as well as costs for open access. We favour the option of open access to give a broad exposure to our work and allow all communities to benefit from our findings.
- **General and administrative costs & other operating expenses:**
- **Travel: 10,000€** These are costs of sampling expeditions, participation to scientific meetings, and movement of staff between the labs of the partners and collaborators.
- **Consumables: 40 000**€ (10 000/year). These include daily expenses for lab-related reagents (Kits, enzymes, chemicals, antibodies, etc…), antibodies, primers and reagents for genome editing and RNAi, needles, bioinformatics equipment (access to server and computing facilities etc…).
- **Overhead fees** (12% of direct costs): 32 572,80€

Partner 2: Mirko Francesconi (LBMC)

- **Staff expenses: 55 440€**

These is a second year of post doc salary for the fellow hired for two years and co-supervised by Francesconi and Khila.

- **Instruments and material costs: 12 000€.**

This cost will cover a computer server for handling and backing-up all high-throughput datasets as well as for giving remote and centralized access to the data and to analysis software for team-members and collaborators.

- **Outsourcing / subcontracting: 35 000€.**

These are costs for 16 single-cell-sequencing samples covering 8 developmental stages and two legs of *Rhagovelia antilleana.* The cost takes into account scRNA-seq library preparation, sequencing and steps of optimization.

- **General and administrative costs & other operating expenses: 10 000€**

Expenses for participation to scientific meetings, and movement of staff between the labs of the partners.

- **Other: 10 000€**

These are both publication and page charges as well as costs for open access. We favour the option of open access to give a broad exposure to our work and allow all communities to benefit from our findings.

- **Overhead fees (12% of direct costs): 14 692,80€**

Partner 3: Jerome Casas (IRBI)

- **Staff expenses: 73 392 €** for an advanced post-doc for one year.
- **Instruments and material costs: 15 000 €** for a platform for moving insect legs in 3D with high precision and speed (model Zaber with adds-on for high speed), force sensors and a platform for measuring wetness.
- **Consumables: 20 000** € for seeding particles for flow visualization, tanks, small mechanical parts for the flow tanks, Micro-and nano-CT scans done at Brücker headquarters in Belgium, with the SKYSCAN 2214 for the Nano-CT scans and the Skyscan 1272 for the micro-CT scans and 3D printing material and access to printers with high definition, such as the Nanoscribe.
- **Outsourcing / subcontracting: 4000€**

These are both publication and page charges as well as costs for open access. We favour the option of open access to give a broad exposure to our work and allow all communities to benefit from our findings.

- **General and administrative costs & other operating expenses: 10 000€**

Expenses for participation to scientific meetings, and movement of staff between the labs of the partners.

- **Overhead fees (12% of direct costs): 14 687,04€**

Requested means by item of expenditure and by partner*

		Partner 1 *IGFL*	**Partner 2** *LBMC*	**Partner 3** *IRBI*
Staff expenses		172 440 €	55 440 €	73 392 €
Instruments and material costs (including the scientific consumables)		44 000 €	22 000 €	35 000 €
Building and ground costs		0 €	0 €	0 €
Outsourcing / subcontracting		35 000 €	35 000 €	4 000 €
General and administrative costs & other operating expenses	Travel costs	10 000 €	10 000 €	10 000 €
	Administrative management & structure costs**	32 572,80 €	14 692,8 €	14 687,04 €
Sub-total		304 012,80 €	**137 132,80 €**	**137 079,04 €**
Requested funding		**578 224,64 €**		

III. Impact and benefits of the project

Scientific outcome: Our research consortium gathers scientists with a history of publishing in top-level scientific journals, including generalist journals (Science, Nature, PNAS, PLoS Biology, Current Biology, Nature Communication) and journals specialised in Fluid mechanics, Development, Evolution, Cell Biology, Genomics, and Organismal biology. We will thus strive to publish the results of our research in top-level scientific journals to increase the exposure and the impact of the French research and strengthen the ties with our international colleagues. In addition, we will present our work at international scientific conferences, such as SMBE, Society for Cell and Developmental Biology, Evolution, ESEB, Euro EvoDevo and EvoDevo-PanAm, Evolutionary Systems Biology and from Functional Genomics to Systems Biology. We will also use social media, preprint servers such as Arxiv or BioRxiv, and opt for open-access options whenever possible to increase the visibility of our research for the scientific community worldwide.

The design of semi-aquatic microrobots is a vibrant field in the micro-robotic community, see for example (Koh et al., 2015), and our understanding of the added value of a fan for improved locomotion will be of high interest to this community. We intend to patent some of the ideas if of immediate interest.

Strategic outcome: A number of leading labs in the study of Development and Evolution have attempted to address questions related to the origin of evolutionary innovations. These questions are known to be hard to pursue due to the difficulty of associating complementary expertise, from cell biology and developmental genetics to applied physics and fluid dynamics, with a tractable study system where experiments revealing the origin of these traits are technically feasible. We have devoted a significant effort to successfully establish this model, which was enthusiastically welcomed by the scientific community (Initially started as minor part of an ERC grant and led to a publication in ***Science*** in 2017). This project will give us the opportunity to share this knowledge with the scientific community while establishing French research as a pioneer in this field. It will also increase significantly the impact our proposal will have on the future of the field of studying evolutionary innovations in general.

Training of highly qualified personnel: During this project, we will train scientists in Molecular, Cellular, Developmental and Evolutionary biology, ecology, behaviour, single cell sequencing and bioinformatics, as well as biomechanics and fluid dynamics. This multi-disciplinarity will help train highly qualified personnel and open-minded future scientists who can advance the field. We will thus train 1 PhD student and two Postdoctoral fellows, Masters and undergraduate students (usually coming from ENS Lyon, INSA Lyon, or Université Lyon 1 and Université de Tours), and highly qualified technicians.

Teaching and public outreach: The research outlined in this proposal will also be used in teaching undergrad and postgraduate students (ENS Lyon and Université Lyon1, University of Tours) and in workshops and summer schools across Europe and elsewhere. Several researchers of our consortium are actively involved in teaching at their respective universities and all partners have been actively involved in international teaching at various summer schools and workshops. These teaching activities will provide ample opportunities to disseminate the results of our work but also the methodologies involved, in particular during already-in-place practical courses in cell biology development, evolution and systems biology.

Beyond university students, our project is a rare example where research can reach a wide audience, because evolutionary innovations provide such a compelling illustration of how organisms can achieve impressive ecological success upon the acquisition of new ecological opportunities. In addition, our research often attracts media attention and has been in display through articles in national and international outlets such as ‘Le Monde’, ‘Quanta magazine’, ‘BBC’ and through generalist articles such as Nature news and views (Gardiner, 2017).

Evolutionary innovations represent powerful illustrations of evolutionary theory, and thus provide an excellent opportunity for outreach based on our research for two reasons. First, the link between development and the ecological function of the trait offers a rare opportunity to explain abstract concepts of evolutionary theory, an important subject to the society. Second, as these traits are hard to explain by descent (a subject of controversy used by proponents of intelligent design to attack the evolutionary theory), our work will appeal to the public and is likely to help raise awareness about fundamental science and the role of insects in teaching us about how life evolves. For this reason, we believe we have a unique opportunity to construct outreach activities revolving around the idea of evolutionary innovations.

IV. References related to the project (**consortium references in bold**)

Aalto, A., Viitasaari, L., Ilmonen, P., Mombaerts, L., Goncalves, J., 2020. Gene regulatory network inference from sparsely sampled noisy data. Nat Commun 11, 3493.

Alles, J., Karaiskos, N., Praktiknjo, S.D., Grosswendt, S., Wahle, P., Ruffault, P.L., Ayoub, S., Schreyer, L., Boltengagen, A., Birchmeier, C., Zinzen, R., Kocks, C., Rajewsky, N., 2017. Cell fixation and preservation for droplet-based single-cell transcriptomics. BMC Biol 15, 44.

Andersen, N.M., 1982. The semiaquatic bugs (Hemiptera: Gerromorpha). Entomonograph 3.

Andersen, N.M., 2000. A new species of Tetraripis from Thailand. TIJDSCHRIFT VOOR ENTOMOLOGIE 142, 185–194.

Armisen, D., Refki, P.N., Crumiere, A.J., Viala, S., Toubiana, W., Khila, A., 2015. Predator strike shapes antipredator phenotype through new genetic interactions in water striders. Nat Commun 6, 8153.

Aubin-Frankowski, P.C., Vert, J.P., 2018. Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. BioRxiV.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., Newell, E.W., 2018. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.

Bergen, V., Lange, M., Peidli, S., Wolf, F.A., Theis, F.J., 2019. Generalizing RNA velocity to transient cell states through dynamical modeling. BioRxiv.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., Satija, R., 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411-420.

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., Adey, A., Waterston, R.H., Trapnell, C., Shendure, J., 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357, 661-667.

Chan, T.E., Stumpf, M.P.H., Babtie, A.C., 2017. Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures. Cell Syst 5, 251-267 e253.

Chaverra-Rodriguez, D., Macias, V.M., Hughes, G.L., Pujhari, S., Suzuki, Y., Peterson, D.R., Kim, D., McKeand, S., Rasgon, J.L., 2018. Targeted delivery of CRISPR-Cas9 ribonucleoprotein into arthropod ovaries for heritable germline gene editing. Nat Commun 9, 3008.

Coifman, R.R., Lafon, S., Lee, A.B., Maggioni, M., Nadler, B., Warner, F., Zucker, S.W., 2005. Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc Natl Acad Sci U S A 102, 7426-7431.

Crumiere, A.J.J., Khila, A., 2019. Hox genes mediate the escalation of sexually antagonistic traits in water striders. Biol Lett 15, 20180720.

Francesconi, M., Di Stefano, B., Berenguer, C., de Andres-Aguayo, L., Plana-Carmona, M., Mendez-Lago, M., Guillaumet-Adkins, A., Rodriguez-Esteban, G., Gut, M., Gut, I.G., Heyn, H., Lehner, B., Graf, T., 2019. Single cell RNA-seq identifies the origins of heterogeneity in efficient cell transdifferentiation and reprogramming. Elife 8.

Francesconi, M., Jelier, R., Lehner, B., 2011. Integrated genome-scale prediction of detrimental mutations in transcription networks. PLoS Genet 7, e1002077.

Francesconi, M., Lehner, B., 2014. The effects of genetic variation on gene expression dynamics during development. Nature 505, 208-211.

Francesconi, M., Lehner, B., 2015. Reconstructing and analysing cellular states, space and time from gene expression profiles of many cells and single cells. Mol Biosyst 11, 2690-2698.

Frantsevich, L., Ji, A.H., Dai, Z.D., Wang, J.T., Frantsevich, L., Gorb, S.N., 2008. Adhesive properties of the arolium of a lantern-fly, Lycorma delicatula (Auchenorrhyncha, Fulgoridae). J Insect Physiol 54, 818-827.

Gao, X.F., Jiang, L., 2004. Water-repellent legs of water striders. Nature 432, 36-36.

Gardiner, J.R., 2017. EVOLUTION. Fan-assisted insects. Nature 550, 464-464.

Haghverdi, L., Buttner, M., Wolf, F.A., Buettner, F., Theis, F.J., 2016. Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods 13, 845-848.

Huynh-Thu, V.A., Geurts, P., 2018. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. Sci Rep 8, 3384.

Jaffar-Bandjee, M., Casas, J., Krijnen, G., 2018. Additive manufacturing: state of the art and potential for insect science. Curr Opin Insect Sci 30, 79-85.

Jaffar-Bandjee, M., Steinmann, T., Krijnen, G., Casas, J., 2020. Leakiness and flow capture ratio of insect pectinate antennae. Journal of the Royal Society Interface 17.

Khila, A., Abouheif, E., Rowe, L., 2009. Evolution of a novel appendage ground plan in water striders is driven by changes in the Hox gene Ultrabithorax. PLoS genetics 5, e1000583-e1000583.

Khila, A., Abouheif, E., Rowe, L., 2012. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. Science 336, 585-589.

Khila, A., Abouheif, E., Rowe, L., 2014. Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. Evolution 68, 2159-2170.

Kim, J., S, T.J., Natarajan, K.N., Won, K.J., 2021. TENET: gene network reconstruction using transfer entropy reveals key regulatory factors from single cell transcriptomic data. Nucleic Acids Res 49, e1.

Koh, J.S., Yang, E., Jung, G.P., Jung, S.P., Son, J.H., Lee, S.I., Jablonski, P.G., Wood, R.J., Kim, H.Y., Cho, K.J., 2015. BIOMECHANICS. Jumping on water: Surface tension-dominated jumping of water striders and robotic insects. Science 349, 517-521.

Krijnen, G.J., Steinmann, T., Jaganatharaja, R.K., Casas, J., 2019. Insect-inspired distributed flow-sensing: Fluid-mediated coupling between sensors, Architectured Materials in Nature and Engineering Springer, Cham, pp. 355-392.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A., Fan, J., Borm, L.E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundstrom, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., Kharchenko, P.V., 2018. RNA velocity of single cells. Nature 560, 494-498.

Liem, K.F., 1973. Evolutionary Strategies and Morphological Innovations: Cichlid Pharyngeal Jaws. Systematic Biology 22 425–441.

Losos, J.B., 2009. Lizards in an evolutionary tree : ecology and adaptive radiation of anoles. . Organisms and environments, 507.

Marbach, D., Costello, J.C., Kuffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., Consortium, D., Kellis, M., Collins, J.J., Stolovitzky, G., 2012. Wisdom of crowds for robust gene network inference. Nat Methods 9, 796-804.

Marbach, D., Prill, R.J., Schaffter, T., Mattiussi, C., Floreano, D., Stolovitzky, G., 2010. Revealing strengths and weaknesses of methods for gene network inference. Proc Natl Acad Sci U S A 107, 6286-6291.

Matsumoto, H., Kiryu, H., Furusawa, C., Ko, M.S.H., Ko, S.B.H., Gouda, N., Hayashi, T., Nikaido, I., 2017. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. Bioinformatics 33, 2314-2321.

Melsted, P., Booeshaghi, A.S., Liu, L., Gao, F., Lu, L., Min, K.H.J., da Veiga Beltrame, E., Hjorleifsson, K.E., Gehring, J., Pachter, L., 2021. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat Biotechnol.

Melsted, P.e.a., 2019. Melsted, P. et al. Modular and efficient pre-processing of single-cell RNA-seq. . BioRxiv 673285.

Moisy, F., Rabaud, M., Salsac, K., 2009. A synthetic Schlieren method for the measurement of the topography of a liquid interface. Exp Fluids 46, 1021-1036.

Packer, J.S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R.H., Murray, J.I., 2019. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science 365.

Peng, J., Serrano, G., Traniello, I.M., Calleja-Cervantes, M.E., Chembazhi, U.V., Bangru, S., Ezponda, T., Rodriguez-Madoz, J.R., Kalsotra, A., Prosper, F., Ochoa, I., Hernaez, M., 2021. A single-cell gene regulatory network inference method for identifying complex regulatory dynamics across cell phenotypes. BioRxiv.

Perez, M.F., Francesconi, M., Hidalgo-Carcedo, C., Lehner, B., 2017. Maternal age generates phenotypic variation in Caenorhabditis elegans. Nature 552, 106-109.

Plass, M., Solana, J., Wolf, F.A., Ayoub, S., Misios, A., Glazar, P., Obermayer, B., Theis, F.J., Kocks, C., Rajewsky, N., 2018. Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. Science 360.

Polhemus, J.T., Polhemus, D.A., 2008. Global diversity of true bugs (Heteroptera : Insecta) in freshwater. Hydrobiologia 595, 379-391.

Qiu, X., Rahimzamani, A., Wang, L., Ren, B., Mao, Q., Durham, T., McFaline-Figueroa, J.L., Saunders, L., Trapnell, C., Kannan, S., 2020. Inferring Causal Gene Regulatory Networks from Coupled Single-Cell Expression Dynamics Using Scribe. Cell Syst 10, 265-274 e211.

Ronco, F., Matschiner, M., Bohne, A., Boila, A., Buscher, H.H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., Jentoft, S., Salzburger, W., 2021. Drivers and dynamics of a massive adaptive radiation in cichlid fishes. Nature 589, 76-81.

Roth, Z., Weil, S., Aflalo, E.D., Manor, R., Sagi, A., Khalaila, I., 2013. Identification of receptor-interacting regions of vitellogenin within evolutionarily conserved beta-sheet structures by using a peptide array. Chembiochem 14, 1116-1122.

Santos, M.E., Berger, C.S., Refki, P.N., Khila, A., 2015. Integrating evo-devo with ecology for a better understanding of phenotypic evolution. Brief Funct Genomics 14, 384-395.

Santos, M.E., Le Bouquin, A., Crumière, A.J.J., Khila, A., 2017. Taxon-restricted genes at the origin of a novel trait allowing access to a new environment. Science 358, 386-390.

Schluter, D., 2000. The ecology of adaptive radiation., 288 p.

Shi, F., Niu, J., Liu, J.L., Liu, F., Wang, Z.Q., Feng, X.Q., Zhang, X., 2007. Towards understanding why a superhydrophobic coating is needed by water striders. Adv Mater 19, 2257-+.

Steinmann, T., Arutkin, M., Cochard, P., Raphaël, E., Casas, J., Benzaquen, M., 2018. Unsteady wave pattern generation by water striders. Journal of Fluid Mechanics, 370-387.

Steinmann, T., Casas, J., 2014. Laser-Based Optical Methods for the Sensory Ecology of Flow Sensing: From Classical PIV to Micro-PIV and Beyond, In Flow sensing in air and water. H. Bleckmann, J. Mogdans & S.L. Coombes. Springer Verlag., pp. 31-62.

Steinmann, T., Casas, J., Braud, P., David, L., 2019. Tomo-PTV measurement of a drop impact at air-water interface. , Proceedings of the 13th International Symposium on Particle Image Velocimetry https://athene-forschung.unibw.de/129163?show_id=128856.

Steinmann, T., Casas, J.M., Braud, P., David, L., 2021a. Coupled measurements of interface topography and three-dimensional velocity field of a free surface flow. Exp Fluids 62.

Steinmann, T., Cribellier, A., Casas, A., 2021b. Singularity of the water strider propulsion mechanisms J. Fluid Mech. 915.

Toubiana, W., Armisen, D., Viala, S., Decaras, A., Khila, A., 2021. The growth factor BMP11 is required for the development and evolution of a male exaggerated weapon and its associated fighting behavior in a water strider. PloS Biology, In Press.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., Rinn, J.L., 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32, 381-386.

Vargas-Lowman, A., Armisen, D., Burguez Floriano, C.F., da Rocha Silva Cordeiro, I., Viala, S., Bouchet, M., Bernard, M., Le Bouquin, A., Santos, M.E., Berlioz-Barbier, A., Salvador, A., Figueiredo Moreira, F.F., Bonneton, F., Khila, A., 2019. Cooption of the pteridine biosynthesis pathway underlies the diversification of embryonic colors in water striders. Proc Natl Acad Sci U S A 116, 19046-19054.

Voise, J., Casas, J., 2010. The management of fluid and wave resistances by whirligig beetles. J R Soc Interface 7, 343-352.

Wagner, A., 2011. The molecular origins of evolutionary innovations. Trends Genet 27, 397-410.

Wagner, A., 2017. Information theory, evolutionary innovations and evolvability. Philos Trans R Soc Lond B Biol Sci 372.

Wagner, G.P., Lynch, V.J., 2010. Evolutionary novelties. Curr Biol 20, R48-52.

Wildeman, S., 2018. Real-time quantitative Schlieren imaging by fast Fourier demodulation of a checkered backdrop. Exp Fluids 59.