Evolution and characterization of Drosophila glue, a model for biomimicry

Manon Monier

Université Paris Cité

Université Paris Cité

Ecole Doctorale Bio Sorbonne Paris Cité (ED 562)
Institut Jacques Monod (UMR_7592) - Equipe Evolution et Génétique

Evolution and characterization of Drosophila glue, a model for biomimicry

Par Manon MONIER

Thèse de doctorat de Génétique
Dirigée par Virginie COURTIER-ORGOGOZO
Et co-dirigée par François GRANER

Présentée et soutenue publiquement à Paris le 10 Octobre 2023
devant un jury composé de :

Patricia GIBERT, directrice de recherche, CNRS, Université Lyon 1, rapportrice
Patrick FLAMMANG, directeur de recherche, Université de Mons, rapporteur
Jean-Luc DA LAGE, chercheur, CNRS, Université Paris Saclay, examinateur
David LABONTE, directeur de recherche, Imperial College London, examinateur
Virginie COURTIER-ORGOGOZO, directrice de recherche, CNRS, Université Paris Cité, directrice de thèse
François GRANER, directeur de recherche, CNRS, Université Paris Cité, co-directeur de thèse

Titre : Evolution et caractérisation de la colle des Drosophiles, un modèle de biomimétisme

Résumé : Les bioadhésifs présentent des propriétés physico-chimiques qui permettent aux êtres vivants d'adhérer à une grande variété de substrats. Développés sur plusieurs millions d'années, ils sont une source d'inspiration pour développer de futurs matériaux plus respectueux de la santé et de l'environnement. C'est le cas de la colle des Drosophiles, une substance produite durant le stade larvaire qui permet à l'animal d'être attaché durant sa métamorphose au stade de pupe et de ne pas être consommé par des prédateurs. Cette colle adhère à une grande diversité de substrats (des feuilles, du bois, des fruits ou du plastique). Elle est produite par les glandes salivaires, sèche en quelques secondes, adhère pendant plusieurs jours dans des conditions d'humidité et température très variables. Jusqu'à présent, la composition et les propriétés adhésives de la colle ont été étudiées chez *Drosophila melanogaster* uniquement. Sa force d'adhésion est de 0.2 Newton pour une surface de 1 mm^2, ce qui correspond à la force de la plupart des rubans adhésifs utilisés dans le commerce. En outre, cette colle est biodégradable, biocompatible, réversible et pourrait posséder des propriétés répulsives. La colle de *D. melanogaster* est composée de huit protéines glycosylées, les Salivary Gland Secretion (Sgs), encodées par les gènes *Sgs*. L'objectif de ma thèse est d'analyser la colle des autres espèces de Drosophiles.

Dans une première partie, j'ai analysé les séquences génomiques de 24 espèces de Drosophiles afin de caractériser l'évolution moléculaire des gènes Sgs sur 30 millions d'années. J'ai annoté 102 gènes *Sgs* et identifié des copies de gènes *new glue* pas encore annotées. J'ai proposé une nouvelle nomenclature pour les gènes *Sgs*, basée sur la conservation des séquences protéiques, leur position au sein du génome et la présence/absence de répétitions internes. Nous avons observé deux dynamiques évolutives au sein des *Sgs*: un premier groupe (*Sgs1*, *Sgs3X* et *Sgs3e*) a connu peu d'évènements évolutifs (duplications, délétions, inversions) alors qu'un second groupe (*Sgs3b*, *Sgs7* et *Sgs8*) en a connu plusieurs. La dynamique du second groupe de gènes a pu être accélérée par la présence de petits gènes adjacents, les gènes *new glue*. Nos travaux montrent que certains gènes de la colle évoluent rapidement et pourraient jouer un rôle dans l'évolution rapide des propriétés de la colle.

Dans une seconde partie, j'ai effectué des tests d'adhésion sur 27 espèces de Drosophiles séparées par 150 millions d'années d'évolution et jusqu'à 6 souches par espèces. En comparant les propriétés adhésives, nous avons observé que la force d'adhésion varie indépendamment de la distance phylogénétique et évolue rapidement dans les deux directions aux cours de l'évolution. La force la plus élevée est observée chez *D. virilis* et elle excède celles des rubans adhésifs commerciaux. L'analyse de la quantité de colle produite par l'animal et son implication dans la force d'adhésion sont en cours. Enfin, nous avons identifié trois groupes d'espèces selon leur adhésion : faible, moyenne et forte.

Dans une troisième partie, nous avons étudié le rôle individuel des gènes *Sgs* dans

les propriétés d'adhésion de la colle de *D. melanogaster*, grâce à des lignées RNAi. Nos résultats préliminaires montrent une forte réduction de l'adhésion en l'absence de *Sgs3*.

Nous avons également publié un article de revue dans le Journal Insects qui présente la colle des Drosophiles comme un modèle prometteur de bioadhésion.

Mes travaux de thèse ont permis de fournir un socle de premiers résultats afin d'établir la colle des Drosophiles comme nouveau modèle pour étudier l'évolution d'un trait adaptatif et pour développer de futurs bioadhésifs.

Mots clefs : bioadhésif, Drosophile, glandes salivaires, Sgs, adhésion, biomimétisme

Title : Evolution and characterization of Drosophila glue, a model for biomimicry

Abstract : Bioadhesives display physico-chemical properties that enable living organisms to attach themselves to a great variety of substrates. They evolved for millions of years and are today a source of inspiration to develop new materials that are safe for human health and the environment. A promising material is Drosophila glue, a bioadhesive produced during the late larval stage that enables the animal to be attached during metamorphosis as a pupa. This glue can adhere to a large range of substrates (leaves, wood, fruits, glass, paper or plastic). It is produced by the salivary glands, it dries within a few seconds, and is adhesive for several days under various humidity and temperature conditions. So far, the precise composition and adhesive properties of the glue has only been studied extensively in the model species *Drosophila melanogaster*. Adhesive strength is approximately 0.2 Newton for 1 mm^2 of glue, which corresponds to the most adhesive tapes available. Moreover, this glue is biodegradable, biocompatible, reversible and could present repellent properties. *D. melanogaster* glue is made of eight glycosylated proteins, named Salivary Gland Secretion (Sgs) proteins, encoded by *Sgs* genes. The goal of my thesis was to examine this glue in other Drosophila species.

In the first part, I analyzed the genome sequences of 24 Drosophila species in order to characterize the molecular evolution of *Sgs* genes across 30 millions of years. I annotated a total of 102 *Sgs* genes and identified *new glue* gene copies not yet annotated. I proposed a new nomenclature for *Sgs* genes based on protein sequence conservation, genomic location and presence/absence of internal repeats. Overall, we observed two evolutionary dynamics among *Sgs* genes: a first group of glue genes (*Sgs1*, *Sgs3X* and *Sgs3e*) underwent relatively few evolutionary events (duplications, deletions, inversions) while genes from the second group (*Sgs3b*, *Sgs7* and *Sgs8*) went through many. The dynamics of the second group of genes may have been accelerated by the presence of adjacent, short *new glue* genes. Our work shows that some glue genes evolve rapidly and could play a role in the rapid evolution of glue properties.

In the second part, I used a pull-off force tests to measure pupa adhesion and compare it between 27 Drosophila species, with up to six strains per species, spanning 150 million years of evolution. We found that adhesion force varies independently of phylogenetic distance, showing that adhesion forces have evolved rapidly in both directions during evolution. The highest adhesion forces observed for *D. virilis* exceed the one of adhesive tapes. Analysis of the amount of secreted glue and whether this parameter affects adhesion strength is ongoing. Overall, we could distinguish three groups of species with a low, medium and strong adhesion force.

In a third part, we investigated the role of each *Sgs* gene in the adhesive properties of the glue of *D. melanogaster*, using RNAi lines. Our preliminary results indicate that the absence of *Sgs3* leads to a very low adhesion force.

Finally, we published a review article in the Insects journal presenting Drosophila glue as a promising model for bioadhesion.

My thesis work has established Drosophila glue as a model to study the evolution of an adaptive trait across species and to develop future bioadhesives.

Keywords : bioadhesive, Drosophila, salivary glands, Sgs, adhesion, biomimicry

Résumé substanciel

Être attaché à un substrat est d'importance primordiale pour les être vivants afin de maintenir leur position, se mouvoir, se nourrir, capturer et maintenir leur proie ou encore pour la copulation (Gorb, 2008). Les organismes ont développé différentes méthodes d'attachement, dont l'adhésion humide qui implique la production d'un bioadhésif, également appelé colle (Bianco-Peled and Davidovich-Pinhas, 2015). Développées au cours de millions d'années d'évolution, les colles ont des propriétés qui diffèrent en fonction des milieux auxquels ils sont adaptés. Leurs composés étant biocompatibles et biodégradables, ils représentent une source d'inspiration pour répondre à des enjeux environnementaux et de santé humaine (Benyus, 1997). Les organismes marins, comme les moules ou les vers marins, ont développés des colles majoritairement composées de protéines, leur permettant d'être fixes face à la houle marine (Bianco-Peled and Davidovich-Pinhas, 2015). La composition de ces colles a inspiré le développement de plusieurs adhésifs utilisés en médecine pour la suture de plaies (Perrini et al., 2016). Les organismes terrestres produisent également des colles leur permettant de contrer la gravité et les forces d'inertie, mais celles-ci restent encore peu étudiées malgré leurs nombreux avantages. En effet, ces colles résistent à de grandes variations de température, d'humidité et sont majoritairement composées de protéines.

Au cours de ma thèse, je me suis intéressée à la colle produite par les Diptères. L'ordre des Diptères comprend les mouches, dont le cycle de vie est divisé en plusieurs stades: le stade embryonnaire, plusieurs stades larvaires, le stade pupal pendant lequel l'animal réalise sa métamorphose et le stade adulte (Ashburner et al., 1989). Durant le stade pupal, l'animal est immobile, ce qui le rend vulnérable aux conditions climatiques et aux prédateurs. Chez de nombreuses espèces de Diptères, la pupe est attachée à un substrat durant toute sa metamorphose qui dure plusieurs jours, jusqu'à ce que l'adulte émerge. Les mécanismes d'adhésion diffèrent entre les espèces. En particulier, plusieurs espèces dont les drosophiles produisent une colle dans les glandes salivaires pendant le troisième stade lavaire, qui est excrétée et répartie à la surface du corps de l'animal et juste avant l'immobilisation de l'animal (Fraenkel and Brookes, 1953). Dans la nature, les pupes sont observées sur une grande variété de substrats comme des feuilles, bois, fruits mûrs, et au laboratoire sur du papier, du plastique ou du verre. Les Drosophiles sont un modèle de recherche en biologie et particulièrement en génétique, le génome de nombreuses espèces est séquencé, leur environnement et mode de vie est bien étudié et de nombreux outils génétiques sont développés chez l'espèce modèle *Drosophila melanogaster*.

La composition de cette colle a été étudiée chez *D. melanogaster*. Elle est composée de huit protéines, nommées Salivary Gland Secretion (Sgs), Sgs1, Sgs3, Sgs4, Sgs5, Sgs5bis, Sgs7, Sgs8 et Eig71Ee (Da Lage et al., 2019; Korge, 1975). Ces protéines sont codées par les gènes *Sgs*, dont l'expression est régulée par l'hormone ecdysone. Les protéines de colle peuvent être distinguées selon deux groupes. Le premier est composé de Sgs1,

Sgs3, Sgs4 et Eig71Ee. Ces protéines sont longues, possèdent des répétitions et plusieurs cytéines, prolines, sérines et thréonines ainsi que des O-glycosylations. Le second groupe est composé de Sgs5, Sgs5bis, Sgs7 et Sgs8. Ces protéines sont riches en cystéines et sans répétitions internes (Da Lage et al., 2019; Farkaš, 2016).

Les gènes *Sgs* sont présents chez d'autres espèces de Drosophiles que *D. melanogaster*, et le nombre de copies varie entre espèces (Da Lage et al., 2019). En particulier, *Sgs3*, *Sgs7* et *Sgs8* forment un cluster de gènes ayant évolué plus rapidement que les autres gènes de la colle, avec plusieurs duplications et délétions au cours de l'évolution.

De précédents travaux effectués au laboratoire ont mesuré la force d'adhésion de pupes de *D. melanogaster* (Borne et al., 2020). Cette force est de 217mN pour 1 mm^2, soit environ 200 kPa, ce qui correspond aux rubans les plus adhésifs utilisés dans le commerce (Du et al., 2021). La force d'adhésion a également été testée pour trois autres espèces (Borne et al., 2021a,b): alors que *D. simulans* possède une force d'adhésion équivalente à celle de *D. melanogaster*, celle de *D. suzukii* est plus faible et celle de *D. hydei* plus élevée.

La fonction précise de la colle et des protéines qui la compose est en encore inconnue. De récentes expériences prouvent que la colle ralentit la prédation des fourmis, les pupes attachées étant indemnes ou consommées sur place alors que des pupes non attachées sont emportées par les fourmis et consommées au nid (Borne et al., 2021b). Les Drosophiles présentes dans de nombreux écosystèmes à travers la planète, nous pouvons supposer que les colles de différentes espèces sont adaptées à une grande diversité de substrats et de conditions environementales. La colle des Drosophiles étant principalement composée de protéines et de sucres, elle est biodégradable et biocompatible. De plus, les outils génétiques développés sur la Drosophile font de cette colle un modèle prometteur pour développer de futurs adhésifs.

L'objectif de mon travail a été de mieux caractériser les propriétés adhésives de la colle et l'évolution des gènes de colle chez plusieurs espèces de mouches, afin de faire de la colle de drosophile un nouveau modèle de bioadhésion.

1. Evolution des gènes de la colle et de leurs régions génomiques.

Dans une première partie, mon travail a consisté à étudier l'évolution des gènes de colle chez plusieurs espèces de Drosophiles.

Tandis que *D. melanogaster* possèdent huit gènes *Sgs*, ce nombre varie pour d'autres espèces de Drosophiles. En particulier, une étude réalisée au laboratoire a révélé que *Sgs3*, *Sgs7* et *Sgs8* forment un cluster de gènes et présentent de nombreuses duplications et délétions en comparaison avec d'autres gènes de la colle (Da Lage et al., 2019). Nous avons focalisé notre étude sur le groupe *Sgs3/Sgs7/Sgs8* et comparé son évolution avec un gène de la même famille, *Sgs1*.

Nous avons examiné les génomes de 24 espèces de drosophile, séparées par 30 millions

d'années d'évolution. Nous avons annoté 102 gènes de la colle et identifié des copies de gènes *new glue* qui étaient jusqu'alors inconnues. Nous avons observé deux dynamiques évolutives. Un premier groupe de gènes (*Sgs1*, *Sgs3b* et *Sgs3e*) ont connu quelques duplications mais peu de délétions, inversions ou conversions génique. Les gènes du second groupe (*Sgs3b*, *Sgs7* et *Sgs8*) ont été dupliqués plusieurs fois et ont subi des évènements d'inversion et conversion génique. Nous avons observé que les régions génomiques voisines des gènes du second groupe possèdent des séquences répétées codantes correspondant à de petits gènes appelés *new glue* (*ng*). Ces gènes sont courts, 243 à 426 pb, ont un signal peptide et codent des protéines riches en threonine. Leur fonction biologique est encore inconnue.

Cette première partie a montré que certains gènes de la colle ont évolué plus rapidement que les autres, ces gènes pourraient jouer un rôle clef dans l'évolution rapide des propriétés de la colle entre espèces.

2. Propriétés adhésives chez plusieurs espèces de Diptères.

Dans une seconde partie, nous nous sommes intéressés aux propriétés adhésives de la colle de plusieurs espèces de Diptères. Jusqu'à présent, seule la colle de *D. melanogaster* a été bien étudiée (Borne et al., 2020). En collaboration avec Jean-Noël Lorenzi, ingénieur au laboratoire, nous avons développé un script d'analyse des données d'adhésion nous permettant de calculer les propriétés d'élasticité, de plasticité et de rigidité de la pupe et de sa colle. Des photos des pupes avant leur détachement ont été prises, afin d'évaluer leur taille et de mesurer la surface de colle en contact avec le substrat d'adhésion.

Nous avons effectué des tests d'adhésion témoins sur *D. melanogaster* qui valident notre protocole expérimental utilisé pour mesurer les propriétés adhésives. Ces tests montrent que la pupe est élastique alors que l'ensemble pupe-colle est plastique. L'adhésion de la pupe est stable au cours de la métamorphose et la pupe se rigidifie au cours de son developement.

Les tests d'adhésion réalisés sur 27 espèces de Diptères montrent une grande variabilité d'adhésion entre les espèces, qui n'est pas corrélée à la phylogénie. La taille de la pupe ne semble pas corréler avec les propriétés adhésives. En revanche, la surface de colle entre la pupe et le substrat varie beaucoup entre espèces et impacte la force d'adhésion. Les pupes ayant une plus grande surface de colle sont attachées plus fermement. Une fois divisée par la surface de colle, la force d'adhésion est équivalente chez la plupart des espèces. Ces résultats signifient que les différences d'adhésion observées entre espèces semblent principalement dûes à des variations dans la quantité de colle produite.

Nos résultats révèlent que *D. virilis* et *D. hydei* sont les espèces ayant la colle la plus adhésive et produite en plus grande quantité. Notre étude a permis de comparer les propriétés adhésives de plusieurs espèces, et de déterminer celles ayant un intérêt pour de futures applications industrielles.

3. Rôle des gènes de la colle dans l'adhésion des pupes.

Les protéines Sgs sont la composante majeure de la colle. Pourtant, leur rôle dans l'adhésion est encore inconnu. Dans cette troisième partie, nous avons inhibé individuellement l'expression des gènes de la colle grâce à des outils génétiques, et réalisé des tests d'adhésion. Ce travail a été effectué avec l'aide d'Isabelle Nuez, ingénieure au laboratoire, et en collaboration avec Kelly Ten Hagen, aux National Institutes of Health (Bethesda, USA), qui a analysé en parallèle les effets de l'inhibition des gènes de la colle sur l'aspect de la colle en microscopie électronique.

Nous avons utilisé deux outils génétique d'inhibition des gènes, la technique UAS-GAL4 RNAi et Gene-Switch. Grâce à la première technique, nous avons observé une diminution significative de l'adhésion lorsque *Sgs3* est inhibé et une légère diminution non significative lorsqu'un des gènes *new glue*, *CG33272*, est inhibé. La technique Gene-Switch n'a pas été concluante et nous prévoyons d'améliorer le protocole.

Cette étude révèle que le gène *Sgs3* ainsi que les genes *new glue* jouent un rôle dans les propriétés adhésives de la colle.

En conclusion, ce travail a permis de mieux caractériser l'évolution des gènes de la colle, leur fonction et les propriétés adhésives de la colle chez plusieurs espèces de drosophile. Dans mes travaux de recherche, j'ai utilisé la colle des Drosophiles comme un modèle d'étude en évolution. J'ai mis en évidence de courtes régions génomiques, les *new glue* genes, qui pourraient causer des dynamiques évolutives accélérées des gènes de la colle *Sgs3/7/8*. J'ai également montré que les propriétés adhésives des pupes évoluent rapidement entre plusieurs espèces de Diptères et en particulier la surface de colle entre l'animal et son substrat. Mes résultats montrent qu'un phénotype peut évoluer rapidement entre espèces proches, sans lien apparent avec leur distance évolutive.

La colle des drosophiles constitue un modèle d'étude prometteur des bioadhésifs, de part ses propriétés remarquables et la diversité de substrats auxquels elle adhère. Mes travaux de thèse ont permis d'explorer la diversité d'adhésion de plusieurs espèces et révèlent que certaines espèces, dont *D. hydei* et *D. virilis*, sont très adhésives et produisent une grande quantité de colle. Nos travaux s'inscrivent dans une démarche biomimétique, ces espèces étant une source d'inspiration pour le développement de futurs bioadhésifs.

Remerciements

Tout d'abord je souhaite remercier Virginie de m'avoir accueillie dans son laboratoire et m'avoir fait confiance. Merci pour ta disponibilité, le partage de tes connaissances, tes encouragements, et d'avoir été un modèle de femme scientifique. Merci à mon co-directeur de thèse, François, d'avoir insufflé une nouvelle dimension à ce projet et pour ta patience envers la novice en physique que je suis. Je suis très heureuse d'avoir travaillé à vos côtés et d'avoir bénéficié de l'exemple de deux chercheurs passionnés.

Merci aux membres du jury d'avoir accepté d'évaluer mon travail, en particulier Patricia Gibert et Patrick Flammang, rapporteurs de cette thèse.

Merci à tous les membres de l'équipe, passés et présents, Isabelle pour ton aide précieuse et ta bonne humeur, Michael Lang, Alexis, Stéphane, Michael Rera et Sunitha pour les discussions et conseils. Merci à Roshan pour les discussions, scientifiques ou non, et d'avoir été mon conseiller de carrière. Merci à Flora de m'avoir accompagnée durant mon stage et début de thèse, et Bénédicte pour ton soutien. Un merci particulier à Jean-Noël, pour ta pédagogie, ta disponibilité, et m'avoir fait croire qu'il est facile de coder. Merci aux stagiaires Louis, Sabrina et Romane pour la qualité de leur travail.

Je voudrais remercier toutes les personnes m'ayant aidée à mener à bien ce projet et particulièrement Jean-Luc Da Lage et Kelly Ten Hagen. Je remercie également Cyprien Gay pour ses conseils concernant les adhésifs.

Merci à l'ensemble des membres de l'Institut Jacques Monod et en particulier à l'équipe Konstantinides pour leurs conseils lors des labmeetings partagés.

Merci à la plateforme d'imagerie ImagoSeine et Vincent Contremoulins pour leurs conseils, leur disponibilité et leur bonne humeur. Merci à l'équipe laverie, Théo, Vincent et Clarisse, pour leur dévouement envers nos mouches, leur aide si précieuse et pour animer nos journées au labo. Je souhaite également remercier l'ensemble des équipes techniques de l'Institut Jacques Monod.

Merci aux membres de mon comité de suivi de thèse, Laurent Corté, Jean-Michel Gibert et Pierre Kerner, pour leur accompagnement et intérêt.

Merci à la team Méditerranée du 4ème étage pour le soleil, l'accent, et le soutien. Merci à Maëlys pour ta générosité et avoir été d'un grand soutien tout au long de la thèse.

D'un point de vue personnel, une page ne suffira pas à remercier tous ceux qui m'ont accompagnée tout au long de cette thèse. Merci aux Fidèles pour leur amitié infaillible, aux copaings de Toulouse et aux 1m55 pour leur soutien. Merci à Samuel d'avoir tant contribué à cette thèse. Merci à ma famille de m'encourager et à mes parents, pour leur amour et soutien inconditionnel.

Contents

Table of figures

Table of tables

Abbreviations

aa: amino acid

APF: After Puparium Formation

Bab2: Bric-à-Brac 2

bHLH: basic helix-loop-helix

bp: base pair

Br: Broad-Complex

Cys: Cysteine

DOPA: 3,4-dihydroxyphenylalanine

EM: Electron Microscopy

NGS: Next-Generation Sequencing

Ser: Serine

Sgs: Salivary Gland Secretion

Thr: Threonine

A toad can die of light!
Death is the common right
Of toads and men,
Of earl and midge
The privilege.
Why swagger then ?
The gnat's supremacy
Is large as thine

Life, is a different thing,
So measure wine,
Naked of lask, naked of cask,
Bare Rhine,
Which ruby's mine ?

Emily Dickinson

Chapter 1

Introduction

1.1 Bioadhesives in nature

1.1.1 Introduction to bioadhesion

Definitions of bioadhesion in nature

Being attached is of major importance for many living animals to fulfil biological functions: (i) maintenance of position, (ii) locomotion requiring strong adhesion, (iii) attachment to a substrate while feeding, (iv) prey capture and holding the captured prey, (v) temporary attachment between two body parts, (vi) maintenance of mechanical contact with the mating partner during copulation, and (vii) particle manipulation (Gorb, 2008). The characteristics of the surfaces on which these organisms live have shaped their attachment methods and they can be grouped into eight mechanisms: (i) hooks, (ii) lock or snap, (iii) clamp, (iv) spacer or expansion anchor (v) suction, (vi) dry adhesion, (vii) wet adhesion (glue/cement, capillarity), and (viii) friction (Gorb, 2008). These bioadhesion mechanisms can be used alone or combined (Büscher and Gorb, 2021). They are very different both quantitatively and qualitatively between organisms from the aquatic or terrestrial realm (Ditsche and Summers, 2014).

Bioadhesion is defined as the state in which two materials, at least one of which has a biological origin, are held together for an extended period of time by interfacial forces (Bianco-Peled and Davidovich-Pinhas, 2015). Three regions are involved in bioadhesion: the surface of the bioadhesive material, the surface of the substrate and the interface between these two (Peppas and Buri, 1985). In a natural environment, bioadhesion can occur at different scales, for instance between cells in a tissue, between organs, or between an organism and a substrate. We distinguish 'dry adhesion', caused by the physical structure of two surfaces, and 'wet adhesion' where a third component is present, a substance referred to as glue. A bioadhesive is a glue produced or obtained from living organisms. In this thesis, we will mainly focus on wet adhesion.

Quantification of bioadhesion

Quantification of bioadhesion is mostly done by measuring its adhesive strength through in vitro techniques. Adhesive strength is measured as the intensity of the external force required to detach the bioadhesive from its substrate. This force can be applied during a tensile test, shearing test or peeling test (Figure 1.1).

The tensile and shearing test force directions are respectively perpendicular and parallel to the surface of contact between the adhesive and the substrate. The peel force is applied at the edge of the bioadhesive and its direction is perpendicular to the surface joint. For both tensile and shear tests, the force is equally distributed along the bioadhesive while for peeling test, it is concentrated at the edge. Tensile tests are often used to assess bioadhesion (Bianco-Peled and Davidovich-Pinhas, 2015). Adhesion forces are given in Newton. They are divided by the contact surface area to measure the adhesive capacity of the material in Pascal (i.e. in Newton per square meter) independently of its quantity (van den Boogaart et al., 2022). The external force applied during adhesion tests can be global, e.g. with gravitational or centrifugal forces, or it can be applied to a part of the animal through a tether (van den Boogaart et al., 2022).

When studying the adhesion of living animals, we must consider the behavior of the animal during the assay. Indeed, previous experiments revealed that adhesion can be voluntary and that the animal can react over the adhesion tests device (Eisner and Aneshansley, 2000).

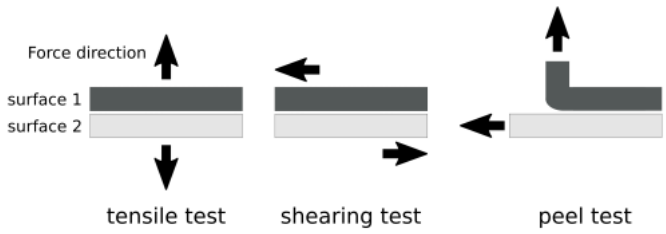


Figure 1.1: **Schematic representation of tensile, shearing and peel tests.** Adhesion is measured between a surface 1 and a surface 2. The adhesive is between these surfaces (not shown) (Bianco-Peled and Davidovich-Pinhas, 2015).

Mechanisms of bioadhesion in nature

A bioadhesive consists in a substance able to wet and diffuse on the two surfaces to create bonds between them with physicochemical forces (Khanlari and Dubé, 2013). The bioadhesive and the substrate are brought together through two stages. During the contact stage, the bioadhesive and the substrate form an intimate contact. The consolidation stage creates a stable joint between the two surfaces, able to resist mechanical stretching encountered in nature.

The bioadhesive material is attracted to the substrate via intermolecular forces ex-

isting between them. These forces are influenced by the nature of the molecules, their distance from their surface and their environment. Various interfacial forces are involved in the consolidation stage and are both chemical and physical. Steric repulsion is the strongest force involved and corresponds to the contact between two solids where their respective electrons do not interpenetrate. Covalent bond corresponds to electron sharing between two molecules. This bond is strong and stable at ambient temperature. Ionic and hydrogen bonds are respectively electron and proton transfer from one molecule to another; they have an intermediate strength. Van der Waals force is weak, short-range and corresponds to attractive forces between molecular electric dipoles (Israelachvili, 2011).

The environmental characteristics in which an organism lives put a strain on its morphology and lead to its adaptation to the forces encountered (Flatt and Heyland, 2011). In particular, aquatic and terrestrial organisms have evolved according to their respective environments. Mechanical properties of aquatic and aerial environments differ greatly. Water is 800 times more dense and 50 times more viscous than air. Terrestrial organisms are subject to gravity but their locomotion is eased by the low density of the atmosphere. On the other hand, aquatic organisms are less impacted by gravity because it is balanced by buoyancy, but deal with water flow and its hydrodynamic forces (Delroisse et al., 2023). Bioadhesives play an important role in the adaptation of organisms to their environment. They also represent a source of inspiration to develop more sustainable adhesive biomaterials as millions of years of evolution have selected diverse efficient biodegradable substances.

In the following sections, we will investigate what are the main wet adhesion strategies in the aquatic and terrestrial environment, with wet adhesion corresponding to the presence of glue in adhesion. We will refer to a terrestrial environment as surrounded by air, even if humidity may be present in the air. Adhesion in the aquatic environment is characterized by the attachment organ and the attachment surfaces, at least, being surrounded by water. In nature, immersed attachment can also occur in a terrestrial environment, for example an insect attached to a surface with a drop of water present at the insect/substrate interface. Even if there are overlaps between aquatic and terrestrial conditions, most of the known adhesion techniques occur in only one of the two systems (Ditsche and Summers, 2014).

1.1.2 Examples of aquatic wet adhesion

Various attachment methods of marine animals

Adhesion in marine environment can be either permanent, transitory, temporary or instantaneous. Permanent adhesion implies the secretion of a cement such as the one produced by bivalve molluscs. Transitory adhesion enables locomotion during adhesion, this is the case of marine flatworms using a viscous film they produce to attach to a substrate and

move on it. Temporary adhesion refers to a firm adhesion that enables to attach and detach quickly, like the echinoderm tube feet. Instantaneous adhesion corresponds to a fastly secreted adhesive from single-use adhesive organs, it is observed in ctenophores for prey capture and sea cucumbers as a defense mechanism (Flammang et al., 2016; Delroisse et al., 2023).

It was proposed that adhesives produced by marine organisms are mainly made of an association between protein and glycans (Tyler, 1988), permanent adhesives being mostly composed of proteins and nonpermanent adhesives made of proteins and carbohydrates (Flammang et al., 2016). In both cases, adhesion relies on glycosylated proteins, rich in glycine, cysteine and serine.

In an aquatic environment, substrates are usually covered with a biofilm, a layer of microorganisms growing on a surface, or fouling organisms attached to floating objects. Biofilms change the properties of the initial surface and can inhibit or enhance attachment of other marine organisms on it (Ditsche and Summers, 2014). We will now examine three different examples of aquatic wet adhesion in marine environment.

Mussels

Mussels (*Mytilus* genus) are sessile bivalve molluscs that live attached to a great variety of surfaces in their natural intertidal and subtidal ocean habitats (rocks, wood, ships, seaweed or other animals) but also to materials found in a laboratory (glass, plastics, Teflon, metals) and biological materials (teeth, bones, cells, and tissues). So far, there is no material created by humans that can compete with such a wide range of adhesion substrates (von Byern and Grunwald, 2010).

Permanent attachment is done by the byssus, a mussel structure composed of a stem that divides in several byssal threads ending with a byssal plaque. These plaques are adhesive and made of thread matrix proteins, prepolymerized collagens and a few proteins called Mfp for mussel foot proteins numbered Mfp-1 to Mfp-6 (Figure 1.2). The exact functions of each of these proteins are not known yet. They are rich in the modified amino-acid DOPA (3,4-dihydroxyphenylalanine), Mfp-3 and Mfp-5 having the highest content (respectively 20% and 30%). DOPA can form multiple interactions, such as hydrogen bonds, ionic bonds, or hydrophobic interactions, with a large variety of substrates (Bianco-Peled and Davidovich-Pinhas, 2015). DOPA also favours cross-links contributing to the cohesive strength of the adhesive plaque (Benedict and Waite, 1986). In particular, DOPA catechol side chain forms highly reactive O-quinone in alkaline seawater or after oxidation, which creates covalent bonds. Mfp-2 is also rich in cysteine (6 mol%) able to make disulfide bonds (Rzepecki et al., 1992) and Mfp-3 in glycine (25–29 mol%) (Zhao et al., 2006).

Mussel foot proteins are secreted in closed environment, which protects them against dissolution in seawater. Before secretion, the foot is pressed against the surface, creating a cavity in which the mussel monitors the chemistry conditions and by suction generates a

temporary attachment. In particular, pH and redox conditions preserve both reduced and oxidised forms of DOPA. Managing pH and redox conditions to maintain DOPA activity is energetically costly for mussels, but worthy as DOPA is an essential compound for adhesion (Waite, 2017; Yu et al., 2011). Secretion of Mfp-3, Mfp-5 and Mfp-6 occurs within seconds and yields a mixture denser than water. Mfp-3S, a variant of Mfp-3, is able to self-coacervate (Waite, 2017) : a coacervate is an aggregation of macromolecules separated from an aqueous phase by liquid-liquid phase separation after a change in pH, temperature or ionic strength. Self-coacervation, done by a single molecule bearing opposite charges on its chain, is distinguished from complex coacervation, occurring by the interaction of oppositely charged polyelectrolytes. In both cases, a dense phase (coacervate) and a dilute phase (supernatant) are separated (Delaney and Fredrickson, 2017). In the case of mussel glue, the coupling of oppositely charged sites in Mfp-3S leads to hydrophobic interactions and fluid–fluid phase separation which prevents dilution in sea water (Waite, 2017; Wei et al., 2014). Once the secreted fluid is solidified, the plaque detaches from the substrate and the mussel is held by the glue thread (Waite, 2017).

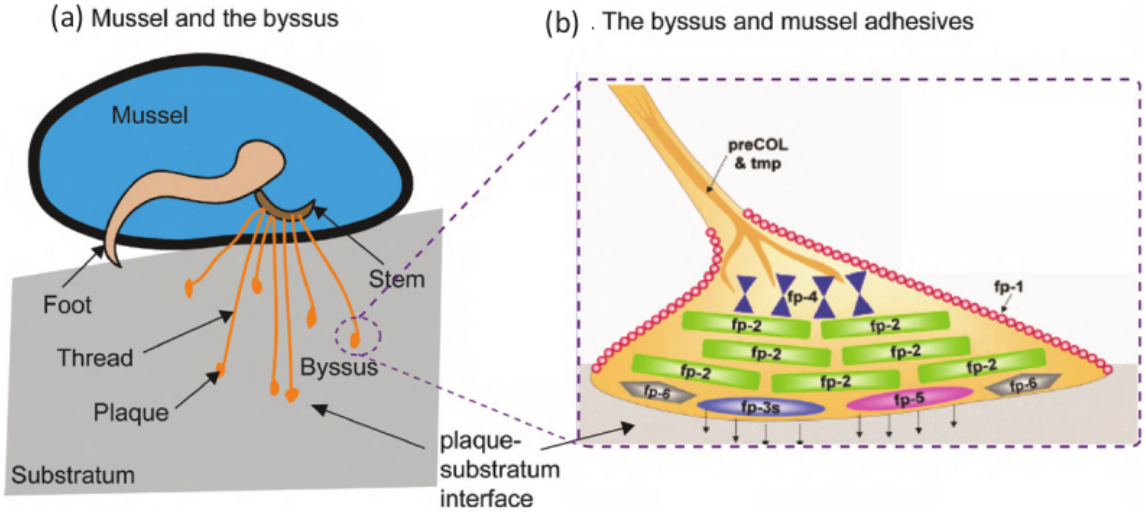


Figure 1.2: **Schematic of mussel adhesive system.**
(a) Adult mussel attached to a subtrate by its byssus; (b) Byssal plaque proteins organisation (Bianco-Peled and Davidovich-Pinhas, 2015).

Sandcastle worms

While mussels and other invertebrates live in mineral shells that they produce themselves, the sandcastle worm (*Phragmatopoma californica*) and other closely related species assemble sand particles to form a tubular shell (Stewart et al., 2011). These small gregarious marine worms build their own underwater dwellings in colonies by excreting a glue which cements together the collected particles. The final shape of individual tubes attached together in a foam like pattern gave the common name of the sandcastle worms (Figure

1.3). They live in the intertidal zone of the Californian coast where water flow is sufficient to bring the building sand particles. The cement must then be resistant to intertidal environmental constraints (Stewart et al., 2011).

Upon secretion, the glue solidifies within seconds (Stevens et al., 2007; Stewart et al., 2004). It is made of at least three constitutive proteins named Phragmatopoma californica proteins, Pc1, Pc2 and Pc3 (Waite et al., 1992; Zhao et al., 2005) and large amounts of Mg^{2+} and Ca^{2+} (Stewart et al., 2004). Like for the mussel glue proteins, Pc1 and Pc2 are basic and harbour lysine and glycine-rich repeats with DOPA (Waite et al., 1992). The third protein Pc3 is rich in phosphorylated serines which yields acidic properties and contains DOPA too (Stewart et al., 2004; Zhao et al., 2005).

Pcs being water-soluble, the two oppositely charged proteins (Pc2 and Pc3) are stored in separate secretory granules, which prevents the formation of a coacervate in the glands, prior to secretion (Stewart et al., 2011). Sandcastle worms secrete their glue directly in their open surrounding environment and their secretion in seawater would lead to their dissolution. Complex coacervation of polyanions (polyphosphoserine-rich protein) and polycations (lysine-rich proteins) prevents water dissolution and makes the glue a coacervate (Stewart et al., 2004; Wei et al., 2014). Moreover, at seawater pH (8,2), the interaction between Mg^{2+}, Ca^{2+} and Pc3 phosphates are less soluble, which provides the excreted glue a gel-like structure. Finally, DOPA residues oxidise to yield quinones which create cross-links with cysteines (Zhao et al., 2005).

Figure 1.3: **Aggregation of sandcastle worms: *Phragmatopoma californica.*** Tubes are made of sand and glue and are aggregated together forming a 'sandcastle' structure. From www.inaturalist.org.

Sea stars

The phylum of echinoderms is known for its extensive use of adhesion mechanisms. They are mostly benthic animals, their lifestyle is adapted to interact with the sea floor. Among them, asteroids (sea stars) are striking in using both non-permanent and permanent adhesion through their life cycle. As adults, they are non-permanently attached for locomo-

tion, feeding or burrowing (Santos et al., 2005). Tube feet are external appendages of the water-vascular system and are used for temporary adhesion as they are able to both attach strongly to a substrate and detach to create a new adhesion further away (Flammang, 2020; Thomas and Hermans, 1985). Each tube foot is made of a basal hollow cylinder, a stem and a disc at its flat extremity. Sea star attachment was once thought to be related to suction (Mooi, 1986) but its morphology studies revealed that it relies on tube feet secretions (Flammang et al., 1994; Thomas and Hermans, 1985). Tube feet discs are flat and do not have a suction cavity (Thomas and Hermans, 1985; Hennebert et al., 2012). Moreover, the detachment force and tenacity of single tube feet do not vary with pulling angle or surface perforation, as would be expected for a sucker (Hennebert et al., 2012).

The mode of attachment is based on a duo-gland system (Hermans, 1983). It refers to two types of cells producing antagonist secretions: an adhesive one and a de-adhesive one. In sea stars this mechanism has been inferred from the observation of adhesive cells releasing a secretion during attachment to a substrate while the de-adhesive gland did during detachment (Flammang and Walker, 1997; Santos et al., 2005). Adhesive secretions are released onto the substrate and the disc adapts its shape to the substrate roughness (Jangoux and Lawrence, 1996). As a result, the adhesive strength of the animal increases with the surface roughness and the contact area (Santos et al., 2005). The tube foot adhesive strength ranges from 0.17 to 0.43 MPa in asteroids (Santos et al., 2005; Flammang and Walker, 1997; Hennebert et al., 2010). After detachment of the tube foot, the adhesive secretions remain on the substrate as a footprint that can be stained with crystal violet (Thomas and Hermans, 1985).

The composition of the adhesive secretion is known for the sea star *Asteria rubens*, its footprints dry weight being made of proteins (20.6%), carbohydrates (8%) and a large inorganic fraction (about 40%). The protein fraction contains both charged (mainly acidic) and uncharged residues, with high levels of glycine, threonine, serine, proline, and alanine. SDS-PAGE gel electrophoresis on the proteins extracted from footprints reveals eight proteins bands, named sea star footprint proteins (Sfps) (Hennebert et al., 2011). Two of the Sfps (Sfp-290 and Sfp-210) are glycosylated with mannose, galactose, N-acetylgalactosamine, fucose and sialic acid residues attached to their chains (Hennebert et al., 2011). In the inorganic fraction, sulphate groups represent 2,5% of the footprint dry weight (Flammang et al., 1998) but it is not known whether they are attached to the proteins or the carbohydrates.

Transcriptomic analysis on the expressed tube foot mRNAs and proteomic analysis of mucuous and adhesive secretion of *A. rubens* reveals that the adhesive secretion is made of 34 proteins. Due to their resemblance to lectin-like and mucin-like proteins, most of these proteins are supposed to be adhesive (Hennebert et al., 2015). Only Sfp1 is fully characterized and is supposed to be involved in the cohesion of the adhesive (Hennebert et al., 2014).

Unlike in mussels and sandcastle worms glue, DOPA has not been found in the composition of the adhesive. Disulphide bonds, hydrophobic and electrostatic interactions are thought to contribute to the adhesive crosslinks and its insolubility (Flammang et al., 2016). The animal detaches by the foot cuticle from the adhesive material with a de-adhesive secretion which cleans the tube foot from its adhesive substance. The composition of this detachment fluid is still unknown (Flammang et al., 1994; Tyler, 1988).

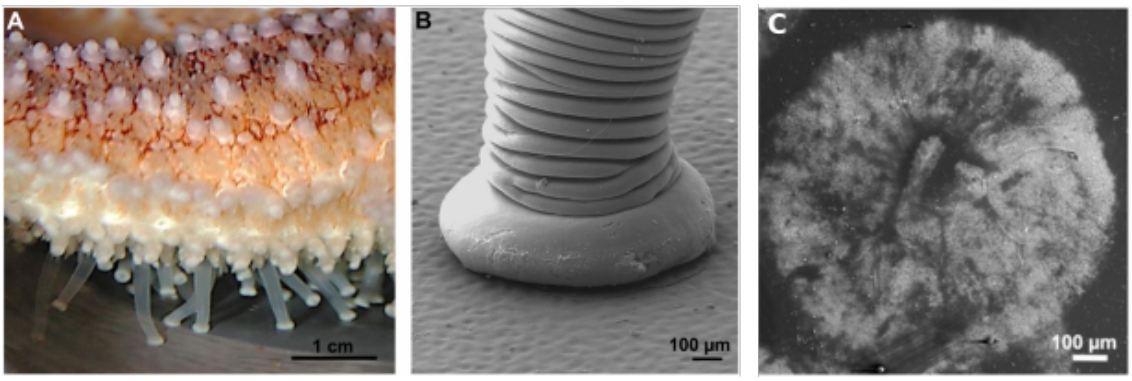


Figure 1.4: ***A. rubens* podia adhesion**
(A) Lateral view of tube feet. The stem enables movements while the distal part (the disk) is attached to the substrate. (B) Scanning electron microscopy (EM) photograph of tube foot attached to a substrate. (C) Scanning EM adhesive secretion footprints left on the substrate after detachment. Adapted from Hennebert et al. (2014)

1.1.3 Examples of terrestrial wet adhesion

Various attachment methods of terrestrial animals

Like for marine animals, terrestrial animal adhesion can be permanent, transitory, temporary or instantaneous. In a terrestrial environment, gravity is one of the main forces, together with inertial forces, against which organisms develop adhesive strategies. We can find examples of wet adhesion in diverse taxa. For example, many fungi secrete a glue, often composed of glycoproteins, for adhesion on a host or substrate (Epstein and Nicholson, 2016) and chameleon's tongue is covered by mucous secretions that, once projected, stick a prey (Brau et al., 2016).

In insects, for example the eggs of the Philippine leaf insect *Phyllium philippinicum* display adhesive properties. Upon contact with water, fibrillary adhesive structures on their surface unfold and a glue film made of glycoproteins interacts with the substrate. As a vegetal analog, the seeds of the ivy gourd *Coccinia grandis* display the same morphology and adhesion mechanisms, having fibrillary adhesive structures activated by water and a glue secretion made of polysaccharides (Büscher and Gorb, 2021; Büscher et al., 2020). While these two species have converged to similar adhesive mechanisms, they perform different functions: the insect eggs above the ground are avoiding parasitic wasps living in the soil and the seeds are dependent on the contact with the soil for a successful

germination (Büscher and Gorb, 2021).

Organs used for temporary or permanent adhesion have been found in all stages of insect life cycle (Betz, 2010). Holometabolous insects undergo complete metamorphosis, with four developmental stages: the egg, larva, pupa and adult. During the pupal stage, the animal is enclosed in a cocoon or a pupal case, and metamorphoses through structural changes by which the larva is transformed into an adult. We distinguish pupa and cocoon by the fact that the pupal case is made from the larval cuticle and a cocoon is made of silk spun before pupation. Hemimetabolous insects have an incomplete metamorphosis with three stages, egg, nymph and adult.

In terrestrial insects, the duration and type of adhesion depends on the developmental stage. Eggs and pupae are permanently attached as these stages can last several days and are very critical for the survival because the animal is immobile. As the larva and the adult have no need of their former developmental envelopes (pupal case or cocoon), the pupa and the egg can be permanently attached without impacting further life stages. In this case, the adhesive secretion is a cement. The adult stage is usually mobile and characterized by the need of an instantaneous adhesion mechanism that enables locomotion and can surpass gravity force. It requires specific physical properties of the adhesive fluid produced to enable efficient adhesion and locomotion without slipping.

We discriminate different types of bioadhesives in insects and will present three different examples of terrestrial wet adhesion.

Glue secretion around *Bombyx mori* silk

Silk is an ectodermal secretion stored as a hydrated jelly within cells that polymerises into water-insoluble filaments during passage to the external environment (Sehnal and Sutherland, 2008). It is a protein fibre coated with glue-like proteins and is used by insects or spiders to build cocoons protecting juvenile stages, protect eggs or capture prey (Siri and Maensiri, 2010). This material is found across all terrestrial arthropods. It has been extensively studied in Lepidoptera larvae for commercial interests. Caterpillars secrete silk from labial glands to form cocoons in which they undergo metamorphosis to become an adult. In particular, the species *Bombyx mori* has been extensively used by humans to produce silk for the textile industry. Its silk is a continuous filament of 90-180 nm width (Putthanarat et al., 2000) that is composed of elastic fibroin proteins (65 to 85%) coated by glue-type sericin proteins (15 to 35%) (Figure 1.5) (Akai et al., 1993; Cao and Zhang, 2016).

Sericin proteins cement fibroin proteins and act as a glue during cocoon formation. Sericins are a family of glycoproteins, ranging from 65 to 400 kDa in size and having a high serine and aspartic acid content in the species *Bombyx mori* (Teramoto and Miyazawa, 2005). They are hydrophilic and the hydroxyl groups of serine create ester bonds with the aspartic or glutamic acid of fibroin. The serine content from sericin forms hydrogen

bonds with the arginine from fibroin (Betz, 2010). Sericin is resistant to oxidation and ultraviolet radiation, is able to regulate humidity, and has antibacterial properties (Zhang, 2002).

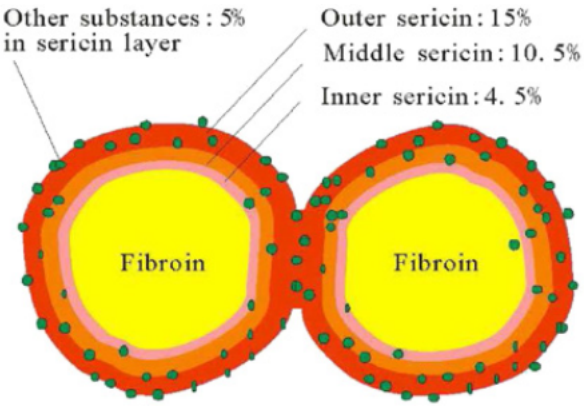


Figure 1.5: ***Bombyx mori* silk structure.** Fibroin proteins are coated by several layers of sericin proteins. Adapted from (Zhang, 2002).

Weaver ants secretion

The weaver ant genus *Oecophylla* lives in arboreal nests made of leaves. These nests are subject to rain, wind and predators. To stabilise the structure, leaves are bound by worker ants and are attached together by larval silk. The workers hold larvae between their mandibles and press them onto leaf margins to make them secrete their silk (Hölldobler and Wilson, 1977) (Figure 1.6). The silk is produced in the labial salivary gland from the final instar larva and strikingly in some species might only be used for nest construction, leaving the pupa cocoon-less (Dorow et al., 1990; dos Santos-Pinto et al., 2022). Once secreted, it creates filaments with zig-zag patterns.

Weaver ant silk is mainly made of fibroin but differs in composition with the one produced by silkworms (Siri and Maensiri, 2010). The four proteins from the *Oecophylla* weaver ant silk are named Weaver Ant Fibroin (WAF1-4). They are small proteins (about 400 amino-acids), and do not have repeated primary sequences. They have a high level of acidic amino acid and few glycine residues. The average thickness of the thread is 39 µm. Freshly produced silk is hydrophilic and becomes hydrophobic with time, protecting the nest against rain (Siri and Maensiri, 2010). It is shaped in a coiled-coil structure instead of a helicoidal structure like for silkworm species (Sutherland et al., 2010). In particular, in the species *Camponotus textor*, a fibroin protein was found to be O-glycosylated. No sericin proteins were found by proteomic analysis, electron microscopy of the fibers reveals the gelatinous structure of the outer layer (dos Santos-Pinto et al., 2022).

The weaver ant larva secretion is still poorly known and the precise mechanisms of silk adhesion have not been described. Interestingly, the production of glue is at the centre of a

Figure 1.6: **Weaver ant silk in nest building.**
Weaver worker ants hold leaves together while pressing larvae to secrete their silk and attach leaves (Reddy et al., 2011) .

social interaction between workers and larvae, with some workers holding leaves together, and others pressing larvae to join the leaves with silk (Hölldobler and Wilson, 1977). In this situation, another kind of adhesion is present, the insect adhesive pad adhesion of workers holding leaves that involves both wet and dry adhesion.

Insect pad secretion

Like lizards, frogs and spiders, insects are able to walk on vertical surfaces thanks to their pads. Two different structures are observed, hairy and smooth pads. In both cases, attachment is done by maximising the contact surface area with the smooth or rough substrate. These organs can be found on the final tarsal segment or along the leg (Gorb et al., 2002; Beutel and Gorb, 2001).

Hairy pads are covered with setae (stiff hair or bristle) or acanthae (spine structures) (Richards and Richards, 1979). These structures are flexible and maximise the contact area with the substrate and its asperities (Gorb, 2007). Smooth pads are 'pillow-like' structures, made of a soft cuticle that adapts as a whole to the surface. Because they have been more often observed than the hairy adhesive system, it is thought to have developed earlier in evolution (Beutel and Gorb, 2006). These adhesion structures can both involve an adhesive secretion (Dirks, 2014). According to Dirks and Federle (2011) all insects studied secrete an adhesive fluid on their pad. Their composition, adhesive mechanisms and biological function are not yet fully understood because of the small contact area between the pad and the substrate, and of the small amount of fluid secreted (Dirks and Federle, 2011). The secretion is produced by pore canals of the pad and the fluid can be re-absorbed during detachment from the surface (Gorb, 1998; Dirks and Federle, 2011). Field observation reveals that the beetle *H. cyanea* is not firmly attached to its substrate unless it is disturbed by an external stimulus. In this case, it uses adhesion to secure its

position by a firm foothold (Eisner and Aneshansley, 2000). Insect pad adhesive liquids are also involved in self-cleaning properties (Clemente et al., 2010).

In beetles species, the hairy pad secretion is a nonvolatile emulsion of a lipid-like substance within a watery liquid. It can be observed as footprints stained with Sudan Black (Gorb et al., 2002; Dirks and Federle, 2011). For smooth pads (e.g. ants or cockroaches), the secretion is a water-in-oil emulsion, where hydrophobic lipid-like droplets are dispersed in a watery phase. It is mainly composed of hydrocarbons, fatty acids, carbohydrates and amino acids (Dirks and Federle, 2011).

Insect attachment relies not only on adhesion (perpendicular to the substrate surface) but also on friction (parallel to the substrate surface) forces (Figure 1.7). Friction is the force generated under the action of the frictional shear stresses. It resists the motion of two elements sliding against each other. Shear forces help animals to support their body on vertical surfaces and are also essential for walking upside-down.

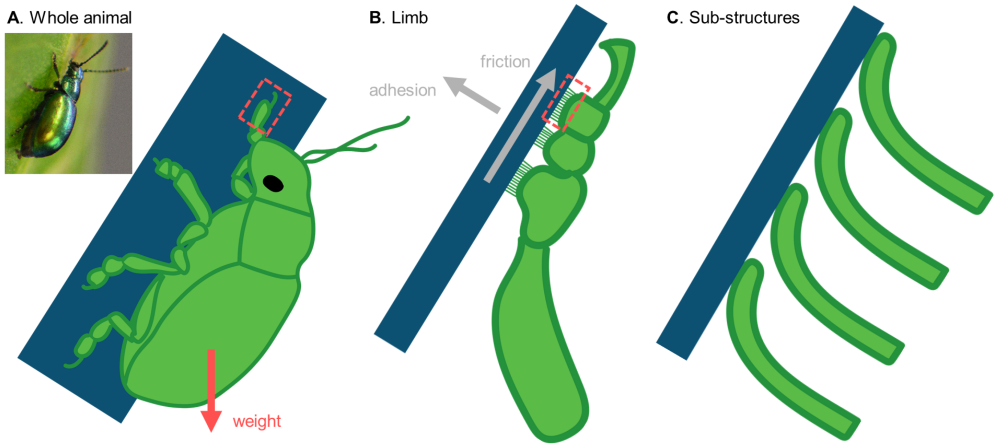


Figure 1.7: **Beetle pad adhesion at different scales**
Adhesion is displayed at the level of the whole animal (A), the limb (B) and the setae sub-structures (C). Adhesion and friction forces are respectively perpendicular and parallel to the substrate (B). Friction is oriented in opposition to the sliding direction (van den Boogaart et al., 2022).

On a smooth surface, an excess of fluid volume decreases the adhesion and friction forces. On rough a surface, where the fluid enhance the surface of contact area, these forces increase and generate traction force (Drechsler and Federle, 2006). It would seem natural to suppose that the presence of a thick fluid layer (about 100 nm) would reduce the shear forces and cause the insect to slip (Dirks and Federle, 2011; Federle et al., 2002). However, insects are still able to walk on vertical surfaces thanks to the following physical properties of their pad secretion.

Smooth pad secretions are shear-thinning Bingham fluids (Dirks et al., 2010). Shear-thinning is a non-Newtonian fluid behavior where the viscosity decreases with shear strain. A Bingham fluid behaves as a rigid body at low stresses and as a viscous fluid at high stresses. Overall, these properties allow the pad to yield and slide when large shear forces

are acting and thus prevents stresses from reaching levels that would damage the cuticle. These properties are due to the water-in-oil emulsion composition of the fluid; emulsions are more viscous than their pure components taken separately and present non-Newtonian properties (Barnes, 1994). The emulsion non-Newtonian properties induce static friction, the friction between two elements not moving relatively to each other, and prevents an object to slide (Dirks and Federle, 2011).

The hairy pad secretion in beetles and flies was found to behave like a Newtonian fluid (Abou et al., 2010; Peisker et al., 2014). Their pad fluid high viscosity (110 mPa.s, hundred times more than water) increases the friction forces which might reduce the insect speed locomotion, but also the fluid evaporation (Abou et al., 2010).

1.2 Applications in industry

1.2.1 Bioinspiration

Definitions of bioinspired approaches

Over 3,85 billion of years of evolution, living organisms have adapted to their life conditions by developing efficient strategies. This wide diversity encountered in the living world is a reservoir of ideas for humanity. Although human activity is a threat for biodiversity, several approaches aim to consider biodiversity as a source of inspiration. We can distinguish four main approaches defined by the ISO 18458:2015 standard : bioinspiration, biomimetic, bionic and biomimicry (Figure 1.8) (Biomimetics, 2015).

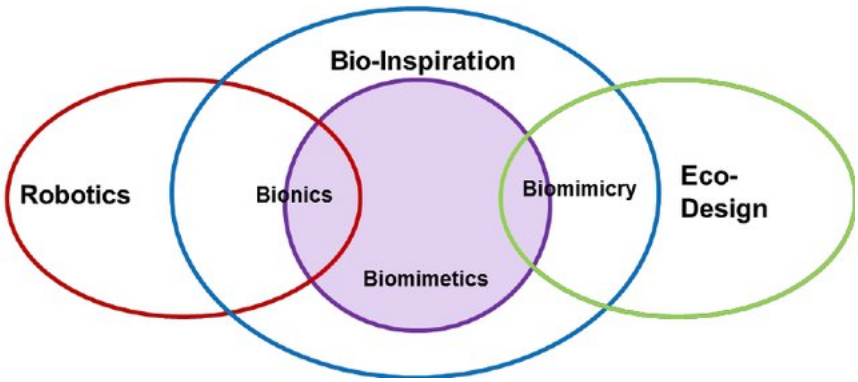


Figure 1.8: **Map of bioinspiration and other related concepts** (Fayemi et al., 2014)

Bioinspiration is a creative approach based on the observation of biological systems but does not necessarily copy the biological function. For example, the symbol of a car brand can be a jaguar but the car engine is not inspired from the animal speed capacity. Biomimetic is an interdisciplinary cooperation of biology and technology or other fields of innovation with the goal of solving practical problems through the function analysis

of biological systems, their abstraction into models, and the transfer into application of these models to the solution. While bioinspiration is influenced or informed by biology, biomimetic imitates a biological phenomenon (Rawlings et al., 2012). Bionic is a technical discipline aiming to replicate, increase, or replace biological functions by their electronic and/or mechanical equivalents. In 1997, Janine M. Benyus introduces the concept of biomimicry in her book Biomimicry: Innovation Inspired by Nature (Benyus, 1997). It is defined as a philosophy and interdisciplinary design approach taking nature as a model to meet the challenges of sustainable development (social, environmental, and economic). My thesis involves biomimetic and biomimicry motivations.

Advantages of using biomimetic adhesives

Most adhesives used in the industry are petroleum based. Environmental awareness, the need to reduce oil consumption, and regulations are driving research to replace these adhesives or improve their sustainability (Heinrich, 2019).

Natural bioadhesives possess many advantages when compared with petroleum based ones. They are mainly made of proteins hence are biodegradable and biocompatible. As bio-based structures are encouraged in industry, renewable materials are used as bulk components for diverse purposes. Bioadhesives are then interesting because of their compatibility with these substrates, due to similar functional groups. Their water soluble properties can also avoid the use of organic solvents. Bioadhesive proteins can potentially be produced at a lower cost: as they are often macromolecules linked with complex bonds, there is no need to artificially reproduce a polymeric structure. Proteins produced as animal waste can also be obtained at low price. Biopolymers found in nature are very similar to the ones present in the human body hence they can have many applications in biomedicine (Heinrich, 2019).

Finally, there is a need for highly performing bioadhesives that can withstand heat, shock, UV or moisture. Such constraints are encountered by living organisms in the wide diversity of environments to which they have adapted. Because terrestrial animals have adapted to the same environmental conditions than ours, their adhesive mechanisms are of great interest to imagine future human-made bioadhesives.

Next-generation sequencing techniques are catalysing bioadhesive design

Recent advances in genome sequencing techniques lead to the advent of next-generation sequencing (NGS) technologies that represent new powerful tools to identify bioadhesive components and create biomimetic materials. Since the Sanger dideoxynucleotide sequencing method in 1977 (Sanger et al., 1977), several DNA and RNA sequencing methods have been developed. While a single DNA fragment is sequenced at a time with the Sanger method, NGS techniques developed in 2000-2015 enable to sequence millions

of fragments simultaneously. NGS techniques are also called massively parallel and second generation sequencing methods (McCombie et al., 2019). Illumina technology, Ion Torrent and 454 pyrosequencing are different methods available (Slatko et al., 2018).

The third generation of sequencing techniques sequence longer DNA or RNA fragments compared to the second generation. Indeed, these techniques are able to sequence a single DNA molecule instead of breaking it into smaller fragments. They are commercialised by the Pacific Biosciences company (PacBio) or by the Oxford Nanopore Technologies (Slatko et al., 2018).

Proteins are the structural components of bioadhesives and biological material in general. They are also controlling the growth and assembly of inorganic components into biocomposites (Guerette et al., 2013). Their sequence information is then of high importance to understand the structure and properties of a bioadhesive and represent a template for biomimetic engineering. Recent NGS advances represent opportunities to obtain high quality genome assembly and identify the genes encoding for the adhesive proteins. One of the examples of bioadhesive understanding improvement thanks to NGS is mussel glue. First studies determined DOPA as the main contributor to adhesion in mussel glue, and most mussel-inspired adhesives developed for commercialization involve DOPA polymer (Lee et al., 2011). As decribed above, mussel foot proteins also contribute to byssal adhesion. Transcriptomics and proteomics analysis using Illumina sequencing retrieved full Mpf orthologs sequences and 15 highly expressed proteins still uncharacterised in *Mytilus californianus* (DeMartini et al., 2017; Guerette et al., 2013). This example shows that a model organism long studied for its biomaterial benefits from recent advances in sequencing techniques (Davey et al., 2021).

1.2.2 Applications from marine bioadhesives

Research on bioadhesives has been mainly based on marine animals (Melrose, 2022). The human body and the marine environment present similar features such as continuous mechanical, chemical and biological stress due to variation of salinity, temperature and pH or even the flow of tides for the marine environment and of heartbeats in human body. For these reasons, marine bioadhesives represent a source of inspiration for medical purposes because they are adapted to the physiological constraints of human body (Bianco-Peled and Davidovich-Pinhas, 2015).

Mussel adhesion is of great interest and has been a model to develop various bioengineering and medical applications thanks to its capacity to attach to a variety of substrates. So far, mussel glue is one of the best characterized bioadhesive and its study led to bioinspired adhesives by the production of recombinant mussel foot proteins or synthesized polymers in which catechol groups are added (Kord Forooshani and Lee, 2017). In particular, mussel glue has inspired medical sealants. Wounds are often mechanically

closed by sutures or staples, but these techniques might not be ideal in certain situations. Cyanoacrylates, gelatin and fibrin-based glues are thus commonly used but they can have cytotoxic effects or low cohesive strength (Burke et al., 2007). A new adhesive material has been developed, created by the reaction between the thermally triggered release of an oxidizing reagent to form hydrogel from a soluble DOPA-modified biocompatible polymer such as polyethylene glycol (PEG). The gel formed can stand shear bond strengths of 35 kPa, five times stronger than commercial fibrin surgical adhesive (Burke et al., 2007). The cell and tissue adhesive from Corning (Cell-Tak™), inspired from mussel proteins, is commercially available. It is developed based on *Mytilus edulis* polyphenolic proteins and is able to attach cells or tissues to a great variety of surfaces like plastic, glass, metal, polymers or other biological surfaces. Mussel glue also has applications in fetal surgery. After surgery, the amniotic sac does not heal. To be sutured, tissue adhesives are necessary but the commonly used ones were unsuccessful (Perrini et al., 2016). Mussel glue appears to be a promising fetal sealant: after in vitro and in vivo experiments with catechol functionalized PEG polymer, it shows no cytotoxicity, promotes fetal membrane sealing and resistance to stand pressure (Haller et al., 2012; Kivelio et al., 2013). Messersmith at the University of California-Berkeley and associates who pioneered this research topic, have patented sealant for fetal membrane repair inspired from mussel glue proteins (Messersmith et al., 2013).

Sandcastle worms glue release mechanism has inspired drug delivery in medicine. Classical tissue adhesives (cyanoacrylates, gelatin and fibrin-based glues) use is limited due to their low viscosity, their hydrophilicity and their dilution in blood and other body fluids (Lee et al., 2015). In order to better control their adhesive activation, a bioadhesive surface is coated by negatively charged alginate to form nanoparticles of glue; the glue viscosity is reduced so it could be easily injected. Once triggered by a positively charged polymer, the nanoparticles would release their content at the desired site (Lee et al., 2015). For example, Tetranite™ is a bone adhesive technology developed by the RevBio company and is inspired by the adhesion of sandcastle worms (Kirillova et al., 2018; Norton et al., 2020). This adhesive is composed of tetracalcium phosphate (TTCP) and bioinspired O-phospho-L-serine (OPLS). OPLS is abundant in sandcastle worm proteins and thought to be responsible for enhanced bone formation.

Sea star bioadhesive is less studied than the above ones because of the small size of secreting glands. A recent study focused on coatings made of two recombinant sea star adhesive proteins based on the sequence of Sfp1 from the sea star *Asterias rubens*. The adhesion of the coating ranged from 70 to 200 pN and was described as soft (Tran et al., 2021). The de-adhesive secretion produced by sea star to detach its feet from the substrate is not yet characterized but could be of interest to envisage a glue which is repositionable at short time scales.

1.2.3 Applications from terrestrial bioadhesives

Terrestrial adhesion is still poorly studied for its potential in industry. So far, only dry adhesion has been much investigated and the gecko foot is one of the examples. To climb on vertical surfaces, geckos rely on their foot toes structure, made of millions of adhesive setae responsible for interfacial adhesion. Each seta ends with 100 to 1000 smaller spatulas responsible for van der Waals forces (Bianco-Peled and Davidovich-Pinhas, 2015). The gecko adhesive mechanism has inspired many applications in robotics (Sun et al., 2023) and is governed by repositionable dry adhesion. The following examples concern industrial applications known for wet adhesion: *Bombyx mori* silk sericin, weaver ants silk and insect adhesive pads.

Bombyx mori silk has been used for centuries for textiles, but also as a suture material to close humans wounds due to its high resistance (Ude et al., 2014). A mucoadhesive has been developed by creating a polymer complex composed of polyacrylic acid and sericin bound together by hydrogen bonding (Ahn et al., 2001). The adhesive force of this complex can reach up to 15 N, which is equivalent to the one of the commercial product (Carbopol 971 PNF). Silk sericin protein and its way of extraction are still under study to develop future bioadhesives (Chirila et al., 2016). However, its wound dressing properties have led to commercial applications. Using a fibroin-deficient *B. mori* strain producing fibers only made of sericin proteins, a gel film was prepared using ethanol as a coagulant (Teramoto et al., 2008). The hydrophilic serines provides the gel with water absorption capacity, which is useful to absorb the wound fluids. The film has elastic deformation that can withstand shocks. Commercialisation of silk for medical purposes is for example done by Sofregen who commercializes SERI Surgical Scaffold manufactured.

The silk produced by weaver ants is still poorly studied for its bioadhesive industrial potential. However, ant silk is considered for tissue engineering as it is cytocompatible and enables the proliferation of cells on its web (Reddy et al., 2011).

Similarly, insect adhesive pad fluid has not yet been studied to develop a bioadhesive. On the opposite, techniques to limit insect adhesion have taken inspiration from insect adhesive pads and plants surfaces. Coevolution between insects and plants have led plants to develop surfaces reducing attachment. Inspired by the porous cuticle of some plants that are able to absorb the insect pad adhesive fluid, several mimicking techniques are under development (Gorb and Gorb, 2017). Oleophilic nanoporous surfaces have been made from different materials, such as biodegradable polymers and metal composites (Boudouris et al., 2008; Su et al., 2021). In another study, a rough porous Al_2O_3 membrane was found to reduce ladybird adhesion (Gorb et al., 2019). Based on anti-adhesive plant mechanisms, a slippery paint (InsectiSlide) which prevents insect attachment, was developed in Walter Federle group at the University of Cambridge.

1.2.4 Future of bioadhesives

An ideal bioadhesive would be multifunctional, able to adapt and perform on different substrates under variable conditions (Wanasingha et al., 2021). Bioadhesives biocompatible and biodegradable properties make them powerful tools for medicinal applications. In particular, bioadhesive are employed in drug delivery to retain the drug at the site of action (Bianco-Peled and Davidovich-Pinhas, 2015). Adhesives play also an important role in the recycling process, as many materials are bonded, developing adhesives that do not restrain the recycling of the primary materials are of interest (Onusseit, 2006).

However, the biodegradability can be a disadvantge for bioadhesives in situations when it is not wanted. This is the case when hydrolytic degradation occurs: the polymer reacts with water molecules to make new chains, the structure of the bioadhesive is then lost (Heinrich, 2019).

One of the important issue with bioadhesive production is to find resources that do not compete with food production. Indeed, when a resource is used both as a food source and in the industry for other means, its production and economic cost will depend on both productions. The resource economical value fluctuation can affect populations relying on it as a food source. One of the alternative for polysaccharide-based adhesives is to use algae: rich in polysaccharides and with a limited food interest for human or animals.

1.3 Diptera pupa adhesion

In this section, we will concentrate on the subject of my research, Diptera pupa adhesion. Diptera order refers to true flies or two-winged flies, with 125 000 species described, among them mosquitoes, black flies, fruit flies and house flies (Merritt et al., 2009; Mayhew, 2007). This order includes pollinators, disease vectors and parasitic species. Diptera flies are holometabolous: they undergo complete metamorphosis and have a four-steps life-cycle (egg, larva, pupa, adult). The egg stage lasts a few days or weeks and is followed by at least three larval instar stages. The pupal stage at which metamorphosis occurs is of variable length. The adult (imago) can live a few hours up to months. The pupal stage is critical for the survival of the animal because it is immobile and vulnerable to its environment threats (climatic conditions, predators). Diptera undergo metamorphosis in a pupal case (e.g. Drosophilidae) or a silken cocoon (e.g. Simuliidae, some Chironomidae and Tipuloidea) and can be attached to a substrate by diverse wet or dry adhesion mechanisms (Adler and Courtney, 2019).

In the following sections we will focus on the attachment mechanisms for different Dipteran species and in particular on the glue secreted at the late larval stage in some species.

1.3.1 Aquatic Diptera

Most of aquatic and terrestrial diptera species live in a humid environment through their life-cycle: some are fully immersed, other temporally (Adler and Courtney, 2019). Aquatic diptera species are in water for only a part of their life cycle, the adult stage is usually terrestrial even if some species are skimming over water (e.g. Chironomidae species). Diptera is the insect group with the most aquatic or semi-aquatic species, with nearly 46000 species living closely to water.

Usually, the eggs of aquatic Diptera are laid by the female in small clusters attached to substrates and are immersed or close to water. Pupae from different species are found in a great variety of habitats and can be categorized as: (i) free-swimming pupae emerging as an adult at the surface, (ii) pupae attached to benthic substrates and emerge as an adult below the surface, (iii) larvae moving to the shore to burrow themselves as a pupa into substrates, and (iv) pupariation, a form of pupation occurring within the integument of the last larval instar (Adler and Courtney, 2019). Adhesion strategies are adapted to this variety of ecological substrates and also to the stream flow which represents a constraint.

So far, little is known about the adhesion mechanisms of aquatic Diptera. A few studies analysed pupae from the simuliid group (black flies), which undergo metamorphosis underwater in a silken cocoon. The pupa produces a silky secretion originating from the salivary glands of the larva that is spread directly on the substrate surface. The cocoon from the species *Simulium ornatum* (simuliid group) is insoluble as it ages (four days old) and is made of proteins up to 200 kDa, two of them being composed of glycine and tyrosine repeats (Kiel and Röder, 2002). Analysis of *Simulium vittatum* silk revealed it is composed of three major silk gland proteins that are phosphorylated (Papanicolaou et al., 2013). It is not known whether these proteins present adhesive properties and are responsible for the cohesion of the silk fibers or the adhesion of the animal.

1.3.2 Terrestrial Diptera

Pupae from Calliphoridae, Drosophilidae, Muscidae, Phoridae and Sarcophagidae families are entirely terrestrial through their life cycle (Figure 1.9). Terrestrial pupae are immobile and are found either on their food sources or close to (Godoy-Herrera et al., 1989; Greenberg, 1990). Species from Calliphoridae, Sarcophagidae, Muscidae and Phoridae families mostly feed on carrion. Their pupae are attached to the decaying flesh, hair or clothes of the corpses (Reigada et al., 2011). These orders are used for forensic purposes since the age of the pupae in the corpses are used to determine the time of death (Reibe and Madea, 2010).

For the Calliphoridae family, a cement produced by salivary glands at the last larval stage is used to attach the pupa (Singh and Greenberg, 1994). The composition of this secretion is unknown.

Drosophilidae are particularly studied as model organisms in biology research. In my personal experience, all the Drosophilidae species raised in laboratory conditions secreted a glue from their mouth part and spread it over their body, attaching the animal to a substrate (plastic vials, paper, glass, cotton). In *Drosophila virilis*, *Chymomyza costata* and *Drosophila montana* species, two different behaviors are observed: either the larvae chew paper with their mouth to produce a mixture made of cellulose fibers and attach themselves to the surface (Babišová et al., 2023), or they directly produce glue from their mouthparts when there is no paper in the tube they are raised in (personal observation).

Drosophilidae species are found to pupate in a large diversity of environments. *D. melanogaster* and *D. simulans* are found in rotten fruits (Sokolowski, 1985; Vandal et al., 2008) while *D. suzukii* is known as a pest threatening fruit culture and can pupate in soil or in fresh fruits (Rota-Stabelli et al., 2013). *D. elegans* is found pupating in flowers (communication from Stéphane Prigent), species from the genus *Scaptomyza* and subgenus *Tantalia* on leaves (Carson et al., 1970). Hawaiian Drosophilidae are buried in the sand (Carson et al., 1970). Interestingly, *D. carcinophila* is living on crabs and its pupae are attached to the mouthparts of the crustacean (Carson, 1967). While this species belongs to a terrestrial group of Dipteran, its life cycle is adapted to both terrestrial and marine conditions.

Terrestrial pupae are also found aggregated to each other. This is the case for the forensic species *Phoridae terranovae* and *Megaselia scalaris* (communication from Flora Borne) that are colonising animal remains and used in crime investigation, and Drosophila species such as *D. melanogaster* and *D. simulans* (Beltramí et al., 2010). When aggregated, pupae are more visible and more susceptible of being found by a predator but are also less accessible because surrounded by other individuals and heavier to be carried away into ant nests, like *Temnothorax nylanderi* ants do with Drosophila pupae (Borne et al., 2021b). In particular conditions, aggregated individuals may thus decrease the risk of attack by a predator. Aggregated *D. subobscura* larvae are able to reach inner part of decaying fruits that single larvae cannot, which seems to protect them from parasitoid wasps (Rohlfs and Hoffmeister, 2004).

1.3.3 Pupation site preference

Pupation site preference depends on abiotic parameters such as moisture, light, temperature, the substrate nature, and also varies between species.

Moisture is an important abiotic factor influencing larvae behavior and pupation site choice (Sokolowski et al., 1984). It was proposed that humidity content of pupation site is of high importance as a dry substrate can lead to desiccation (Sokolowski et al., 1986) and larvae can drown in a watery substrate (Sameoto and Miller, 1968). Indeed, the distribution of pupae in *D. melanogaster* varies depending on the moisture level of the

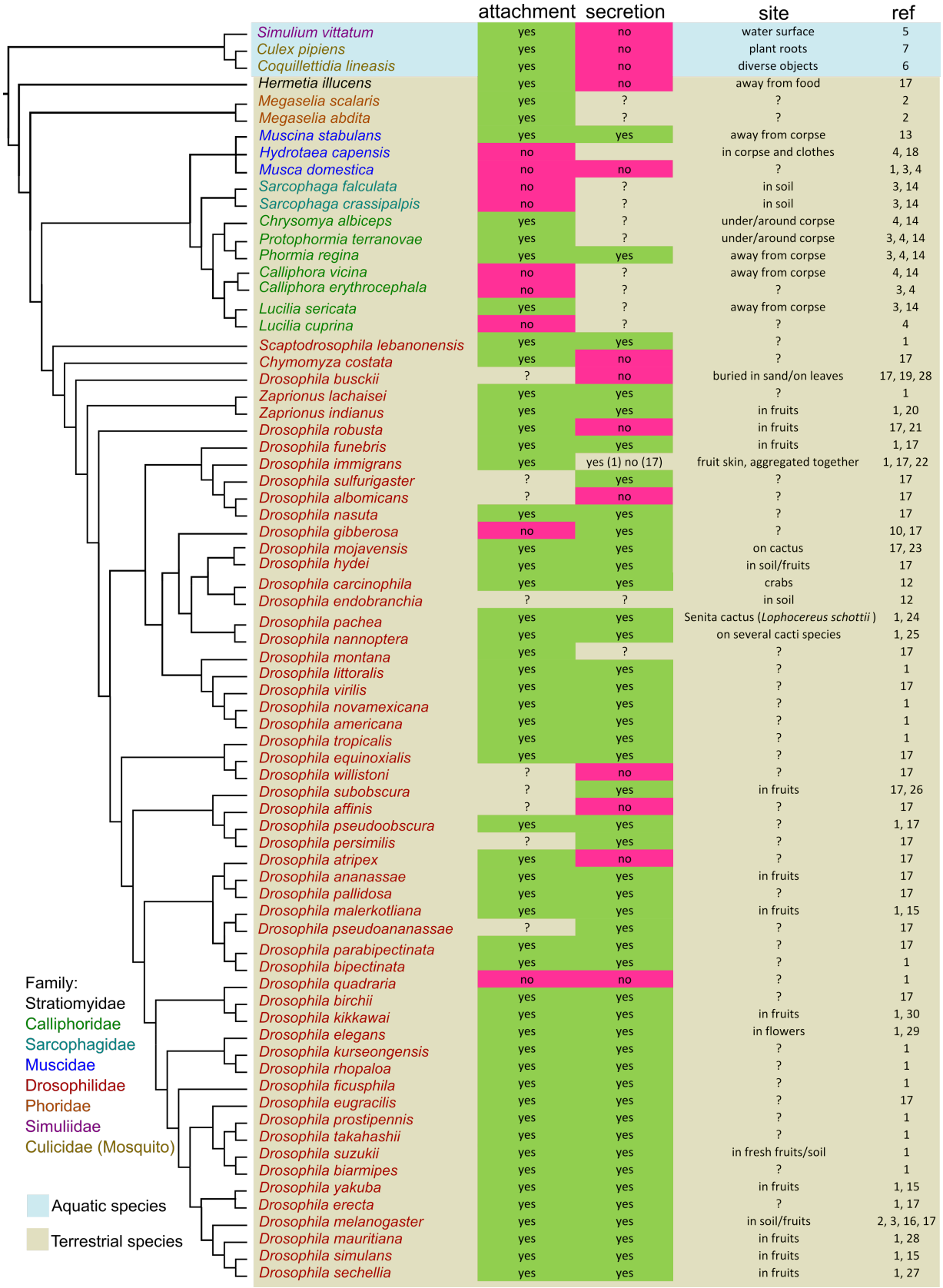
environment. In moist soil conditions, larvae tend to pupate more in the soil, resulting in better survival compared to pupariation in fruits (Sokolowski et al., 1986). Conversely, in dry soil conditions, survival is improved when larvae choose to pupariate in fruits.

Light also influences pupation behavior as an excess of light might lead to desiccation and an increased risk of predation. Behaviors vary between species. *D. melanogaster* prefers to pupate in the dark whereas *D. willistoni* and *D. simulans* tend to pupate in illuminated areas (Manning and Markow, 1981; Rizki and Davis Jr, 1953). It was proposed that the difference in this behaviour was due to interspecific competition for pupation substrate.

In the lab, where flies are raised in culture vials, the position of pupation depends on temperature (Schnebel and Grossfield, 1992). For elevated temperatures (20 - 38°C) species studied in laboratory conditions from the four groups *virilis*, *repleta*, *melanogaster* and *willistoni* are pupating within the food. At low temperatures (down to 4.5°C) larvae from *virilis* group crawl on the walls of the tubes to pupariate while the species from *melanogaster* group pupate up to 4-4.5 mm above food level. At intermediate temperatures comprised between 18.5°C and 20°C, pupation heights differ between groups, some species staying in the food while others would crawl up.

The choice of pupation substrate differs among species. As described above, some Drosophila species pupate in specific micro-environments while others are more generalists (Vandal et al., 2008). In laboratory conditions, when they are given the choice, *D. simulans*, *D. yakuba*, *D. mauritiana* and *D. malerkotliana* tend to pupariate on fruits. On the other hand, *D. melanogaster*, *D. ananassae*, *D. virilis*, *D. novamexicana* and *D. hydei* prefer to pupariate on glass. *D. rajasekari* chooses the cottons plugs of the culture vials. The substrate properties are also influencing pupation site choice. While *D. busckii* and *D. hydei* prefer to pupariate on smooth surfaces, *D. melanogaster* and *D. simulans* prefer rough substrates (Godoy-Herrera and Silva-Cuadra, 1998).

Pupation behavior might be different in the wild and in laboratory conditions, where the environmental inputs are limited.



	attachment	secretion	site	ref
Simulium vittatum	yes	no	water surface	5
Culex pipiens	yes	no	plant roots	7
Coquillettidia lineasis	yes	no	diverse objects	6
Hermetia illucens	yes	no	away from food	17
Megaselia scalaris	yes	?	?	2
Megaselia abdita	yes	?	?	2
Muscina stabulans	yes	yes	away from corpse	13
Hydrotaea capensis	no		in corpse and clothes	4, 18
Musca domestica	no	no	?	1, 3, 4
Sarcophaga falculata	no	?	in soil	3, 14
Sarcophaga crassipalpis	no	?	in soil	3, 14
Chrysomya albiceps	yes	?	under/around corpse	4, 14
Protophormia terranovae	yes	?	under/around corpse	3, 4, 14
Phormia regina	yes	yes	away from corpse	3, 4, 14
Calliphora vicina	no	?	away from corpse	4, 14
Calliphora erythrocephala	no	?	?	3, 4
Lucilia sericata	yes	?	away from corpse	3, 14
Lucilia cuprina	no	?	?	4
Scaptodrosophila lebanonensis	yes	yes	?	1
Chymomyza costata	yes	no	?	17
Drosophila busckii	?	no	buried in sand/on leaves	17, 19, 28
Zaprionus lachaisei	yes	yes	?	1
Zaprionus indianus	yes	yes	in fruits	1, 20
Drosophila robusta	yes	no	in fruits	17, 21
Drosophila funebris	yes	yes	in fruits	1, 17
Drosophila immigrans	yes	yes (1) no (17)	fruit skin, aggregated together	1, 17, 22
Drosophila sulfurigaster	?	yes	?	17
Drosophila albomicans	?	no	?	17
Drosophila nasuta	yes	yes	?	17
Drosophila gibberosa	no	yes	?	10, 17
Drosophila mojavensis	yes	yes	on cactus	17, 23
Drosophila hydei	yes	yes	in soil/fruits	17
Drosophila carcinophila	yes	yes	crabs	12
Drosophila endobranchia	?	?	in soil	12
Drosophila pachea	yes	yes	Senita cactus (*Lophocereus schottii*)	1, 24
Drosophila nannoptera	yes	yes	on several cacti species	1, 25
Drosophila montana	yes	?	?	17
Drosophila littoralis	yes	yes	?	1
Drosophila virilis	yes	yes	?	17
Drosophila novamexicana	yes	yes	?	1
Drosophila americana	yes	yes	?	1
Drosophila tropicalis	yes	yes	?	1
Drosophila equinoxialis	yes	yes	?	17
Drosophila willistoni	?	no	?	17
Drosophila subobscura	?	yes	in fruits	17, 26
Drosophila affinis	?	no	?	17
Drosophila pseudoobscura	yes	yes	?	1, 17
Drosophila persimilis	?	yes	?	17
Drosophila atripex	yes	no	?	17
Drosophila ananassae	yes	yes	in fruits	17
Drosophila pallidosa	yes	yes	?	17
Drosophila malerkotliana	yes	yes	in fruits	1, 15
Drosophila pseudoananassae	?	yes	?	17
Drosophila parabipectinata	yes	yes	?	17
Drosophila bipectinata	yes	yes	?	1
Drosophila quadraria	no	no	?	1
Drosophila birchii	yes	yes	?	17
Drosophila kikkawai	yes	yes	in fruits	1, 30
Drosophila elegans	yes	yes	in flowers	1, 29
Drosophila kurseongensis	yes	yes	?	1
Drosophila rhopaloa	yes	yes	?	1
Drosophila ficusphila	yes	yes	?	1
Drosophila eugracilis	yes	yes	?	17
Drosophila prostipennis	yes	yes	?	1
Drosophila takahashii	yes	yes	?	1
Drosophila suzukii	yes	yes	in fresh fruits/soil	1
Drosophila biarmipes	yes	yes	?	1
Drosophila yakuba	yes	yes	in fruits	1, 15
Drosophila erecta	yes	yes	?	1, 17
Drosophila melanogaster	yes	yes	in soil/fruits	2, 3, 16, 17
Drosophila mauritiana	yes	yes	in fruits	1, 28
Drosophila simulans	yes	yes	in fruits	1, 15
Drosophila sechellia	yes	yes	in fruits	1, 27

Figure 1.9: **Pupa attachment for different Diptera species.**
Pupae attached are indicated by 'yes' and pupae not attached by 'no' following the name of the species. Whether glue is observed in salivary glands or on the pupa cuticle is indicated by 'yes' or 'no'. *Simulium vittatum* is attached by surface tension, *Culex pipiens* by

piercing and *Coquillettidia lineasis* with silk and hooks. Missing information is indicated by '?'. Phylogenetic tree was obtained from TimeTree Kumar et al. (2022) and adapted from Suvorov et al. (2022). Tree branch lengths do not represent real distances. Species names are colored according to their species family and highlighted according to their larval habitat (blue for aquatic and brown for terrestrial). Sources of observation: *1*: personal observations; *2*: communication from Flora Borne; *3*: Fraenkel and Brookes (1953); *4*: communication from Damien Charabidze; *5*: Stuart and Hunter (1995); *6*: Johnson and Russell (2019); *7*: Mulla et al. (1968); *8*: Kress (1982); *9*: Berendes (1965); *10*: Shirk et al. (1988); *11*: Ramesh and Kalisch (1988); *12*: Carson (1967); *13*: Greenberg (1990); *14*: Arnott and Turner (2008); *15*: Vandal et al. (2008); *16*: Sokolowski et al. (1986); *17*: Babišová et al. (2023); *18*: Vanin (2016); *19*: Godoy-Herrera and Silva-Cuadra (1998); *20*: Marchiori and Silva (2003); *21*: Carson and Stalker (1951); *22*: Beltramí et al. (2010); *23*: Coleman et al. (2018); *24*: Lefèvre et al. (2022); *25*: Lang et al. (2014); *26*: Josso et al. (2011); *27*: Erezyilmaz and Stern (2013); *28*: Ideo et al. (2008); *29*: communication from Stéphane Prigent; *30*: Costa et al. (2003).

1.4 Glue in Drosophila species

1.4.1 *Drosophila melanogaster* life cycle and pupa morphology

D. melanogaster has been studied in the laboratory for more than a century and represents now a powerful model organism to study evolution and genetics. The life cycle of this species starts with an egg laid by the female. After one day, it goes through three larval stages (instar) which last respectively one day, one day and two days at 25 °C in laboratory conditions (Figure 1.10). About 24 hours before the end of the third instar (L3), the larva enters the wandering stage: it stops feeding and starts searching for a site to pupariate. The wandering stage lasts between 6 and 8 hours.

Then the larva begins pupariation. Pupariation lasts between 4 to 6 hours and is the process by which the animal goes from a larva to a prepupa. During this period, the animal (called puparium) is immobile, its cuticle is hardening but remains soft, anterior spiracles are extruded and the whole body is reshaped to be more rounded. After pupariation, the animal is a prepupa.

The next process, called pupation, is characterised by a new molt and the formation of the pupal case, an external protective layer of the pupa, separated from the new cuticle of the animal, then called a pupa. In *D. melanogaster*, the pupa stage will last 4 days before the adult emerges from the pupal case (Ashburner et al., 1989).

Life cycle stages may vary in length according to the species. A recent study showed that the timing of pupation varies among Drosophila species, with two main trends (Babišová et al., 2023). The first group is composed of species from the *melanogaster*

species group (*ananassae* subgroup and *saltans-willistoni* clade included) which pupates 11 to 14 hours after puparium formation (APF). The second group includes Nearctic and Palearctic species (*D. lebanonensis*, *D. pseudoobscura*, and *D. robusta*, the virilis-repleta radiation, *D. montana* and *Chymomyza costata*). These species pupate between 15 to 22 hours APF.

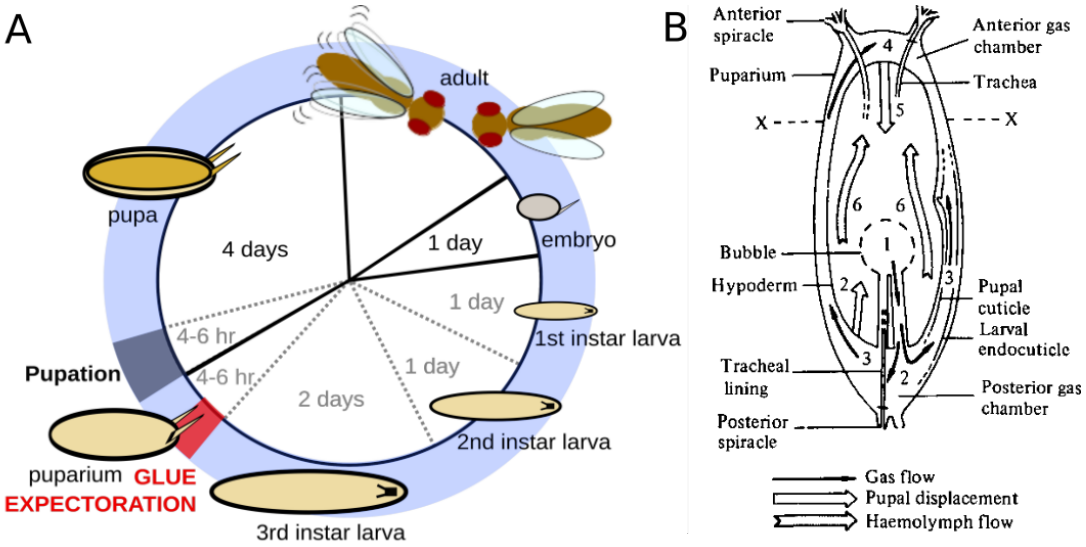


Figure 1.10: ***Drosophila melanogaster* life cycle and pupa morphology.** (A) *Drosophila melanogaster* life cycle. Sketch from Flora Borne, adapted from Bénédicte Lefèvre. (B) Pupa morphology. Both anterior and posterior spiracles are everted and a gas bubble is formed inside the pupa case (Ashburner et al., 1989).

Spiracles are openings connected to the tracheal system internally and to the epidermis externally (Figure 1.10 B). They function as gatekeepers to control gas flow and water loss. Shortly before pupariation, the anterior spiracles of the larva evert. Apart from the gas exchanges via spiracles, the prepupa and pupa are enclosed from their surrounding environment. The prepupa undergoes apolysis, the separation between the cuticle (external layer) and the epidermis (inner layer). Apolysis creates a gas bubble inside the abdomen. The cuticle transforms from soft and white to rigid and brown. Inside the pupal case, the larval body organization is modified, the tissues and organs are moving, floating in liquid with dorsal and ventral contractions. The pupa encloses an animal that is more condensed, with the adult body organization (legs, wings and eyes) visible under the cuticle. Although the pupa is immobile, it is very active inside the pupal case, as the metamorphosis is a complete re-organization of the body morphology. Movements of air bubbles, tissues and muscles contractions are observed (Ashburner et al., 1989).

According to our observations, the morphology of the pupal case varies between species. In Figure 1.11, pupae belonging to four species differ greatly in size, shape and color. Anterior spiracles size also varies between species, in particular *D. hydei*, *D. funebris*, *D. albomicans*, *D. sulfurigaster*, *D. immigrans*, and *D. robusta* have long spiracles compared with other Drosophila species from *melanogaster* group (Babišová et al.,

2023). Pupal case colour is linked to the activity of dopamine N-acetyltransferase gene (*Dat*) in *D. virilis*, *D. americana*, *D. novamexicana* and *D. lummei* (Ahmed-Braimah and Sweigart, 2015). *Dat* encodes for an enzyme involved in the insect pigmentation pathway. For these species, a high expression of *Dat* is responsible for the brown pupal case color in *D. americana*, *D. novamexicana* and *D. lummei* whereas its low expression is observed in *D. virilis* and associated with black pupal case.

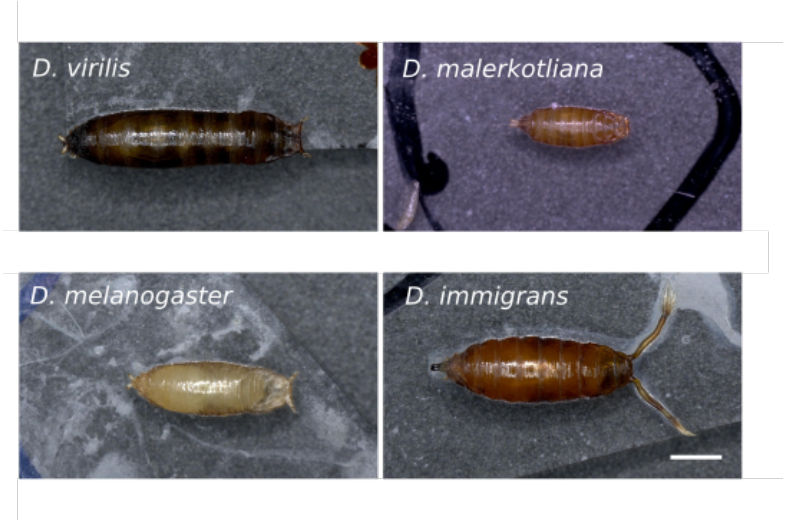


Figure 1.11: **Drosophila pupa morphology of four species.** Wandering larvae were deposited on glass slides and pupae pictures were taken between 15 to 21 hours after deposition. Scale bar is 1 mm.

At the end of the wandering stage, a glue is produced by the salivary gland cells, is secreted in the salivary gland lumen and then expelled from the mouth parts at the beginning of pupariation. This phenomenon was observed for the first time in *Drosophila melanogaster* in 1953 and the substance was described as a glue because of its adhesive properties to attach the pupa to a substrate (Fraenkel and Brookes, 1953). The larval peristaltic movements spread the fluid glue between the body and the substrate, then the glue dries in a few seconds (Heredia et al., 2021). When the adult emerges, the empty pupal case remains attached to the substrate.

1.4.2 The salivary glands

Description

The larval salivary glands (SGs) of *D. melanogaster* are a pair of secretory organ connected to the pharynx through the salivary duct (Ashburner et al., 1989). Each gland is made of a single layer of epithelial cells forming a tubular secretory tissue of ectodermal origin. The salivary glands of *D. melanogaster* are simple, they are composed of only two types of cells: secretory cells that synthesize and secrete high levels of proteins, and cuboidal epithelial duct cells that form tubes connecting the secretory cells to the larval mouth.

Salivary gland cell number is stable along larval development, individual cells grow by increasing the volume without further cell division. Each of them contains roughly 130 secretory cells in *D. melanogaster* (Ashburner et al., 1989; Farkaš, 2016). In a recent study, salivary glands from 34 Diptera species have been studied (Babišová et al., 2023). The salivary gland cell number varies between them, from 48 cells per gland for *Chymomyza costata* to 426 for *Lucilia cuprina*. For species from the *melanogaster* group, cell number per salivary gland is comprised between 107 and 133, comparable to the one of *D. melanogaster*.

Salivary glands are mostly known for their polytene chromosomes. These massive chromosomes are made of hundreds of sister chromatids tighly packed together and are the result of repeated rounds of endoreplication. This replication of the nuclear genome occurs without mitosis, leading to polyploidy (Ashburner et al., 1989).

Salivary gland development

Salivary gland formation starts at the embryonic stage. In *D. melanogaster*, at 25°C, the embryonic stage is divided into 17 steps, lasts 24 hours after egg laying and ends with the hatching of the first instar larva. At the stage 11 of embryonic development, salivary glands primordia invaginate from the ventral labial segment (Campos and Hartenstein, 1985). The salivary duct cells are the last ones to invaginate and arise from the most ventral regions of the salivary gland primordia (Andrew et al., 2000). They present salivary ducts at stage 13, ultimately connected together at stage 15. Despite the absence of mitotic activity, gland cells grow during embryogenesis, secretory cells growing faster than duct cells.

At the first instar larva stage, about 42 hours after egg laying, salivary glands are placed on either side of the body, below body wall muscles (Ashburner et al., 1989). The duct is connected to the alimentary tract and the fat body is attached to the lobe along larval development. Glands are increasing in length by cell volume increase and nuclear diameter is enlarged due to the level of chromosome polyteny. Very few vacuoles are observed within the cells at this stage. During the second instar larva, salivary glands continue to grow and vacuoles number is increasing with few granules visible.

At the beginning of the third instar larva stage, about 72 hours after egg laying, an anterior-posterior gradient in cell size and nuclear volume is observed. At the middle of third instar larva stage, granules appear in posterior cell cytoplasm and continue to grow from anterior to posterior. In late third instar larvae, posterior cells present a chromosome polyteny level twice that of anterior cells. When the larva stops feeding, posterior cells nuclear diameter remains stable while it continues to increase in anterior cells. Cell and vacuole volumes are increasing until puparium formation. Before puparium formation, glands reach their maximal size and the highest polyteny level.

Salivary gland function

In *D. melanogaster*, the exact role of salivary glands in larva is not known yet. Salivary glands display secretory activity, supposedly to help the larva escape the extracellular egg membranes, but their function at this developmental stage in still unknown (Ashburner et al., 1989). At the first, second and early third instar larva stages, traces of amylase were found. It was proposed that salivary glands might be involved in digestive enzymes production at these stages, however these enzymes are mostly found in the gut.

A recent study reported that salivary gland ablation leads to a decrease in pupal volume, in the size of wing disc, fat body cells, and brain in the larva, and in body size, body weight, and wing size at the adult stage (Li et al., 2022). This phenomenon is due to the salivary glands producing a hormone regulating systemic growth. Insect systemic growth is mainly regulated by insulin/insulin-like growth factor signaling (IIS) and mechanistic target of rapamycin (mTOR), two nutrient-sensing pathways. In Drosophila, insulin-producing cells (IPCs), present in the larval brain, secrete insulin-like peptide 2 (Dilp2), acting as a regulator of systemic growth. Both the release of Dilp2 from IPCs and IIS/mTOR signaling were found to be regulated by a salivary gland-derived secreted factor (Sgsf). Sgsf is encoded by the gene *Sgsf* and expressed in the salivary glands to be secreted into the hemolymph. In *Sgsf* mutant flies, IIS/mTOR signaling and Dilp2 expression were reduced. Overall, salivary glands thus seem to regulate the animal growth.

Salivary glands are sequentially producing three types of secretions from the third instar until the late prepupal stage.

First, the synthesis of glue proteins in *D. melanogaster* begins during the second half of the third larval instar. Glue proteins are folded in the endoplasmic reticulum and transported to the Golgi apparatus (Reynolds et al., 2019). Some glue proteins are glycosylated while in the Golgi apparatus, and all of them are packaged into granules before leaving the trans Golgi network. These granules are initially about 1 µm in diameter and subsequently fuse with one another to form large mature granules that range from 3 to 8 µm in diameter (Ji et al., 2018). In *D. melanogaster*, each salivary gland contains between 2500 and 3000 individual secretory granules (Farkaš, 2016). The granule content is then released in the lumen by exocytosis. Exocytosis is the process by which secretory vesicles are expelled out of the cell into the extracellular space. A few hours later, once the animal finds a site for metamorphosis, expectoration of the glue occurs. Expectoration is the release of secretions from the mouth. Several Diptera species (*D. albomicans*, *D. immigrans*, *D. busckii*, *D. robusta*, *Lucilia cuprina*, *Hermetia illucens*, *Chymomyza costata*, *D. affinis*, *D. willistoni* and *D. atripex*) have been found to not produce the glue protein granules or in very small quantity (Babišová et al., 2023). Factors that trigger the expectoration of the glue are not known yet.

In a second time, at the prepupal stage, salivary glands produce a secretion distinct

from the glue directly into the periexuvial cavity instead of the alimentary tract like observed for glue secretion. This is an apocrine secretion, an alternative extrusion mechanism where a part of the cytoplasm surrounds and pinches off the granules. Apocrine secretion occurs with a partial cytoplasm loss (Farkaš et al., 2014). This apocrine secretion has been shown to have an immune and defensive role in *D. melanogaster*. Indeed, mutant flies unable to produce apocrine secretion had a low survival rate when infected with bacteria, yeast and fungi (Beňová-Liszeková et al., 2021). Although the puparial case acts as a physical barrier against microbial challenges, the salivary glands apocrine secretion acts as a primary defense. Aprocrine secretion was found to be evolutionary conserved through 34 Diptera species, contrarily to glue exocytosis. Apocrine secretion would then be an ancient mechanism conserved through evolution as its critical immune role is of high importance in pupa survival. Since exocytosis appeared later, we can suppose that the glue secretion might not be essential for pupa survival.

Finally, at the late prepupal stage and one hour after apocrine secretion, an additional post-apocrine secretion was observed in a few species (*D. willistoni*, *D. mojavensis*, *D. subobscura*, *D. albomicans*, *D. eugracilis*, *D. sulfurigaster*, *D. equinoxialis*). This secretion is made of dark brown concrements that are not derived from apocrine secretion (Babišová et al., 2023). Its composition and its formation mechanisms are not known yet. After post-apocrine secretion, no further material is produced and the lumen is cleared. The possible roles of apocrine and post-apocrine secretions in pupa adhesion are not known yet.

Salivary glands fate after secretion

Once the glue is secreted at the beginning of pupariation, salivary gland lumen is becoming narrow (Ashburner et al., 1989). It gradually increases again for apocrine secretion and finally, the lumen is emptied after the apocrine fluid is released. After these events (at about 130 hours after egg laying) salivary glands are undergoing apoptosis. They are degenerating, starting from the posterior cells which nuclear membranes are breaking down. Both autolysis and phagocytosis are involved in salivary gland cells break down. Autolysis is the enzymatic digestion of cells by its own enzymes while phagocytosis corresponds to the engulfment of a particule by a cell plasma membrane. Salivary glands are broken down within a few hours and the duct disappears. Once histolysis is completed (at 133 hours after egg laying) the differentiation of new imaginal discs begins. From these structures, the future adult salivary glands will originate from a pair of imaginal rings found at the tips of the larval salivary glands. Salivary glands in the adult fly are still poorly studied and their precise function is unclear. However, no secretory granules like the ones found in larval glands are observed, suggesting that adult glands have a different function.

1.4.3 Glue genes

In *D. melanogaster*

The glue is composed of proteins, named Salivary Gland Secretion (Sgs) proteins. Sgs1, Sgs3, Sgs4, and Sgs5 proteins were initially identified in *D. melanogaster* through the migration of the glue content on acid-urea electrophoresis gels for both salivary gland and saliva extracts. In the presence of 90-95% ethanol, the saliva in the gland lumen forms a visible, opaque plug that can be readily separated from the remaining gland tissue (Korge, 1975). The Sgs proteins were named in the order of their increasing electrophoretic mobility, as there are five major bands. It was later discovered that the second band is not exclusive to the salivary glands and probably corresponds to a protein that is not specific to the glue (Korge, 1975). In certain strains of *D. melanogaster*, a sixth band was detected, and the corresponding protein was named Sgs6, although the corresponding gene has still not been identified and therefore is not considered as a glue gene (Akam et al., 1978). The relative content of each protein in the secreted glue is not known (Farkaš, 2016). The genes responsible for the diverse protein bands, *Salivary Gland Secretion* (*Sgs*) genes, were mapped on polytene chromosomes by observation of natural variations in *D. melanogaster* stocks, as well as the correlation between protein secretions and the occurrence of major puffs in the salivary gland polytene chromosomes (Korge, 1975). Puffs are bands on polytene chromosomes that are enlarged and form swellings during transcription.

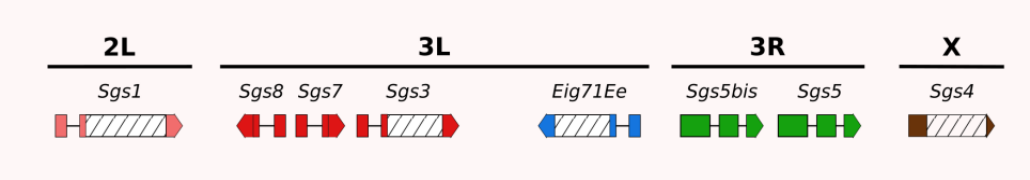


Figure 1.12: **Structure and distribution of the *Sgs* genes in *D. melanogaster*.** Horizontal lines represent introns and boxes exons. Genes represented with the same color have high similarities in their sequences. Shaded boxes correspond to repeated gene sequences. Chromosomes arms are numbered and separated by lines. Source: Flora Borne, PhD thesis.

Sgs were cloned and well-described in *D. melanogaster* during the 1980s. *Sgs7* and *Sgs8* genes were discovered as they clustered with *Sgs3* (Garfinkel et al., 1983; Hofmann et al., 1991). Additionally, the Ecdysone Induced 71Ee (*Eig71Ee*) gene was identified as part of a cluster of 11 genes regulated by the ecdysone hormone at different times around metamorphosis. *Eig71Ee* is expressed at the same time as the glue genes in the salivary glands and shares similar characteristics with the glue genes *Sgs1*, *Sgs3*, and *Sgs4*, thus being considered a glue gene (Wright et al., 1996). A recent study by the laboratory and collaborators has identified a new *Sgs* gene called *Sgs5bis* (*CG7587*), located adjacent

and sharing homology with *Sgs5* (Figure 1.12) (Da Lage et al., 2019). As In total, eight glue genes are described in *D. melanogaster*, all with a signal peptide, meaning that the Sgs proteins are transported out of the cell membrane for secretion.

Among these eight glue proteins, we can distinguish two groups. The first group, including Sgs1, Sgs3, Sgs4 and Eig71Ee, consists of relatively long proteins. They contain multiple cysteines and amino acid repeats, with a high abundance of prolines, serines, and threonines and presents O-glycosylations. On the other hand, the second group comprises relatively short proteins, namely Sgs5, Sgs5bis, Sgs7, and Sgs8. These proteins lack internal repeats and are notably rich in cysteines (Da Lage et al., 2019; Farkaš, 2016). Overall, proteins and sugars respectively represent 70% and 30% of the glue content (Ashburner et al., 1989; Korge, 1975).

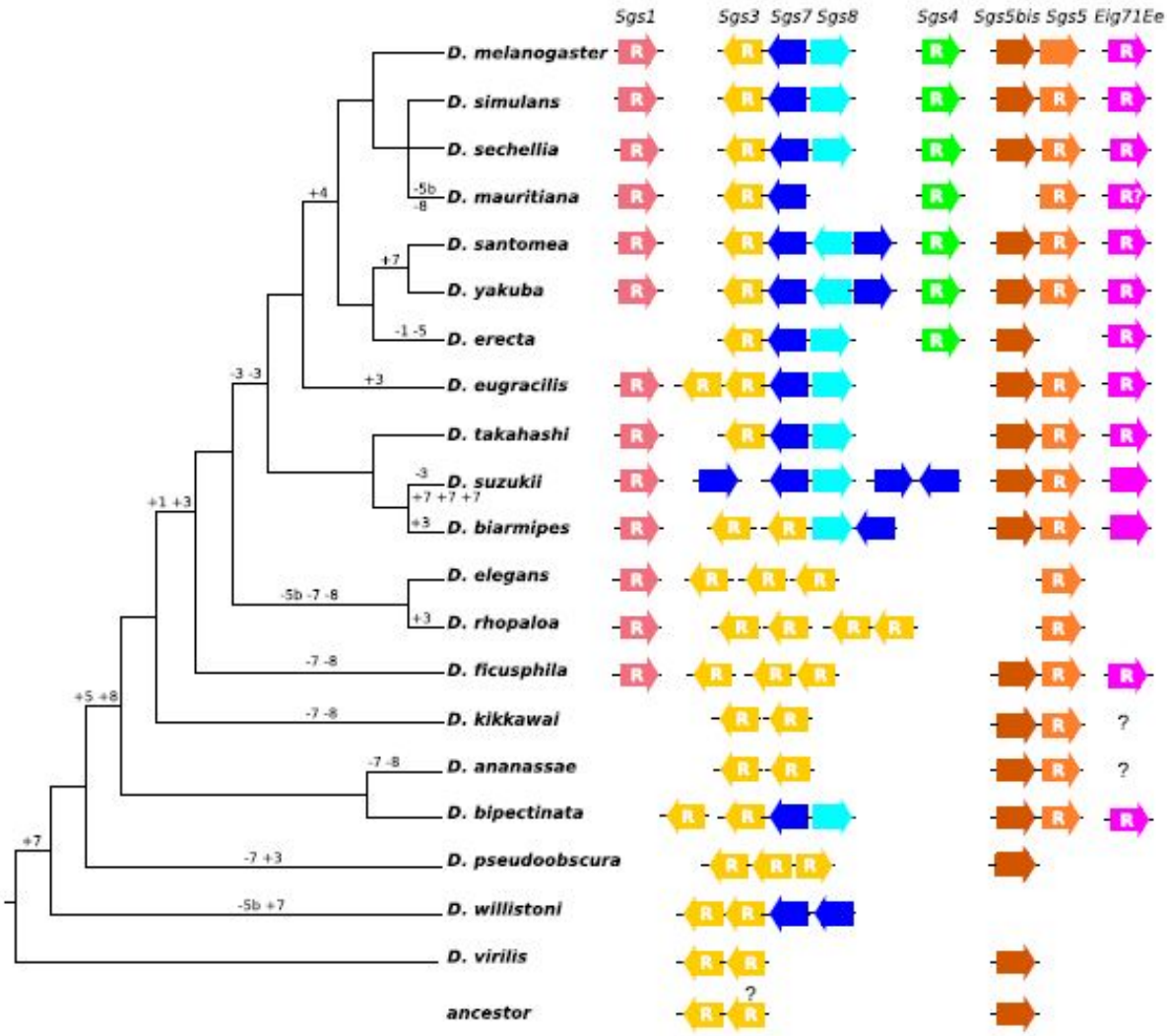


Figure 1.13: **Glue genes in 20 Drosophila species.**
Each arrow represents a gene copy. Each of the eight genes has a specific color. R indicates that the gene contains repeats. + and - respectively indicate gain and loss of one *Sgs* gene copy, followed by a number representing the *Sgs* concerned (e.g. -7 means a loss of one copy of *Sgs7*). Question marks indicate a lack of information (Da Lage et al., 2019).

In other Drosophila species

In a recent study done by the laboratory and collaborators, glue genes were annotated for 20 Drosophila species, from the *melanogaster* group and spanning until *D. virilis* (Figure 1.13) (Da Lage et al., 2019). *Sgs* sequences from *D. melanogaster* were used to annotate ortholog genes on genome assemblies derived from Illumina shortreads. No match were obtained between *D. melanogaster Sgs* sequences and genomes from species outside of *Drosophila* genus. Numerous events of gene gain and loss have been characterized. In particular, *Sgs3* is present in at least one copy in all species or several (4 copies in *D. rhopaloa*) except in *D. suzukii* where it is absent. *Sgs3* neighbouring glue genes, *Sgs7* and *Sgs8* have also experienced duplications (2 copies of *Sgs7* in *D. santomea*) or gene inversions (*Sgs8* in *D. santomea* and *Sgs8* in *D. yakuba*). The cluster of genes *Sgs3/Sgs7/Sgs8* appears to have evolved more rapidly compared to other glue genes. *Sgs3* and *Sgs5bis* are present in most studied species (Figure 1.13). On the other hand, *Sgs1* and *Sgs4* are the glue genes found in fewer species, *Sgs4* being restricted to the *D. melanogaster* subgroup and *Sgs1* spanning until *D. ficusphila*.

A positive correlation between the number of *Sgs* genes and the number of secretory cells per salivary gland lobe is observed (Babišová et al., 2023). Indeed, species with a low number of secretory cells per gland lobe present one to three *Sgs* genes (*Sgs3*, *Sgs5* or *Sgs5bis* and/or *Sgs7*) in their genome. For example, *D. pseudoobscura*, *D. ananassae* and *D. willistoni* salivary glands have 69, 75 and 78 cells per gland respectively, and have copies of the genes mentioned above. Interestingly, these species usually have duplications of *Sgs3* in their genome, but not of the other glue genes. On the other hand, species with a high number of cells per salivary gland lobe exhibit an elevated number of *Sgs* genes. For example, *D. melanogaster* and *D. simulans* salivary glands display 134 and 121 cells per gland respectively and their genomes possess the eight glues genes in single copies.

In addition, species with a low number of *Sgs* genes tend not to attach themselves to a substrate. Altogether, these indications suggest that *Sgs* number and diversity is linked to salivary gland secretory cells and the use of their secretion (Babišová et al., 2023).

Glue gene regulation

During metamorphosis, pulses of steroid hormone 20-hydroxyecdysone (hereafter referred to as ecdysone) are regulating salivary glands protein production, their histolysis and imaginal discs differentiation into future adult structures. This regulation needs to be tissue and stage specific in order to coordinate the biological responses. In the salivary glands, ecdysone pulses play a critical role in the regulation of *Sgs* gene expression during metamorphosis in Drosophila. These pulses are also responsible for larval tissue degradation and imaginal discs differentiation (Ashburner et al., 1989). At the mid-L3 stage, a first ecdysone pulse induces the expression of genes called intermolt primary-response

genes, which includes the glue genes. At the larval-prepupal stage, a second ecdysone pulse induces the expression of intermolt secondary-response genes. At the late prepupal stage, a third pulse induces the expression of intermolt late-response genes, which lead to the destruction of larval tissues. The regulation of gene expression by ecdysone pulses through time is responsible for the temporal pattern of polytene chromosome puffs, enlarged bands on polytene chromosomes, in the salivary glands (Andres et al., 1993; Duan et al., 2020). Glue genes expression starts at the first pulse and stops at the second one.

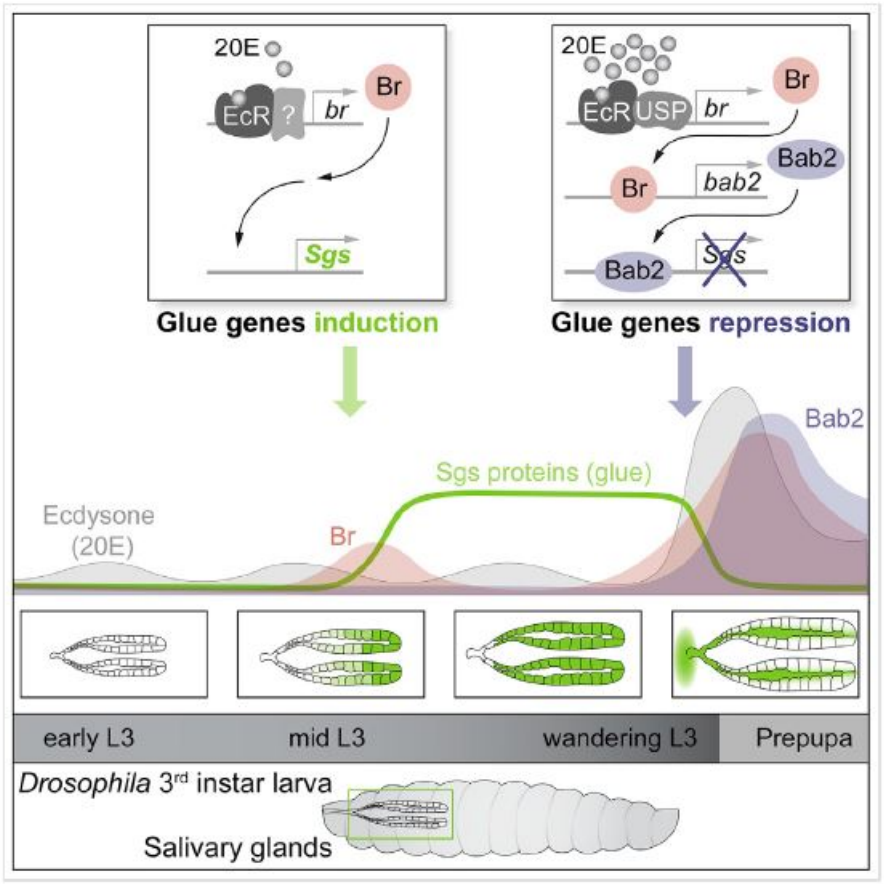


Figure 1.14: **Glue genes regulation model orchestrated by Bab2.** The mid-L3 pulse of ecdysone induces a medium expression of *Br* involved in upregulation of the glue genes. L3-prepupa pulse of ecdysone induces a high expression of *Br* which upregulates *Bab2*, downregulating the glue genes by directly binding their regulatory sequences (Duan et al., 2020).

This ecdysone activity is regulated by *Broad-Complex* (*Br-C*), a primary-response gene which encodes for multiple protein isoforms. The mRNAs coding for these isoforms are generated by alternative splicing of a single pre-messenger RNA transcript. The proteins encoded by *Br-C*, also named *Br*, are DNA binding proteins (Karim et al., 1993; von Kalm et al., 1994). *Br* activity is involved twice in glue genes regulation. The first ecdysone pulse at mid-third instar larva induces *Br* activity which triggers intermolt gene induction, including glue genes. *Br* mutations induce a decrease in glue genes RNAs

synthesis. According to a recent study, at the second ecodysone pulse, *Br* activates a secondary-response gene *Bric-à-Brac 2* (*Bab2*) (Figure 1.14) (Duan et al., 2020). *Bab2* mediates the downregulation of the glue genes, Bab2 protein directly binds to the regulatory sequences of the glue genes, particularly *Sgs3*, resulting in the downregulation of their expression.

The eight glue genes are among the ten most expressed genes in the salivary glands in *D. melanogaster* (Borne et al., 2021a). The two other genes are *CG6770*, predicted to be involved in negative regulation of cell cycle, cell population proliferation and in the regulation of transcription by RNA polymerase II; and *CG11300*, whose function is unknown.

In *D. melanogaster*, twenty other genes are specifically expressed in the salivary glands at the same stage as the glue genes and might be involved in glue functions (Li and White, 2003). Whether the products of these genes are secreted in the salivary gland lumen is unknown. These genes are highly enriched in the salivary glands 18 hours before pupa formation and downregulated from 18 hours before and until pupa formation. Among them, we find *CG12715* and *Muc68Ca*, predicted to code for an enzyme with galactosyltransferase activity and for an extracellular matrix structural constituent, respectively. The gene *sage* is also part of this cluster of genes and will be discussed hereafter. The biological processes in which the other genes (*CG10918*, *CG7606*, *CG9040*, *CG12310*, *CG13560*, *CG12360*, *CG14265*, *CG13947*, *CG13461*, *CG17362*, *CG11300*, *CG10782*, *pig1*, *new glue 1*, *new glue 2*, *new glue 3*) are involved are not known.

By analysing the 2000 bp upstream sequences of each of these twenty genes, a shared cis-regulatory region site was revealed (Li and White, 2003). For each gene, the cis-regulatory elements (promoters, enhancers, and silencers) are non-coding DNA regulating the transcription of closely located genes. In this case, this cis-regulatory region corresponds to the binding site of a basic helix-loop-helix (bHLH) transcription factor. Further analysis showed that *salivary gland-expressed bHLH* (*sage*) is the only bHLH transcription factor-endoding gene in the genome to be co-expressed with the 28 genes in the salivary glands at the third instar larva stage (Li and White, 2003). When *sage* is silenced, 8 out of the 28 genes are downregulated, and 2 are upregulated, indicating that *sage* is required to regulate some of these genes. *Sage* is only expressed in the salivary glands, from the embryo until the adult stage, and its highest level of expression is reached at the third instar larva stage.

To summarize, glue genes have been widely studied and their regulation is well described, but their precise function in the glue and their role in pupa adhesion is still unknown. Drosophila is a promising organism to study bioadhesion, as it is a model in research in genetics. In this thesis, we will compare the glue genes and the adhesive properties for several Drosophila species with the aim of developing Drosophila glue as a new model for bioadhesion.

1.5 Review - "Drosophila glue: a promising model for bioadhesion"

This part has been published in the journal Insects in 2022 August 16th. This review was requested by the Insect journal for a special issue "Contributions of Women in Insect Science". I mostly wrote the four parts dedicated to glue genes (6. Identification of the glue genes in *D. melanogaster*, 7. Characteristics and functions of the glue proteins in *D. melanogaster*, 8. Glue genes and proteins in other Drosophila species, 9. Role of Drosophila glue in natural environments).

The review and the supplementary materials can be found here : https://www.mdpi.com/2075-4450/13/8/734

insects

MDPI

Review

Drosophila Glue: A Promising Model for Bioadhesion

Manon Monier and Virginie Courtier-Orgogozo *

CNRS, Institut Jacques Monod, Université Paris Cité, 75013 Paris, France
* Correspondence: virginie.courtier@normalesup.org

Simple Summary: Before entering metamorphosis, the larvae of Drosophila flies expel a transparent glue from their mouth, which solidifies in contact with air within seconds and fixes the animal to a substrate (wood, leaves, fruits, stones, etc.) for several days until the adult emerges. This glue displays interesting adhesive properties, as it can adhere to various substrates with strengths similar to strongly adhesive commercial tapes. We review here the production, aspect, composition and role of this glue in the model organism *Drosophila melanogaster* and in other Drosophila species. The glue is made of several proteins, which have diversified rapidly during evolution. With the large diversity of substrates and environmental conditions where fly species undergo metamorphosis, Drosophila glue provides a large source of inspiration for the development of biomimetic adhesive materials. We propose several potential avenues of research for the future development of Drosophila-inspired adhesive materials.

Abstract: The glue produced by Drosophila larvae to attach themselves to a substrate for several days and resist predation until the end of metamorphosis represents an attractive model to develop new adhesives for dry environments. The adhesive properties of this interesting material have been investigated recently, and it was found that it binds as well as strongly adhesive commercial tapes to various types of substrates. This glue hardens rapidly after excretion and is made of several proteins. In *D. melanogaster*, eight glue proteins have been identified: four are long glycosylated mucoproteins containing repeats rich in prolines, serines and threonines, and four others are shorter proteins rich in cysteines. This protein mix is produced by the salivary glands through a complex packaging process that is starting to be elucidated. Drosophila species have adapted to stick to various substrates in diverse environmental conditions and glue genes appear to evolve rapidly in terms of gene number, number of repeats and sequence of the repeat motifs. Interestingly, besides its adhesive properties, the glue may also have antimicrobial activities. We discuss future perspectives and avenues of research for the development of new bioadhesives mimicking Drosophila fly glue.

Keywords: bioadhesion; glue; Drosophila; fly; Sgs; biomimetism; evolution; salivary gland; glycoprotein; mucin

Citation: Monier, M.; Courtier-Orgogozo, V. Drosophila Glue: A Promising Model for Bioadhesion. *Insects* **2022**, *13*, 734. https://doi.org/10.3390/insects13080734

Academic Editors: Christen Mirth and Lynn M. Riddiford

Received: 28 July 2022
Accepted: 12 August 2022
Published: 16 August 2022

1. Introduction

Bioadhesives are materials naturally produced by living organisms that can stick two separate items together and resist their separation [1]. These materials present singular physicochemical properties that have gone through millions of years of evolution. They are of commercial interest because they are made of proteins and sugars and hence are safe for the human body and the environment. Dental, medical and industrial applications often require adhesion in wet environments, and the marine mussel's byssus has become a leading model for biomimetic wet adhesion [2,3]. Still, mussel-inspired bioadhesives are so far only used in research and have not been tested through clinical trials [2]. They have been used as a model to perform sutureless wound closure or to seal a fetal membrane. These glue components may also have anticancer and antimicrobial applications thanks to their sticky properties at the cellular level that enable them to target cancer or microbe cells.

They also possess antifouling properties that can be used to control absorption of cells or proteins into a surface [3].

In contrast, bioadhesives that work in dry environments are less well characterized. The glue produced by Drosophila flies to stick themselves to a substrate for several days during metamorphosis appears to be a promising model for biomimetic dry adhesion. This glue is produced by the animal at the third instar larval wandering stage [4], a developmental stage during which the Drosophila larva does not feed and is searching for an appropriate site to undergo metamorphosis [5]. The glue is secreted by exocrine cells and accumulates into a pair of salivary glands. Just before entering into metamorphosis, the larva expectorates the entire content of the glands within a minute and the glue is spread all over the body. The glue solidifies rapidly and forms a transparent dry material located at the interface between the substrate and the animal [4]. After expectoration of the fluid, the larval skin hardens and encloses the now-immobile animal. The process that includes the hardening of the larval skin and the adoption by the animal of a characteristic barrel-like shape is named pupariation [5]. At the end of pupariation, the animal is a prepupa. Between 4 and 6 h after pupariation, the epidermis comes off the puparium cuticle, and a gas bubble appears in the abdomen. Eight hours later, the animal molts and technically becomes a pupa [5]. A few days later, an adult emerges and moves out from the pupal case. The glue allows the animal (as a prepupa and then as a pupa) to remain attached for several days onto its substrate, despite temperature variation, wind, rain and other environmental factors. There has likely been strong evolutionary pressure for firm attachment, as it allows the animal to remain within the environment it chose for metamorphosis, thus increasing its chances of survival.

Given the wide diversity of environmental conditions in which they live and the variety of substrates to which they attach, the numerous fly species that produce glue represent a large source of inspiration for biomimetism. Just within the Drosophila genus, more than 1600 species have been described, and they are widely spread around the world [6], with some species present on all continents while others are specific to an island, tropics or deserts [7]. The glue being located at the interface between the animal and the pupariation substrate, its composition and properties might be adapted to the nature of the pupa's microhabitat. Experiments in the laboratory have found that different species and strains choose distinct pupariation sites according to humidity, light, temperature, larval density, substrate texture and substrate consistency [5,8]. For example, in laboratory conditions, *D. busckii* and *D. simulans* prefer to pupariate on humid surfaces, while *D. melanogaster* and *D. hydei* prefer dry substrates [9]. *D. simulans*, *D. yakuba*, *D. mauritania* and *D. malerkotliana* are found to pupariate in fruits rather than on glass walls, whereas *D. melanogaster*, *D. ananassae*, *D. virilis*, *D. novamexicana* and *D. hydei* prefer to pupate on the vial walls [10]. Interestingly *D. carcinophila* and *D. endobranchia* have adapted to a humid environment as their pupae attach to the surface of the external mouthparts of land crabs [11,12]. Unfortunately, due to their small size (on the order of 1–2 mm length) and their brown color, which is usually hardly distinguishable from the environmental background, pupae are difficult to spot in nature and there is little information about pupariation sites in the wild for the various Drosophila species. While species with narrow ecological niches are expected to have precisely defined pupariation sites, others appear to stick to a large range of substrates. *D. melanogaster pupae* have been found adhered to multiple substrates, including the dry parts of various rotten fruits, grape stalks and wood [9,10,13,14]. *D. simulans* and *D. buzzatii* pupae have been observed on the dry parts of *Opuntia ficus-indica* cactus [15]. The invasive species *D. suzukii* and many Hawaiian Drosophila species often pupariate several centimeters deep in the soil [12,16,17]. The pupariation sites described in the literature might constitute only the most visible locations, while other pupa microhabitats, which are not easily accessible, may not be recorded.

In the 1970s and 1980s, the proteins that make up the *Drosophila melanogaster* glue were characterized biochemically, and their corresponding *salivary gland secretory* (*Sgs*) genes were identified. The glue genes then became a premier model to study the regulation of

gene expression, with several ecdysone pulses triggering their expression at defined developmental stages. Such studies were facilitated by the presence of polytene chromosomes in the salivary gland cells [18]. Polytene chromosomes are giant chromosomes visible with classical light microscopy that are made of hundreds of sister chromatids packed together, resulting from multiple rounds of endoreplication. A larva possesses one pair of salivary glands, with about 130 secretory cells per gland in *D. melanogaster* [19]. Each secretory cell contains about a thousand chromatids for each chromosome, thus allowing the production of large amounts of adhesive glue within a short amount of time [20].

Although the regulation of glue gene expression has been extensively studied, comparatively very little is known about the function and the adhesive properties of the glue, in *D. melanogaster* or in any other Drosophila species. We review here the aspect and ultrastructure of the glue, its adhesive properties, function, its composition in *D. melanogaster* and other Drosophila species and its potential for developing bioadhesives.

2. Research Interest in the Adhesive Properties of Drosophila Glue Is New

We searched for "Drosophila glue" in PubMed on 7th June 2022 and retrieved 152 research articles (Table S1). Among them, 32 were not relevant and 120 dealt with the glue produced by Drosophila larvae to attach the animal to a substrate during metamorphosis. We attributed to each article one of the following research topics: glue gene expression, glue gene identification, glue secretion, glue of other Drosophila species, salivary gland physiology and glue ultrastructure and adhesion.

The earliest sets of papers, starting from 1975, focused on glue secretion and glue gene identification (Figure 1). Papers published before 1975 were not retrieved by this PubMed search because abstracts are not included in the PubMed database for most articles published before 1975 (https://pubmed.ncbi.nlm.nih.gov/help/, accessed on 8 June 2022). Note that this review article also includes older papers and publications not found with these keywords.

More than half of the collected Drosophila glue papers were devoted to the regulation of glue gene expression, with a peak in publication number in the 1980s (Figure 1A, File S1). A few papers, classified as "salivary gland physiology", examined diverse aspects of the salivary glands, including programmed cell death and movements of various ions and metabolites occurring after glue secretion at later stages during metamorphosis (Figure 1E). Although the role of the glue in fixing the animal to a substrate was proposed by G. Fraenkel and Victor J. Brookes in 1953 [4], research interests in the adhesive properties of this glue are fairly recent. Surprisingly, we found only three papers focusing on Drosophila glue ultrastructure and adhesive properties, published in 2019–2021 (Figure 1F).

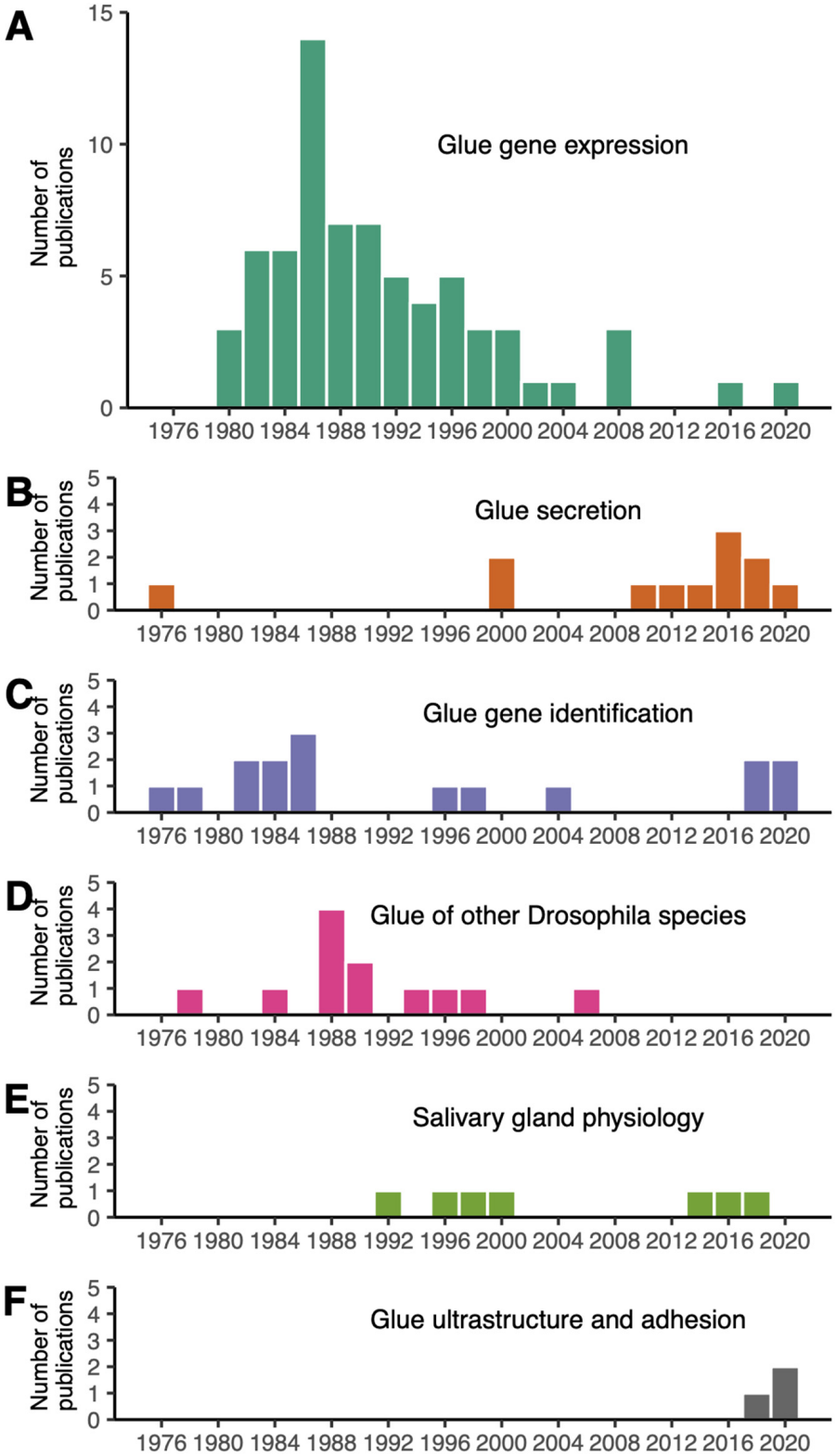


Figure 1. Distribution of Drosophila glue papers by year of publication and research topic. Five research topics are distinguished: (**A**) glue gene expression (70 articles), (**B**) glue gene identification (15 articles), (**C**) glue secretion (12 articles), (**D**) glue of other Drosophila species (12 articles), (**E**) salivary gland physiology (7 articles), (**F**) glue ultrastructure and adhesion (3 articles).

3. Aspect and Ultrastructure of Drosophila Glue

When left on glass slides, *D. melanogaster* animals usually attach on their ventral side, which presents a relatively flat surface, whereas the lateral and dorsal sides are more curved (Figure 2A–C). The glue forms an oval-shaped patch of solid transparent material of approximately 2 mm in length and 0.5 mm in width located at the surface of contact between the animal and the substrate (Figure 2C) [21]. Due to the overall barrel shape of the pupa, glue thickness varies from 0 µm (at the confocal microscope detection limit, in the middle of the surface of contact) to 20 µm (on the edges of the surface of contact). In addition, the glue can be detected when spread as a thin layer of about 0.1 µm onto the substrate outside of the surface of contact, and it also covers the surface of the pupal case that is not in contact with the substrate. It is thus reasonable to assume that the glue has good wetting properties both on the pupa and natural substrates, which probably means that the glue is highly hydrophilic. Further investigation on wetting properties and contact angle measurements will be valuable.

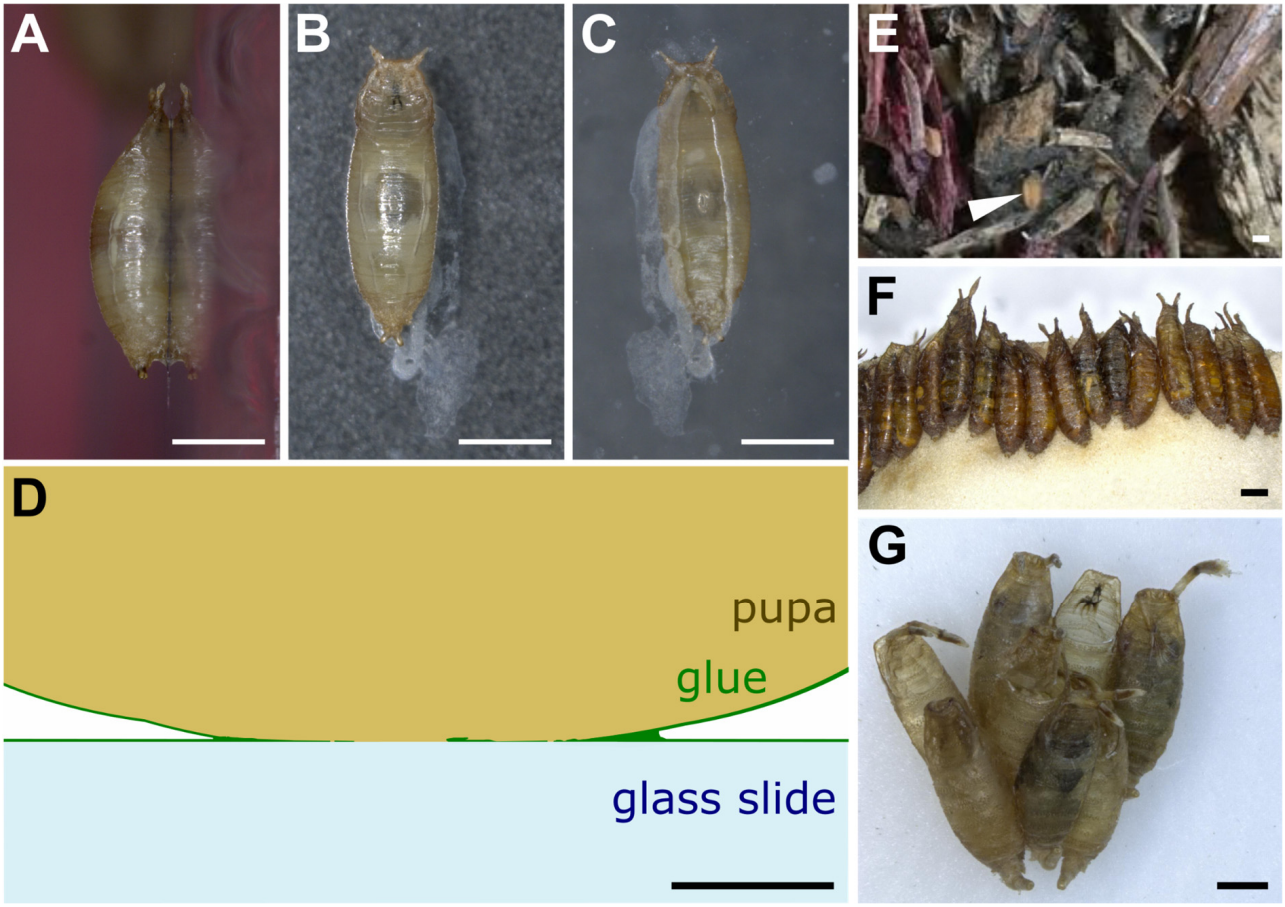


Figure 2. Drosophila pupae attached to various substrates. (**A**–**C**) *D. melanogaster* pupa attached with its own glue to a glass slide. Three pictures were taken of the same individual: (**A**) side view, (**B**) dorsal view, (**C**) ventral view throughout the glass slide. Anterior is up. White prints correspond to glue tracks on the glass slide or secretion from the larva before pupation. (**D**) Schematic transverse section of part of a *D. melanogaster* pupa attached to a glass slide with its own glue. Drawing based on confocal microscopy sections of *Sgs3:GFP* pupae obtained as in [21]. (**E**) First centimeters of soil made of wood chips from Bassevelle, France, where a Drosophila pupa (arrowhead) was found in July 2018. (**F**) Array of *D. hydei* pupae naturally attached to the plug within a laboratory vial. (**G**) Cluster of *D. acanthoptera* (Cornell University Drosophila Species Stock Center, stock #15090-1693.00) pupae found on the plastic wall of a laboratory vial. Scale bar is 1 mm in all panels except panel D, where it is 100 µm.

On the substrate near the posterior part of the animal is often found a whitish material that appears to be expelled by the intestine prior to metamorphosis and that mixes partly with the glue that is expectorated from the anterior part of the larva (Figure 2B,C) [21,22]. Whether this whitish material contributes to the adhesive properties of the glue is unknown.

In the wild and in the laboratory, pupae can be attached to the substrate either in isolation or in clusters (Figure 2D,E) [14,23,24]. When clustered, pupae are usually aligned with their anteroposterior axes pointing to the same direction (Figure 2D,E) [25]. In the wild, when pupating on Opuntia cactus, *D. buzzatii* and *D. simulans* tend to form species-specific aggregations in different locations on the cactus [15]. This clumping behavior, in which individuals of a given species closely group with each other, might improve the animals' attachment, as their glue may combine with the glue of other already attached individuals.

Scanning electron microscopy reveals that the surface of the glue is uniformly smooth and that its internal aspect is complex and structured [21,22]. The glue seems to be organized in thin layers separated by air bubbles [21] and is made of a multidirectional arrangement of thick fibers of various densities ranging from 30 to 90 nm diameter [22]. The variability in inner glue fiber thickness may serve as an elastic buffer that can accommodate mechanical stress exerted onto the animal and may allow its firm attachment to the substrate [22].

4. Adhesive Properties of the Glue

The adhesion force at the bonding surface of a bioadhesive can be measured in the laboratory as the force required to detach one material from the other under the application of a shearing, tensile or peeling force [1]. Recently, three studies of our group and collaborators used an automated pull-off adhesion test program to evaluate the force required to detach pupae from a substrate [21,26,27]. Third instar wandering larvae were let to pupariate on glass slides and kept in a box with wet paper. Fifteen to twenty-one hours later, pull-off force assays were conducted on pupae naturally attached to glass slides with their own glue using a universal test machine with a 5 N force sensor covered with double-sided tape. The glass slide with a pupa attached on it was placed under the force sensor. The force sensor was moved down until reaching the pupa, pressing onto it until a determined maximal force of 0.07 N, stilled at 0.03 N for 10 s and finally moved up at a constant speed of 0.2 $mm \cdot s^{-1}$ until a determined position. During the assay, three variables were measured: time (seconds), position of the force sensor (extension in mm) and force (N). The maximal value of the force reached during the assay when the pupa detaches from the glass slide was considered as the glue adhesion force for the individual.

The pull-off force for *D. melanogaster* pupae on glass slides was found to range from 151 mN to 269 mN with an average of 217 mN (15,500 times the weight of a pupa) [21]. By dividing the force value by the area of the pupa–substrate interface, which is approximately 1.1 mm^2, the adhesive strength is thus estimated at 137–244 kPa (1 Pa = 1 N/m^2) [21]. Adhesive strengths of the same order of magnitude (hundreds of kilopascals) are found for commercial adhesive tapes and for the mussel-inspired epoxy bioadhesives (136 kPa) [28]. In comparison, the cyanoacrylates composing highly adhesive glues, known as Super Glue, have a lap-shear force 100 times higher, of about 13.7 MPa [29].

Interestingly, diverse substrates, non-coated, Poly-L-lysine-coated (PLL-coated), poly-L-lysine–polyethylene glycol-coated (PLL–PEG- coated) and oxygen-activated glass slides present very similar adhesion forces ranging from 184 mN to 229 mN. In these cases, the break most often occurs between the pupal case and the glue, indicating that the bond between the glue and the substrate is stronger than the bond between the glue and the animal. This also suggests that the assay measures the adhesion force between the animal and its bioadhesive and not with its pupariation substrate. This observation can explain the similarity in adhesion forces between different substrates. Similarly, with substrates of increasing roughness, the break usually occurs between the pupal case and the glue, and no significant amelioration of the adhesion is detected [21]. As a control, when a low-stickiness substrate such as polytetrafluoroethylene (PTFE, Teflon) is used as a pupariation substrate,

during the assay the glue completely detaches from the substrate and remains on the pulled pupa, and the average adhesion force is significantly lower (42 mN).

Taken together, these results mean that the bond between the glue and many different types of substrates is stronger than 200 mN, except for Teflon. No effect of humidity, temperature, atmospheric pressure and age of the pupa was found on the adhesion force measure [21], but the ranges in temperature and age of the pupa were small (respectively, 23.5–27.9 °C and 3.5–23 h), so it is possible that more extreme temperatures and older pupae exhibit differences in adhesion forces.

Relatively high variation in adhesion forces was detected between pupae for the same substrate, even within the same strain, ranging for example from 80 mN to 430 mN between *D. melanogaster* individuals for glass slides [27]. This variation cannot be solely attributed to the measurement error, as the universal test machine has an accuracy of ±0.5%. The adhesion assay described here measures the adherence of naturally attached pupae, and it is possible that factors that are not controlled in the experiment greatly influence adhesion, such as animal size, shape, weight, position of the pupa on the substrate or the amount of glue produced. Ideally, it would be good to develop adhesion assays on extracted glue. Unfortunately, there is currently no means to trigger glue expectoration. A recent study [22] that managed to collect glue monitored the larvae under a stereomicroscope until glue expectoration and required action within a few seconds before the glue solidified completely, which is a time-consuming approach.

Comparison of 12 *D. melanogaster* lines from different geographical regions revealed that adhesion can also vary between strains of the same species [27]. Besides *D. melanogaster*, glue adhesion strength has been reported in only three Drosophila species so far [26]. *D. simulans* detach at a similar force (median of 234.2 mN) [26] to *D. melanogaster* (median of 217 mN) [21]. *D. hydei* have the highest force (median of 482.6 mN) and *D. suzukii* the lowest (median of 78.7 mN). Furthermore, the adhesion force correlates with the glue contact area between the pupa and the substrate for these three species [26] but not for *D. melanogaster* [21].

Noticeably, most Drosophila researchers who manipulated pupae in vials know from experience that the glue displays an interesting reversible adhesiveness property. Pupae can be detached from the glass or plastic vial to which they stick by adding a small drop of water, waiting about one minute for the glue to swell in the water and then using a small paintbrush to gently detach the pupa. Such detached pupae can then be placed in another location within the vial. When dried, the glue will strongly adhere again to the tube.

In conclusion, assays have been developed recently to evaluate the force of detachment of naturally glued pupae. These assays will be very useful in future years to assess the range of adhesion forces across various fly species, diverse substrates and various environmental conditions.

5. Production and Expectoration of the Glue

The glue is made of water and several proteins named glue proteins [30]. The production of glue proteins in *D. melanogaster* begins during the second half of the third larval instar with their synthesis in the endoplasmic reticulum, where they are folded and then transported to the Golgi apparatus via the formation of Tango1-mediated rings that act as docking points between the endoplasmic reticulum and Golgi [31]. There, some of the glue proteins are glycosylated, and all the glue proteins are packaged into vesicles, also named granules. As they leave the trans Golgi network, these granules are about 1 µm in diameter, and they will fuse with each other to give large mature granules, about 3 to 8 µm in diameter [32]. Each salivary gland in *D. melanogaster* contains between 2500 and 3000 individual secretory granules [19]. In the granules, three main ultrastructural components are observed: a paracrystalline component made of electron-dense filament bundles, electron-lucent discs and a fine particulate or electron-opaque matrix [19,33]. The formation, composition and properties of these individual components is starting to be studied [33]. Progressively, glue proteins appear to be densely packed and dehydrated in

large vesicles, in a process involving granule acidification, chloride ions, calcium ions and glycosylation [33].

Four to five hours prior to expectoration, a pulse of ecdysone triggers exocytosis, and granules release their content into the salivary gland lumen in an actomyosin-dependent process [34–36]. Once the secretion begins, the paracrystalline structure is lost and the lumen is filled with an amorph secretion [37,38]. A rise in pH and disappearance of calcium ions lead to the unfolding and hydration of the granule contents with water coming from the hemolymph, increasing the total volume [39,40]. At the end of the third instar larval stage, the salivary gland becomes bloated and full of glue.

When a larva finds an appropriate substrate for pupariation, it expectorates the glue, and the content of the lumen of the salivary gland is expelled through the mouth [5]. The process of glue expectoration has been described recently in exquisite detail based on movies of *D. melanogaster* larvae expressing *Sgs3:GFP* fluorescent glue [41]. At the end of the larval stage, the larva everts its pair of anterior spiracles, which are respiratory openings through which air will pass during metamorphosis, and it moves less and less. The animal also acquires a characteristic barrel shape through increasingly strong whole-body contractions and then enters a tetanic contraction phase where ventral anterior segments contract and slightly arch the anterior half of the larva for 17–70 s. Then, an anterior peristaltic wave propagates from segment T2 to A2 in approximately 3 s, further squeezing the anterior segments. A few milliseconds later, the glue is expelled from the lumen of the salivary gland to the exterior of the animal. While the glue is being released, a series of coordinated peristaltic movements propagate forwards and backwards, starting from segment A2, and lead to the spreading of the glue throughout the whole body. Furthermore, during expectoration, the animal usually moves forward about half of its length, reaching its final pupariation site, where it typically waves its anterior end left and right a few times. From the tetanus phase to the head waving, about 60–70 s have elapsed. Then, occasional whole-body contractions occur for about 50 min: they help remodel the puparium shape and lead to the formation of the operculum (the part of the pupal case that will be opened up by the adult fly when it emerges at the end of metamorphosis), and the cuticle starts to harden. The same suite of behavioral events accompanying glue expectoration was observed in *D. virilis* [41], which diverged from *D. melanogaster* about 45 million years ago [42]. After expectoration, the glue is liquid and it hardens in a few seconds, depending on air humidity, and becomes completely dry and solid after 3–5 min [4,22]. The movements of the larva during expectoration allow the glue to completely wet the body and increase contact with the animal surface topography by flowing into the folds and crevices of the cuticle, thus maximizing adhesiveness between the animal and the substrate. To our knowledge, the behavior of larvae that stick themselves to already attached pupae has not been described.

In summary, the stickiness of the pupae to their substrate results not only from the biochemical properties of the glue but also from the behavior of the larvae, including its body shape remodeling and its spreading of the glue via peristaltic movements. The glue proteins display remarkable properties, allowing them to be packed and dried into granules, fluidified in the gland lumen and then solidified in contact with air.

6. Identification of the Glue Genes in *D. melanogaster*

The glue of *D. melanogaster* was first isolated in 1948 from the salivary glands by placing the glands into an ethanol solution and then dissecting the solid plug of precipitated glue [43]. However, it was only in 1975 that the composition of the secretion was studied [44]. Using acid–urea gel electrophoresis, it was found that the glue separates into several bands, corresponding to different proteins. In *D. melanogaster*, bands were labeled from one to five according to their increasing electrophoretic mobility, thus from large to small size [44]. In *D. melanogaster*, eight glue genes were found in total (Table 1). The genes responsible for each protein band, named salivary gland secretion (*Sgs*) genes, were cytogenetically mapped based on polytene chromosomes and polymorphism be-

tween several *D. melanogaster* strains, with presence/absence of certain bands on glue gel electrophoresis correlating with the presence/absence of certain puffs on polytene chromosomes. Puffs are enlarged regions on polytene chromosomes that form swellings where active transcription takes place [45]. Then, in the 1980s, thanks to DNA cloning and restriction mapping, the glue genes were among the first developmentally regulated genes whose DNA gene sequence was identified [46]. *Sgs1*, *Sgs3*, *Sgs4* and *Sgs5* gene sequences were thus found at four distinct chromosomal locations (Table 1). Band 2 was considered as a contamination, and no further analysis of this band was performed [47]. *Sgs5bis*, *Sgs7* and *Sgs8* and *Eig71Ee* were found in later studies and were named without correlation to the electrophoretic mobility of their corresponding proteins (Table 1). Another glue gene, named *Sgs6*, has not been identified yet. The corresponding protein is present in only some *D. melanogaster* strains such as Canton S [48,49], and its nucleotide sequence located in region 71C3-4 is still unknown today [49,50]. *Eig71Ee* was first studied for its ecdysone-induced gene expression [51] and was later found to be expressed in the salivary glands at the late third instar larvae stage [52]. The Eig71Ee protein is O-glycosylated [53], contains internal repeats similar to Sgs3 and Sgs4 and is rich in cysteines (8%) like Sgs4 [52]. *Eig71Ee* is also expressed in the hemocytes and gut, where it is involved in immunity and clotting [53].

Table 1. List of the main glue genes of *D. melanogaster* and their characteristics. See FlyBase (http://flybase.org, accessed on 8 July 2022) for further information.

Sgs Gene Name	Band	Chromosome	Cytogenetic Map	Other Gene Names	Number of Amino Acids	Amino Acid Composition and Glycosylation State	Reference
Sgs1	1	2L	25B4	*CG3047*	1286	Presence of repeats PTTTTPR/STTTTSTSR. Rich in cysteines, prolines, serines and threonines. Glycosylated.	[54]
Sgs3	3	3L	68C11	*CG11720*	307	Presence of repeats KPTTT. Rich in cysteines, prolines, serines and threonines. Glycosylated.	[55]
Sgs4	4	X	3C11-12	*CG12181*	297	Presence of repeats. Rich in cysteines, prolines, serines and threonines. Glycosylated.	[46]
Sgs5	5	3R	90B3-8	*CG7596*	163	No repeat. Rich in cysteines, prolines and serines.	[56]
Sgs5bis	-	3R	90B5	*CG7587*	142	No repeat. Rich in cysteines and prolines.	[50]
Sgs7	-	3L	68C11	*CG18087*	74	No repeat. Rich in cysteines.	[47]
Sgs8	-	3L	68C11	*CG6132*	74	No repeat. Rich in cysteines.	[47]
Eig71Ee	-	3L	71E5	*CG7604 VII I71–7 gp150*	393	Presence of repeats CTCTESTT/(R/K)TNPT. Rich in cysteines, prolines, serines and threonines. Glycosylated.	[52]

In total, the sequence of eight glue genes has been described in *D. melanogaster*. These eight genes are among the ten most highly expressed genes in salivary glands at the wandering third instar larval stage [27]. It is possible that *D. melanogaster* glue contains other proteins that have not been characterized yet. Highly expressed genes in wandering third instar salivary glands include genes involved in transcription and translation, as well as several small uncharacterized genes encoding for secreted peptides with the same tissue-specific, stage-specific gene expression as the glue genes [27]. These genes may encode for additional components of the glue.

7. Characteristics and Functions of the Glue Proteins in *D. melanogaster*

The eight glue proteins identified in *D. melanogaster* present a signal peptide, so that the resulting proteins are all destined to the secretory pathway. We can distinguish two groups of glue proteins: Sgs1, Sgs3, Sgs4 and Eig71Ee are relatively long proteins containing multiple cysteines and amino acid repeats that are rich in prolines, serines and threonines, whereas Sgs5, Sgs5bis, Sgs7 and Sgs8 are relatively short proteins that do not have internal repeats and are rich in cysteines [19,50]. The relative amount of each protein within the glue is not known [19,48], and their respective roles in the various steps of glue production (granule maturation, hydration in the salivary gland lumen, lubrication during expectoration, glue cementing, glue adhesion) have so far mostly been inferred based on their amino acid sequence.

In the first group of proteins, repeats containing serines and threonines are subject to O-glycosylation and are characteristic of secreted mucins [19,50]. Mucins are highly glycosylated proteins present in animal mucus that protect the epithelia from physical damage or pathogens [57]. Glycosylation makes the molecules very hydrophilic, which enhances solubility, adhesion and is probably important for rehydration of the secreted content of the granules in the salivary gland lumen during glue production [30,40,58]. Sgs3 is O-glycosylated in the T-rich region and the PTTTK repetitive domain, and this glycosylation is in part accomplished by PGANT9A and PGANT9B enzymes [32]. The exact nature of the sugars covalently attached to the serines and threonines of the glue proteins has not been characterized. Computer predictions of protein structure reveal that the repeated regions are intrinsically disordered: they lack α-helices and β-sheets and do not have a fixed three-dimensional structure [19,50]. They are enriched in prolines like other intrinsically disordered regions [59], and they may form long threads [50]. The number of repeats and total protein length vary across *D. melanogaster* strains [55,60]. Overall, the long, disordered and highly glycosylated glue proteins of the first group may help to increase solubility at high concentrations, allow the rapid rehydration of the vesicles content after exocytosis in the salivary gland lumen, enhance fluidity of the mixture during expectoration and improve adhesive properties of the glue once released [19,33].

RNAi-mediated reduction of O-glycosylation leads to more tightly packed electron-dense fibers within the salivary gland granules, suggesting that adjacent fibers are repelled via their negatively charged sugars [33]. In RNAi loss-of-function mutants of *Sgs1* and *Sgs3*, the electron-lucent discs and the filament bundles are, respectively, gone. This shows that these glue proteins are involved in the intense packaging of molecules into the vesicles, and the authors propose that Sgs1 forms the disc structures while Sgs3 adopts a bundled filament structure.

Glue proteins of the second group contain α-helices and β-sheets. They may be involved in the nucleation of the densification process in Golgi vesicles [19]. The multiple cysteines present in these glue proteins and in those of the first group can allow the formation of disulfide bridges intramolecularly to build up the three-dimensional structure of each protein and also between glue proteins, for example, by cysteine oxidation when the glue comes in contact with air, to create a complex fibrous macromolecular material [19,50].

Except *Sgs4* and *Eig71Ee*, all the glue genes are only expressed in the salivary glands and at the third instar larval stage [35,61,62], suggesting that their function is restricted to the making up of the glue. *Eig71Ee* is also involved in immunity and clotting in the

hemocytes and the gut [53], while *Sgs4* is expressed in proventriculus and salivary glands from late second to late third instar larval stages [63], but its exact role in these tissues has not been characterized. The protein sequences of Sgs4 and Eig71Ee may thus also be subjected to other functional constraints.

In summary, according to their amino acid sequences, the eight glue proteins of *D. melanogaster* appear to display remarkable biochemical properties. Further work is needed to decipher the respective roles of the various molecular components of the glue in glue production, hardening, adhesion strength and adhesion reversibility.

8. Glue Genes and Proteins in other Drosophila Species

Besides *D. melanogaster*, fly glue has been mostly studied in *D. virilis*, and it appears to be composed of fewer proteins than in *D. melanogaster* (Table 2). Compared to the five bands present on electrophoresis gel in *D. melanogaster*, only three bands are found [64]. The first band protein is encoded by the gene *Lgp1* [64], an ortholog of *D. melanogaster Sgs3* also named Sgs3a in a more recent study [50]. In the *D. virilis* genome, it is adjacent to another glue gene, named *Lgp3* or *Sgs3b* [50,58]. *Sgs3a* and *Sgs3b* result from a recent duplication in the *D. virilis* lineage [50]. Lgp1/Sgs3a and Lgp3/Sgs3b are major components of the glue and together represent 90% of its content [65]. The remaining 10 % correspond to a weakly glycosylated 15-kDa protein named Lgp2 whose sequence was not characterized at the time [66]. By BLAST, only three glue genes were identified recently in the genome of *D. virilis* [50], *Sgs3a*, *Sgs3b and Sgs5bis. Sgs5bis* protein has no internal repeat and is expected to be 15.9 kDa, so we suggest that Lgp2 and Sgs5bis are the same.

Besides *D. melanogaster* and *D. virilis*, glue protein composition has been examined in *D. gibberosa* and in seven species from the *D. nasuta* group, which all diverged about 45 million years ago from *D. melanogaster* [42] (Table 2): *D. n. nasuta*, *D. n. albomicans*, *D. n. kepulauana*, *D. kohkoa*, *D. s. albostrigata*, *D. s. bilimbata* and *D. s. sulfurigaster*. Using gel electrophoresis, multiple protein bands were found, some of them being glycosylated [67], but the corresponding gene sequences were not characterized. The number of bands ranged from nine in *D. n. nasuta* up to seventeen in *D. gibberosa* (Table 2) [68]. Intraspecific polymorphism in the number of bands was observed in *D. nasuta nasuta* and *D. s. neonasuta* collected in the wild [69].

The level of glycosylation of the glue proteins appears to vary between Drosophila species. For example, *D. virilis* glue is rich in different sugars such as glucose, mannose and galactose [70], whereas the one from species of the *D. suzukii* subgroup, *D. suzukii*, *D. rajasekari* and *D. lucipennis*, contains no or low amounts of glycosylation [71].

In two recent studies, the glue genes from 22 Drosophila species spanning from the *D. melanogaster* subgroup to *D. virilis* and *D. mojavensis*, which diverged about 45 million years ago from *D. melanogaster* [42], were uncovered by BLAST based on sequence similarity with *D. melanogaster* glue genes [19,50]. Among them, *D. virilis* has the lowest number of glue genes (only three) while *D. santomea* and *D. yakuba* have the highest (nine in total) [50]. Interestingly, each species has at least one representative for each gene group: one encoding a long protein rich in cysteines, prolines, serines and threonines and containing repeats and one encoding a short protein rich in cysteines (Table 2).

Table 2. List of the glue genes and glue protein bands identified in Drosophila species. *Sgs* genes written in bold correspond to proteins with internal repeats rich in serines, threonines and prolines. * An updated genome assembly [72] shows that *D. suzukii* actually contains one *Sgs3* gene and only one copy of *Sgs7* (data not shown). Nd: not determined.

Species	Number of Bands	Glue Gene Sequences Identified	Reference
D. simulans *D. sechellia*	nd	***Sgs1***; ***Sgs3***; ***Sgs4***; *Sgs5; Sgs7; Sgs8;* ***Eig71Ee***	[50]
D. mauritiana	nd	***Sgs1***; ***Sgs3***; ***Sgs4***; ***Sgs5***; *Sgs7; Eig71Ee*	[50]
D. santomea *D. yakuba*	nd	***Sgs1***; ***Sgs3***; ***Sgs4***; ***Sgs5***; *Sgs5bis; Sgs7; Sgs7bis Sgs8;* ***Eig71Ee***	[50]
D. erecta	nd	***Sgs3***; ***Sgs4***; *Sgs5bis; Sgs7; Sgs8;* ***Eig71Ee***	[50]
D. eugracilis	nd	***Sgs1***; ***Sgs3***; ***Sgs3bis***; ***Sgs5***; *Sgs5bis; Sgs7; Sgs8;* ***Eig71Ee***	[50]
D.takahashii	nd	***Sgs1***; ***Sgs3***; ***Sgs5***; *Sgs5bis; Sgs7; Sgs8;* ***Eig71Ee***	[50]
D. suzukii *	nd	***Sgs1***; *Sgs4;* ***Sgs5***; *Sgs5bis;* 4 copies of *Sgs7; Sgs8; Eig71Ee*	[50]
D. biarmipes	nd	***Sgs1***; ***Sgs3***; ***Sgs3bis***; ***Sgs5***; *Sgs5bis; Sgs7; Sgs8; Eig71Ee*	[50]
D.elegans	nd	***Sgs1***; ***Sgs3a***; ***Sgs3b***; ***Sgs3c***; ***Sgs5***	[50]
D. rhopaloa	nd	***Sgs1***; ***Sgs3a***; ***Sgs3b***; ***Sgs3c***; ***Sgs3d***; ***Sgs5***	[50]
D. ficusphila	nd	***Sgs1***; ***Sgs3a***; ***Sgs3b***; ***Sgs3c***; ***Sgs5***; *Sgs5bis;* ***Eig71Ee***	[50]
D. kikkawai *D. ananassae*	nd	***Sgs3a***; ***Sgs3b***; ***Sgs5***; *Sgs5bis*	[50]
D. bipectinata	nd	***Sgs3a***; ***Sgs3b***; *Sgs7; Sgs8; Sgs5; Sgs5bis;* ***Eig71Ee***	[50]
D. pseudoobscura	nd	***Sgs3a***; ***Sgs3b***; ***Sgs3c***; *Sgs5bis*	[50]
D. willistoni	nd	***Sgs3a***; ***Sgs3b***; *Sgs7a; Sgs7b*	[50]
D. virilis	3	***Sgs3a*** (or *Lgp1*); ***Sgs3b*** (or *Lgp3*) *Sgs5bis* (or *Lgp2*)	[50]
D. mojavensis	nd	*Sgs4; Sgs5; Sgs7*	[19]
D. persimilis	nd	*Sgs5; Sgs7; Sgs8*	[19]
D. n. nasuta	9	nd	[68]
D. n. albomicans	10	nd	[68]
D. n. kepulauana	12	nd	[68]
D. kohkoa	10	nd	[68]
D. s. albostrigata	12	nd	[68]
D. s. bilimbata	14	nd	[68]
D. s. sulfurigaster	13	nd	[68]
D. gibberosa	17	nd	[73]

9. Glue Gene Evolution

The various glue genes that have been identified in Drosophila species can be grouped into three gene families based on their sequence similarities: one composed of *Sgs5* and *Sgs5bis*, one with *Sgs4* and one with the remaining genes (*Sgs1*, *Sgs3*, *Sgs7*, *Sgs8* and *Eig71Ee*) [50]. Genes of the last group show C-terminal and N-terminal sequence similarities

and have an intron at the same position and phase, with the codon disrupted by the intron encoding for an alanine or valine at position 10 [50,58].

Sgs sequences can differ in length between species. Overall, the glue genes with internal repeats (*Sgs1*, *Sgs3*, *Sgs4* and *Eig71Ee)* vary much more in length than the other glue genes, due to variation in the number of repeats and in the size of the repeated motif [50]. The number of repeats can vary rapidly. For example, *Sgs1* contains about 13 repeated motifs in *D. mauritiana* and 40 in *D. simulans* [50], which diverged some 300 000 years ago [42]. Furthermore, some genes can have internal repeats while their paralogs do not, suggesting that a glue gene devoid of repeats can acquire repeated sequences and/or that a glue gene can lose all of its repeats during evolution. For example, *Sgs5* does not appear to contain repeats in *D. melanogaster* but does in *D. simulans* [50]. In *D. melanogaster*, three glue genes are located at the same chromosomal location, and they share sequence identity, suggesting that they come from ancient duplications: one, *Sgs3*, contains repeats, whereas the other two, *Sgs7* and *Sgs8*, do not. This suggests that the presence/absence of repeats can change across evolution. As described above, the presence/absence of repeats is associated with two distinct glue functions: the multiple prolines, serines and threonines present in repeats appear to generate long glycosylated filaments, whereas the shorter proteins devoid of repeats may contribute to the scaffolding of the glue via disulfide bonds. Two types of functional glue proteins may thus be formed within the same gene family.

Noticeably, the *Sgs1-3-7-8-Eig71Ee* gene family has experienced a higher rate of gene losses and gene duplications than other gene families present in the *Drosophila* genomes [50]. The rapid evolution of glue gene sequences, in terms of gene number, number of repeats and repeat motifs, may be related to the rapid adaptation of glue adhesiveness to various environmental conditions.

10. Role of Drosophila Glue in Natural Environments

Drosophila salivary glands form during embryogenesis, and it is unclear whether they produce digestive enzymes during early larval stages [74–76]. In any case, salivary glands appear to be dispensable during larval life since individuals carrying salivary glands as closed sacs devoid of ducts due to mutation in the *eye gone* gene survive until the pupal stage and then die as late pupae or adults [77]. During the prepupal stage, the salivary glands produce a massive secretion distinct from the glue into the peri-exuvial cavity that lies between the metamorphosing pupa and its pupal case [78]. This secretion contains immune-competent and defense-response proteins and acts as a protective barrier against microbial infections.

The main function of salivary glands at the end of the third instar larval stage is the production of a glue that will affix the animal to a substrate [4]. Pupal adhesion can have several functions. First, it can allow the organisms to remain in a favorable environment (in terms of temperature, humidity, background color, etc.), resisting mechanical forces (wind, rain, other animals) that may displace pupae into adverse surroundings. Second, it may help the adult to emerge from the pupal case, although this possibility has not been examined experimentally as far as we know. Third, it can protect the pupae from predation. Recent work from our group showed that in a forest near Paris, *D. simulans* pupae naturally attached to a substrate are taken away less frequently than manually detached pupae [26]. Furthermore, experiments in the laboratory showed that attached pupae are predated less efficiently by ants, which take more time to consume them onsite and are not able to carry them back to the nest [26].

The glue covers the surface of the animal (Figure 2D) for several days, until the adult emerges from the pupal case. So, it is possible that this material has other functions besides stickiness. Pupae are vulnerable not only to predation but also to parasitism [79], fungal or bacterial infection and desiccation. The decaying fruits, which represent a major pupariation site for many Drosophila species [8,10,14], are especially exposed to desiccation and are rich in fungi and bacteria. Besides its adhesive properties, it is thus possible that the glue may have other functions, none of which have been investigated so far, as far as we

know. For example, it may repel predators and parasitoids or make the pupae undetectable to them. Del Pino et al. proposed that components of the salivary gland secretion may act as pheromones [23]. The glue may also act as a preservative and avoid fungal or bacterial infections. The glycoproteins that make up the glue belong to the mucin family, and mucins are known to have antimicrobial properties [39]. In particular, the glue protein Eig71Ee, also named gp150, which is present in hemolymph, lies in structures that entrap bacteria [53]. Scanning electron microscopy showed that yeast-like organisms and coliform bacteria can be found and efficiently trapped within the glue of *D. melanogaster* individuals raised in the laboratory [22].

Further work on the properties and function of Drosophila glue would be extremely useful to get a better idea of the possible applications of future biomimetic adhesives.

11. Perspectives for Future Applications

Bioadhesives inspired from nature may be compatible with the human body and biodegradable and thus offer attractive properties compared to synthetic ones. Furthermore, they may display antifouling or antimicrobial properties. Recent measurements of Drosophila glue adhesiveness showed that it is equivalent to strong commercial adhesive tapes [21]. Indeed, we noticed in our adhesion assays that commercial tapes with low adherence led to detachment of the pupa from the tape and not from its substrate. Research on Drosophila glue may help in the future to develop new bioadhesives for dry environments, on polarizable surfaces.

However, several difficulties remain. First, the volume of glue produced by each larva is relatively small, making it difficult to study the physical and biochemical characteristics of this glue. Second, there is no available method to trigger glue expectoration from the larva. When larvae are manipulated with forceps right before glue expectoration, they can revert the pupariation process, retract their anterior spiracles and start moving again to find another pupariation site [5]. Third, Drosophila glue is produced through a complex granule maturation process, involving pH change, calcium ions and chloride ions [33]. Such a maturation process may be difficult to reproduce in vivo, unless large progress is made in organoid and organs-on-chips research [80,81]. Alternatively, small molecular elements of the glue such as modified amino acids may be found to be key to the adhesion process, and new adhesives may be created by synthesizing polymers containing such molecules. For example, the catecholic amino acid 3,4 dihydroxy phenylalanine (DOPA), which is abundant in mussel adhesive proteins, plays an essential role in strong underwater adhesion, and polyethylene glycol polymers grafted with DOPA are being developed as mussel-inspired tissue adhesives [2].

In future years, the powerful genetic tools of *D. melanogaster* will definitely facilitate the study of the roles of the different players in the formation of the glue and its adhesive properties: glue proteins, glycosylation, pH, calcium ions, chloride ions, etc. The diversity of glues produced by various Drosophila species and adapted to various environments represents a promising reservoir for bioinspiration.

12. Conclusions

Research on the biochemical and physical properties of Drosophila glue is just starting. This is an exciting emerging field where multiple avenues of research are available to learn more about the fascinating biophysical attributes of Drosophila glue, including adhesive properties and antimicrobial activities, as well as its elaborate biosynthesis and secretion. Furthermore, understanding the specificities of the diverse glues produced by Drosophila strains and species in relation to their environments will provide insight into the development of Drosophila-inspired adhesives.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/insects13080734/s1, Table S1. List of papers retrieved from PubMed using a search for "Drosophila glue". File S1. R Script used to make Figure 1.

Author Contributions: Conceptualization, M.M. and V.C.-O.; writing—original draft preparation, M.M.; writing—review and editing, V.C.-O. and M.M.; visualization, V.C.-O.; supervision, V.C.-O.; project administration, V.C.-O.; funding acquisition, V.C.-O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CNRS as part of the MITI interdisciplinary action, "Défi Adaptation du vivant à son environnement" and from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 grant agreement no. 337579) to V.C.-O., M.M. was supported by a PhD fellowship from "Ministère de l'Education Nationale, de la Recherche et de la Technologie" (MENRT) obtained from the BioSPC doctoral school.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Christen Mirth and Lynn M. Riddiford for inviting us to write this review. We acknowledge the Cornell University Drosophila Species Stock Center for fly stocks. We thank Flora Borne, François Graner and Kelly Hagen for comments on this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

A2	Abdominal segment 2
DOPA	3,4 dihydroxy phenylalanine
Sgs gene	Salivary gland secretion gene
T2	Thoracic segment 2

References

1. Bianco-Peled, H.; Davidovich-Pinhas, M. *Bioadhesion and Biomimetics: From Nature to Applications*; CRC Press: Boca Raton, FL, USA, 2015; ISBN 978-981-4463-99-7.
2. Pandey, N.; Soto-Garcia, L.F.; Liao, J.; Zimmern, P.; Nguyen, K.T.; Hong, Y. Mussel-Inspired Bioadhesives in Healthcare: Design Parameters, Current Trends, and Future Perspectives. *Biomater. Sci.* **2020**, *8*, 1240–1255. [CrossRef] [PubMed]
3. Kord Forooshani, P.; Lee, B.P. Recent Approaches in Designing Bioadhesive Materials Inspired by Mussel Adhesive Protein. *J. Polym. Sci. Part Polym. Chem.* **2017**, *55*, 9–33. [CrossRef]
4. Fraenkel, G.; Brookes, V.J. The Process by which the Puparia of Many Species of Flies Become Fixed to a Substrate. *Biol. Bull.* **1953**, *105*, 442–449. [CrossRef]
5. Ashburner, M. *Drosophila: A Laboratory Handbook*, 2nd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2004.
6. O'Grady, P.M.; DeSalle, R. Phylogeny of the Genus Drosophila. *Genetics* **2018**, *209*, 1–25. [CrossRef] [PubMed]
7. Markow, T.; O'Grady, P.M. Evolutionary Genetics of Reproductive Behavior in Drosophila: Connecting the Dots. *Annu. Rev. Genet.* **2005**, *39*, 263–291. [CrossRef]
8. Godoy-Herrera, R.; Cifuentes, L.; Díaz de Arcaya, M.F.; Fernández, M.; Fuentes, M.; Reyes, I.; Valderrama, C. The Behaviour of Drosophila Melanogaster Larvae during Pupation. *Anim. Behav.* **1989**, *37*, 820–829. [CrossRef]
9. Godoy-Herrera, R.; Silva-Cuadra, J.L. The Behavior of Sympatric Chilean Populations of Drosophila Larvae during Pupation. *Genet. Mol. Biol.* **1998**, *21*, 31–39. [CrossRef]
10. Vandal, N.B.; Siddalingamurthy, G.S.; Shivanna, N. Larval Pupation Site Preference on Fruit in Different Species of Drosophila. *Entomol. Res.* **2008**, *38*, 188–194. [CrossRef]
11. Carson, H.L. The Association between Drosophila Carcinophila Wheeler and Its Host, the Land Crab *Gecarcinus ruricola* (L.). *Am. Midl. Nat.* **1967**, *78*, 324–343. [CrossRef]
12. Carson, H.L.; Wheeler, M.R. Drosophila Endobranchia, a New Drosophilid1 Associated with Land Crabs in the West Indies. *Ann. Entomol. Soc. Am.* **1968**, *61*, 675–678. [CrossRef]
13. Sokolowski, M.B. Genetics and Ecology of Drosophila Melanogaster Larval Foraging and Pupation Behaviour. *J. Insect Physiol.* **1985**, *31*, 857–864. [CrossRef]
14. Beltramí, M.; Medina-Muñoz, M.C.; Arce, D.; Godoy-Herrera, R. Drosophila Pupation Behavior in the Wild. *Evol. Ecol.* **2010**, *24*, 347–358. [CrossRef]
15. Beltramí, M.; Medina-Muñoz, M.C.; Pino, F.D.; Ferveur, J.-F.; Godoy-Herrera, R. Chemical Cues Influence Pupation Behavior of Drosophila Simulans and Drosophila Buzzatii in Nature and in the Laboratory. *PLoS ONE* **2012**, *7*, e39393. [CrossRef] [PubMed]
16. Grossfield, J. *Non-Sexual Behavior of Drosophila*; Academic Press: London, UK, 1978; Volume 2, pp. 1–126.

17. Woltz, J.M.; Lee, J.C. Pupation Behavior and Larval and Pupal Biocontrol of Drosophila Suzukii in the Field. *Biol. Control* **2017**, *110*, 62–69. [CrossRef]
18. Crosby, M.A.; Meyerowitz, E.M. Drosophila Glue Gene Sgs-3: Sequences Required for Puffing and Transcriptional Regulation. *Dev. Biol.* **1986**, *118*, 593–607. [CrossRef]
19. Farkaš, R. The Complex Secretions of the Salivary Glands of Drosophila Melanogaster, A Model System. In *Extracellular Composite Matrices in Arthropods*; Cohen, E., Moussian, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 557–600. ISBN 978-3-319-40738-8.
20. Meyerowitz, E.M.; Crosby, M.A.; Garfinkel, M.D.; Martin, C.H.; Mathers, P.H.; Vijay Raghavan, K. The 68C Glue Puff of Drosophila. *Cold Spring Harb. Symp. Quant. Biol.* **1985**, *50*, 347–353. [CrossRef]
21. Borne, F.; Kovalev, A.; Gorb, S.; Courtier-Orgogozo, V. The Glue Produced by *Drosophila melanogaster* for Pupa Adhesion Is Universal. *J. Exp. Biol.* **2020**, *223*, jeb220608. [CrossRef]
22. Beňová-Liszeková, D.; Beňo, M.; Farkaš, R. A Protocol for Processing the Delicate Larval and Prepupal Salivary Glands of Drosophila for Scanning Electron Microscopy. *Microsc. Res. Tech.* **2019**, *82*, 1145–1156. [CrossRef]
23. Del Pino, F.; Jara, C.; Pino, L.; Godoy-Herrera, R. The Neuro-Ecology of Drosophila Pupation Behavior. *PLoS ONE* **2014**, *9*, e102159. [CrossRef]
24. Medina-Muñoz, M.C.; Godoy-Herrera, R. Dispersal and Prepupation Behavior of Chilean Sympatric Drosophila Species That Breed in the Same Site in Nature. *Behav. Ecol.* **2005**, *16*, 316–322. [CrossRef]
25. Ringo, J.; Dowse, H. Pupation Site Selection in Four Drosophilid Species: Aggregation and Contact. *J. Insect Behav.* **2012**, *25*, 578–589. [CrossRef]
26. Borne, F.; Prigent, S.R.; Molet, M.; Courtier-Orgogozo, V. *Drosophila* Glue Protects from Predation. *Proc. R. Soc. B Biol. Sci.* **2021**, *288*, 20210088. [CrossRef] [PubMed]
27. Borne, F.; Kulathinal, R.J.; Courtier-Orgogozo, V. Glue Genes Are Subjected to Diverse Selective Forces during Drosophila Development. *Genome Biol. Evol.* **2021**, *13*, evab248. [CrossRef] [PubMed]
28. Du, D.; Chen, X.; Shi, C.; Zhang, Z.; Shi, D.; Kaneko, D.; Kaneko, T.; Hua, Z. Mussel-Inspired Epoxy Bioadhesive with Enhanced Interfacial Interactions for Wound Repair. *Acta Biomater.* **2021**, *136*, 223–232. [CrossRef]
29. Ebnesajjad, S.; Landrock, A.H. Characteristics of Adhesive Materials. In *Adhesives Technology Handbook*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 84–159. ISBN 978-0-323-35595-7.
30. Kress, H. Temporal Relationships between Leaving Food, Ecdysone Release, Mucoprotein Extrusion and Puparium Formation in Drosophila Virilis. *J. Insect Physiol.* **1974**, *20*, 1041–1055. [CrossRef]
31. Reynolds, H.M.; Zhang, L.; Tran, D.T.; Ten Hagen, K.G. Tango1 Coordinates the Formation of Endoplasmic Reticulum/Golgi Docking Sites to Mediate Secretory Granule Formation. *J. Biol. Chem.* **2019**, *294*, 19498–19510. [CrossRef]
32. Ji, S.; Samara, N.L.; Revoredo, L.; Zhang, L.; Tran, D.T.; Muirhead, K.; Tabak, L.A.; Ten Hagen, K.G. A Molecular Switch Orchestrates Enzyme Specificity and Secretory Granule Morphology. *Nat. Commun.* **2018**, *9*, 3508. [CrossRef]
33. Syed, Z.A.; Zhang, L.; Tran, D.T.; Bleck, C.K.E.; Hagen, K.G.T. Orchestrated Restructuring Events During Secretory Granule Maturation Mediate Intragranular Cargo Segregation. *bioRxiv* **2022**. bioRxiv:2021.08.16.456250.
34. Rousso, T.; Schejter, E.D.; Shilo, B.-Z. Orchestrated Content Release from Drosophila Glue-Protein Vesicles by a Contractile Actomyosin Network. *Nat. Cell Biol.* **2016**, *18*, 181–190. [CrossRef]
35. Duan, J.; Zhao, Y.; Li, H.; Habernig, L.; Gordon, M.D.; Miao, X.; Engström, Y.; Büttner, S. Bab2 Functions as an Ecdysone-Responsive Transcriptional Repressor during Drosophila Development. *Cell Rep.* **2020**, *32*, 107972. [CrossRef]
36. Tran, D.T.; Masedunskas, A.; Weigert, R.; Ten Hagen, K.G. Arp2/3-Mediated F-Actin Formation Controls Regulated Exocytosis in Vivo. *Nat. Commun.* **2015**, *6*, 10098. [CrossRef] [PubMed]
37. von Gaudecker, B. Der Strukturwandel der larvalen Speicheldrüse vonDrosophila melanogaster. *Z. Für Zellforsch. Mikrosk. Anat.* **1972**, *127*, 50–86. [CrossRef]
38. Farkaš, R.; Šuťáková, G. Ultrastructural Changes of Drosophila Larval and Prepupal Salivary Glands Cultured in Vitro with Ecdysone. *Vitro Cell. Dev. Biol.—Anim.* **1998**, *34*, 813–823. [CrossRef] [PubMed]
39. Syed, Z.A.; Zhang, L.; Ten Hagen, K.G. In Vivo Models of Mucin Biosynthesis and Function. *Adv. Drug Deliv. Rev.* **2022**, *184*, 114182. [CrossRef]
40. Lane, N.J.; Carter, Y.R.; Ashburner, M. Puffs and Salivary Gland Function: The Fine Structure of the Larval and Prepupal Salivary Glands OfDrosophila Melanogaster. *Wilhelm Roux Arch. Für Entwickl. Org.* **1972**, *169*, 216–238. [CrossRef]
41. Heredia, F.; Volonté, Y.; Pereirinha, J.; Fernandez-Acosta, M.; Casimiro, A.P.; Belém, C.G.; Viegas, F.; Tanaka, K.; Menezes, J.; Arana, M.; et al. The Steroid-Hormone Ecdysone Coordinates Parallel Pupariation Neuromotor and Morphogenetic Subprograms Via Epider-Mis-To-Neuron Dilp8-Lgr3 Signal Induction. *Nat. Commun.* **2021**, *12*, 1–20. [CrossRef]
42. Kumar, S.; Stecher, G.; Suleski, M.; Hedges, S.B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* **2017**, *34*, 1812–1819. [CrossRef]
43. Kodani, M. The Protein of the Salivary Gland Secretion in Drosophila. *Proc. Natl. Acad. Sci. USA* **1948**, *34*, 131–135. [CrossRef]
44. Korge, G. Chromosome Puff Activity and Protein Synthesis in Larval Salivary Glands of Drosophila Melanogaster. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4550–4554. [CrossRef]
45. Russell, S.; Ashburner, M. Ecdysone-Regulated Chromosome Puffing in Drosophila Melanogaster. In *Metamorphosis*; Elsevier: Amsterdam, The Netherlands, 1996; pp. 109–144.

46. Muskavitch, M.A.; Hogness, D.S. Molecular Analysis of a Gene in a Developmentally Regulated Puff of Drosophila Melanogaster. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 7362–7366. [CrossRef]
47. Meyerowitz, E.M.; Bond, M.W.; Crowley, T.E. The Structural Genes for Three Drosophila Glue Proteins Reside at a Single Polytene Chromosome Puff Locus. *Mol. Cell Biol.* **1983**, *3*, 12.
48. Velissariou, V.; Ashburner, M. Cytogenetic and Genetic Mapping of a Salivary Gland Secretion Protein in Drosophila Melanogaster. *Chromosoma* **1981**, *84*, 173–185. [CrossRef] [PubMed]
49. Akam, M.E.; Roberts, D.B.; Richards, G.P.; Ashburner, M. Drosophila: The Genetics of Two Major Larval Proteins. *Cell* **1978**, *13*, 215–225. [CrossRef]
50. Da Lage, J.-L.; Thomas, G.W.C.; Bonneau, M.; Courtier-Orgogozo, V. Evolution of Salivary Glue Genes in Drosophila Species. *BMC Evol. Biol.* **2019**, *19*, 36. [CrossRef] [PubMed]
51. Restifo, L.L.; Guild, G.M. An Ecdysterone-Responsive Puff Site in Drosophila Contains a Cluster of Seven Differentially Regulated Genes. *J. Mol. Biol.* **1986**, *188*, 517–528. [CrossRef]
52. Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M. Molecular Characterization of the 71E Late Puff in Drosophila Melanogaster Reveals a Family of Novel Genes. *J. Mol. Biol.* **1996**, *255*, 387–400. [CrossRef]
53. Korayem, A.M.; Fabbri, M.; Takahashi, K.; Scherfer, C.; Lindgren, M.; Schmidt, O.; Ueda, R.; Dushay, M.S.; Theopold, U. A Drosophila Salivary Gland Mucin Is Also Expressed in Immune Tissues: Evidence for a Function in Coagulation and the Entrapment of Bacteria. *Insect Biochem. Mol. Biol.* **2004**, *34*, 1297–1304. [CrossRef]
54. Roth, G.E.; Wattler, S.; Bornschein, H.; Lehmann, M.; Korge, G. Structure and Regulation of the Salivary Gland Secretion Protein Gene Sgs-1 of Drosophila Melanogaster. *Genetics* **1999**, *153*, 753–762. [CrossRef]
55. Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M. DNA Sequences, Gene Regulation and Modular Protein Evolution in the Drosophila 68C Glue Gene Cluster. *J. Mol. Biol.* **1983**, *168*, 765–789. [CrossRef]
56. Shore, E.M.; Guild, G.M. Larval Salivary Gland Secretion Proteins in Drosophila Structural Analysis of the Sgs-5 Gene. *J. Mol. Biol.* **1986**, *190*, 149–158. [CrossRef]
57. Marin, F.; Luquet, G.; Marie, B.; Medakovic, D. Molluscan Shell Proteins: Primary Structure, Origin, and Evolution. In *Current Topics in Developmental Biology*; Academic Press: Cambridge, MA, USA, 2007; Volume 80, pp. 209–276.
58. Lanio, W.; Swida, U.; Kress, H. Molecular Cloning of the Drosophila Virilis Larval Glue Protein Gene Lgp-3 and Its Comparative Analysis with Other Drosophila Glue Protein Genes. *Biochim. Biophys. Acta BBA—Gene Struct. Expr.* **1994**, *1219*, 576–580. [CrossRef]
59. Mateos, B.; Conrad-Billroth, C.; Schiavina, M.; Beier, A.; Kontaxis, G.; Konrat, R.; Felli, I.C.; Pierattelli, R. The Ambivalent Role of Proline Residues in an Intrinsically Disordered Protein: From Disorder Promoters to Compaction Facilitators. *J. Mol. Biol.* **2020**, *432*, 3093–3111. [CrossRef]
60. Mettling, C.; Bourouis, M.; Richards, G. Allelic Variation at the Nueleotide Level in Drosophilaglue Genes. *Mol. Gen. Genet. MGG* **1985**, *201*, 265–268. [CrossRef]
61. Andres, A.J.; Fletcher, J.C.; Karim, F.D.; Thummel, C.S. Molecular Analysis of the Initiation of Insect Metamorphosis: A Comparative Study of Drosophila Ecdysteroid-Regulated Transcription. *Dev. Biol.* **1993**, *160*, 388–404. [CrossRef] [PubMed]
62. Li, T.-R.; White, K.P. Tissue-Specific Gene Expression and Ecdysone-Regulated Genomic Networks in Drosophila. *Dev. Cell* **2003**, *5*, 59–72. [CrossRef]
63. Barnett, S.W.; Flynn, K.; Webster, M.K.; Beckendorf, S.K. Noncoordinate Expression of Drosophila Glue Genes: Sgs-4 Is Expressed at Many Stages and in Two Different Tissues. *Dev. Biol.* **1990**, *140*, 362–373. [CrossRef]
64. Swida, U.; Lucka, L.; Kress, H. Glue Protein Genes in Drosophila Virilis: Their Organization, Developmental Control of Transcription and Specific MRNA Degradation. *Development* **1990**, *108*, 269–280. [CrossRef]
65. Kress, H. Ecdysone-Induced Puffing InDrosophila: A Model. *Naturwissenschaften* **1981**, *68*, 28–33. [CrossRef]
66. Kress, H. Biochemical and Ontogenetic Aspects of Glycoprotein Synthesis in Drosophila Virilis Salivary Glands. *Dev. Biol.* **1982**, *93*, 231–239. [CrossRef]
67. Ramesh, S.R.; Kalisch, W.-E. Glue Proteins InDrosophila Nasuta. *Biochem. Genet.* **1988**, *26*, 527–541. [CrossRef] [PubMed]
68. Ramesh, S.R.; Kalisch, W.-E. Comparative Analysis of Glue Proteins in TheDrosophila Nasuta Subgroup. *Biochem. Genet.* **1989**, *27*, 507–520. [CrossRef]
69. Ramesh, S.R.; Shivanna, N. SDS-PAGE Pattern Polymorphism of X-Chromosomal Glue Proteins in Natural Populations of Two Drosophila Nasuta Subgroup Species. *Biochem. Genet.* **1999**, *37*, 1–21. [CrossRef]
70. Perkowska, E. Some Characteristics of the Salivary Gland Secretion of Drosophila Virilis. *Exp. Cell Res.* **1963**, *32*, 259–271. [CrossRef]
71. Thomopoulos, G.N.; Kastritsis, C.D. A Comparative Ultrastructural Study of 'Glue' Production and Secretion of the Salivary Glands in Different Species of TheDrosophila Melanogaster Group. *Wilhelm Rouxs Arch. Dev. Biol.* **1979**, *187*, 329–354. [CrossRef]
72. Paris, M.; Boyer, R.; Jaenichen, R.; Wolf, J.; Karageorgi, M.; Green, J.; Cagnon, M.; Parinello, H.; Estoup, A.; Gautier, M.; et al. Near-chromosome level genome assembly of the fruit pest Drosophila suzukii using long-read sequencing. *Sci. Rep.* **2020**, *10*, 1–14. [CrossRef]
73. Shirk, P.D.; Roberts, P.A.; Harn, C.H. Synthesis and Secretion of Salivary Gland Proteins in Drosophila Gibberosa during Larval and Prepupal Development. *Rouxs Arch. Dev. Biol.* **1988**, *197*, 66–74. [CrossRef] [PubMed]

74. Hsu, W.S. The Golgi Material and Mitochondria in the Salivary Glands of the Larva of Drosophila Melanogaster. *J. Cell Sci.* **1948**, *s3-89*, 401–414. [CrossRef]
75. Gregg, T.G.; McCrate, A.; Reveal, G.; Hall, S.; Rypstra, A.L. Insectivory and Social Digestion InDrosophila. *Biochem. Genet.* **1990**, *28*, 197–207. [CrossRef]
76. Ashburner, M. Function and Structure of Polytene Chromosomes During Insect Development. In *Advances in Insect Physiology*; Elsevier: Amsterdam, The Netherlands, 1970; Volume 7, pp. 1–95. ISBN 978-0-12-024207-8.
77. Jones, N.A.; Kuo, Y.M.; Sun, Y.H.; Beckendorf, S.K. The Drosophila Pax Gene Eye Gone Is Required for Embryonic Salivary Duct Development. *Development* **1998**, *125*, 4163–4174. [CrossRef]
78. Beňová-Liszeková, D.; Mentelová, L.; Babišová, K.; Beňo, M.; Pechan, T.; Chase, B.A.; Farkaš, R. An Apocrine Mechanism Delivers a Fully Immunocompetent Exocrine Secretion. *Sci. Rep.* **2021**, *11*, 15915. [CrossRef]
79. Seyahooei, M.A.; Kraaijeveld-Smit, F.J.L.; Kraaijeveld, K.; Crooijmans, J.B.M.; Van Dooren, T.J.M.; Van Alphen, J.J.M. Closely Related Parasitoids Induce Different Pupation and Foraging Responses in Drosophila Larvae. *Oikos* **2009**, *118*, 1148–1157. [CrossRef]
80. Rossi, G.; Manfrin, A.; Lutolf, M.P. Progress and Potential in Organoid Research. *Nat. Rev. Genet.* **2018**, *19*, 671–687. [CrossRef] [PubMed]
81. Low, L.A.; Mummery, C.; Berridge, B.R.; Austin, C.P.; Tagle, D.A. Organs-on-Chips: Into the next Decade. *Nat. Rev. Drug Discov.* **2021**, *20*, 345–361. [CrossRef] [PubMed]

1.6 Objectives

During my thesis, I addressed three main questions to understand the evolution of glue genes and adhesion properties across several Drosophila species.

1. What is the evolutionary dynamics of several glue genes in Drosophila species ?

My first objective was to further study the evolution of glue genes among several Drosophila species. *D. melanogaster* possesses eight glue genes and previous work has shown that glue genes evolved more rapidly than other genes (Da Lage et al., 2019). Preliminary analyses suggested that *Sgs3*, *Sgs7* and *Sgs8* underwent more gene duplications and deletions than the other glue genes. We annotated *Sgs* genes in the recently published genome sequences of 24 Drosophila species and analysed the number of duplications, deletions, gene inversion and conversion events in *Sgs1*, *Sgs3*, *Sgs7* and *Sgs8*. This bioinformatics analysis was mainly done during the first year of my PhD.

2. What is the variation in adhesion properties between Diptera species ?

My second objective was to assess the adhesive properties of fly species other than *D. melanogaster*. We used an industrial adhesion test machine and followed the protocol previously developed in the lab to measure adhesion (Borne et al., 2021b). We tested the influence of the procotol parameters on adhesion measurements. We compared the adhesion force, energy of detachment, rigidity, elasticity and plasticity of the pupa and its glue for more than 25 species. An R code was developed in collaboration with Jean-Noël Lorenzi, bioinformatician in our team, to analyse and compare the data. In addition to the adhesion tests, we measured pupa size and the surface of contact between the glue and the glass slide. These measurements allowed us to infer correlations between pupa morphology and its adhesion properties.

3. What is the role of each Sgs protein in *Drosophila melanogaster* glue adhesion ?

To assess the role of each *Sgs* gene in the adhesive properties of the glue of *D. melanogaster*, we inhibited individually the expression of the glue genes using two genetic methods. The first technique consisted in using UAS-RNAi lines already available and crossing them to a GAL4 line which drives expression in the salivary glands. For the second one, we used the Gene-Switch technique which can turn off and on the GAL4 activity (Roman et al., 2001). This work was done in collaboration with Isabelle Nuez, engineer in the lab, and Kelly Ten Hagen's team from the National Institutes of Health (Bethesda, USA), who studied the consequence of RNAi on the glue aspect via electron microscopy.

Chapter 2

Higher evolutionary dynamics of gene copy number for Drosophila glue genes located near short repeat sequences

2.1 Abstract

This section aims to study the evolution of some *Sgs* genes in several Drosophila species. Salivary Gland Secretion (Sgs) proteins are the main components of the secreted glue and are thus supposed to be responsible, at least in part, for the glue adhesive properties.

While *D. melanogaster* presents eight *Sgs* genes, this number varies in other Drosophila species. In particular, *Sgs3*, *Sgs7* and *Sgs8* are forming a cluster of genes that underwent many duplications and deletions when compared to other glue genes (Da Lage et al., 2019). We decided to focus our study on the group *Sgs3/Sgs7/Sgs8* and to compare its evolution with respect to a gene from the same gene family, *Sgs1*.

In a previous study of *Sgs* genes evolution, *Sgs* genes were identified using genome assemblies derived from Illumina short reads (Figure 1.13) (Da Lage et al., 2019). Our approach differed in three aspects. First, we took advantage of recently published Drosophila genome assemblies based on Pacbio or Oxford Nanopore technologies, while Da Lage et al. (2019) study used Illumina sequenced genomes. These third generation sequencing techniques generated long reads of DNA (10000 to 30000 bp) while Illumina technique produces short reads of DNA (50 to 300 bp). Hence, genomes used in our study were more accurate in terms of gene order and gene orientation. Second, we used multiple *Sgs* sequences from 24 Drosophila species to search for orthologous sequences by BLAST while the previous study used only *D. melanogaster Sgs* sequences for the BLAST searches. This enabled us to find extra *Sgs* copies not identified in the previous study in species distantly related to *D. melanogaster*. Third, we studied *Sgs* genes synteny, i.e. the neighbouring genes adjacent to *Sgs* genes. In Da Lage et al. (2019) study, only the *Sgs* sequences

were analysed. We used synteny analysis to better understand the evolutionary dynamics of the glue genes genomic region and to compare it between the 24 Drosophila species studied.

Overall, our objective was to better understand the evolution of *Sgs1/Sgs3/Sgs7/Sgs8* across Drosophila species.

2.2 Article - "Higher evolutionary dynamics of gene copy number for Drosophila glue genes located near short repeat sequences"

This work is a preprint manuscript, submitted to BMC Evolution and Ecology and is currently under review. The article and the supplementary materials can be found here : https://www.researchsquare.com/article/rs-2495385/v2

Higher evolutionary dynamics of gene copy number for Drosophila glue genes located near short repeat sequences

Manon Monier
Institut Jacques Monod
Isabelle Nuez
Institut Jacques Monod
Flora Borne
Columbia University
Virginie Courtier-Orgogozo (virginie.courtier@normalesup.org)
Institut Jacques Monod

Research Article

Keywords: glue genes, bioadhesive, Sgs, mucin, gene family, gene diversification, gene turnover, gene loss, gene duplication, synteny, repeat

Posted Date: June 2nd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2495385/v2

Abstract

Background

During evolution, genes can experience duplications, losses, inversions and gene conversions. Why certain genes are more dynamic than others is poorly understood. Here we examine how several *Sgs* genes encoding glue proteins, which make up a bioadhesive that sticks the animal during metamorphosis, have evolved in Drosophila species.

Results

We examined high-quality genome assemblies of 24 Drosophila species to study the evolutionary dynamics of four glue genes that are present in *D. melanogaster* and are part of the same gene family - *Sgs1, Sgs3, Sgs7 and Sgs8* - across approximately 30 millions of years. We annotated a total of 102 *Sgs* genes and grouped them into 4 subfamilies. We present here a new nomenclature for these *Sgs* genes based on protein sequence conservation, genomic location and presence/absence of internal repeats. Two types of glue genes were uncovered. The first category (*Sgs1, Sgs3x, Sgs3e*) showed a few gene losses but no duplication, no local inversion and no gene conversion. The second group (*Sgs3b, Sgs7, Sgs8*) exhibited multiple events of gene losses, gene duplications, local inversions and gene conversions. Our data suggest that the presence of short "new glue" genes near the genes of the latter group may have accelerated their dynamics.

Conclusions

Our comparative analysis suggests that the evolutionary dynamics of glue genes is influenced by genomic context. Our molecular, phylogenetic and comparative analysis of the four glue genes *Sgs1, Sgs3, Sgs7* and *Sgs8* provides the foundation for investigating the role of the various glue genes during Drosophila life.

Background

Genes can be grouped into gene families when they share a common ancestor and are present either in distinct genomes (orthologs and paralogs) or within a single genome (paralogs) due to gene duplications [1]. The increase in gene copy number in a genome can have several fitness advantages: to increase the amount of products (e.g., ribosomal RNAs), to diversify protein activity (e.g., opsins) and to diversify gene expression patterns (e.g., Hox transcription factors) [2]. Gene duplications and gene losses are frequently involved in phenotypic evolution and adaptation [2–5]. In humans, on a per nucleotide basis, gene copy number differences between individuals represent an even larger pool of genetic variation available to selection than single nucleotide polymorphisms [1, 6].

Certain genes are found to exhibit accelerated rates of gene turnover and several factors have been proposed to explain why the pace of gene duplication and gene loss can differ between genes. A first type of explanation relates to the selective forces that act on genes. For example, genes involved in interactions with the environment such as chemoreception, reproduction, stress response or immune defense are generally expected to adapt faster due to conditions that change more rapidly and indeed they are usually observed to undergo faster gene turnover than average genes [1, 7]. In contrast, a few particular genes may require strict stoichiometric balance due to their interactions with other proteins and are less likely to vary in gene copy number [8–10]. A second type of explanation considers the rate of the mutation process itself. Structural changes and thus gene turnover can be facilitated by the presence of certain elements in the genome, such as repeated sequences [11], transposable elements [12] or fragile DNA regions that are more susceptible to DNA breakage [13].

Duplicated gene copies are often clustered at specific genomic locations [14]. Examining the immediate surroundings of gene copies, researchers have often noticed the presence of transposable elements, for example for pigmentation transcription factor genes in maize [15], effector genes in grass powdery mildew [16], insecticide resistance genes in Drosophila [17], amylase genes in Vertebrates [18] and fatty acid metabolic genes in fish [19]. Transposable elements usually flank genes and are oriented in the same direction. They provide regions of high sequence identity that can be used as templates for unequal crossing overs, resulting in the removal or duplication of gene coding sequences between the two elements [12].

The increasing number of available full genome sequences from a variety of organisms offers an unprecedented opportunity to investigate more thoroughly the tempo of gene turnover and the evolutionary forces controlling gene gains and losses. High quality assemblies are required to correctly infer the rates of gene turnover. In case of sequencing errors, certain gene copies and short open-reading frames can be missed. Errors in genome assemblies can also lead to the fragmentation of genes into several individual contigs, the withdrawal of recent duplicates, the split of heterozygous single-copy genes or even sometimes the incorporation of gene sequences from contaminant species [20]. Such incorrect assessment of the number of gene copies within genomes usually lead to higher estimates of the rates of gene gains and losses [21]. On the other hand, comparing species that are too distantly related can overlook rapid duplications followed by the elimination of one of the extra gene copy and lead to an underestimation of gene turnover rates. Overall, gene turnover is best assessed with closely related species and genomes based on long-read sequencing methods. To help in finding ortholog genes and confirming potential gene losses, it can also be useful to perform whole-genome alignments, determine syntenic regions where genes are expected to occur and then search for the presence of the genes of interest in the syntenic region [22].

The Drosophila glue genes, also named Salivary gland secretion (*Sgs*) genes, represent a simple and attractive model system to study the evolutionary forces acting on the evolutionary dynamics of gene copies [23]. These genes encode secreted proteins that make up a bioadhesive that allows the animal to attach itself to a surface for several days while it remains still during metamorphosis [25]. The glue of

diverse Drosophila species is thought to evolve rapidly to stick to various substrates in diverse environmental conditions [25]. The specificity of Drosophila glue genes, with the exception of *Eig71Ee* (see below), is that they have only one known function, glue production. Compared to genes with multiple functions, they are thus presumably subjected to more defined and precise selective forces, which might facilitate our understanding of their evolutionary dynamics. In addition, assessing the diversity of glue genes encoded by different Drosophila species may help to identify key components of Drosophila glue adhesiveness and develop new bioadhesives.

In *Drosophila melanogaster* eight glue genes have been identified [25]. Five of them, - *Sgs1 (2L:25B4), Sgs3 (3L:68C11), Sgs7* (3L:68C11), *Sgs8* (3L:68C11) and *Eig71Ee* (3L:71E5) harbor a phase 1 intron at the same position, which interrupts the signal peptide, and are considered to be part of the same gene family [26], The three other genes - *Sgs4* (X:3C11-12), *Sgs5* (3R:90B3-5) and *Sgs5bis* (3R:90B3-5) - have no intron (for *Sgs4*) or harbor two introns at other positions (for *Sgs5* and *Sgs5bis*). Their relationships with respect to the other glue genes have not been characterized. *Sgs1*, *Sgs3, Sgs4* and *Eig71Ee* encode for long, highly O-glycosylated proteins containing a large, disordered region harboring repeat sequences rich in proline, serine and threonine [25]. The repeat region is characteristic of mucins, which usually form a mucus which can act as a physical barrier against mechanical damage or pathogens [27]. *Sgs5, Sgs5bis, Sgs7* and *Sgs8* genes encode for shorter and more ordered proteins that are rich in cysteine and devoid of internal repeats [25]. All the D. melanogaster glue genes are only expressed in the salivary glands at the third instar larval stage and only known to be involved in glue production [25], with the exception of *Eig71Ee*, which is also expressed in hemocytes and in the gut, where it appears to contribute to coagulation and bacterial entrapment [28]. In a previous study [26], the rate of gene gains and losses for the *Sgs1-Sgs3-Sgs7-Sgs8* gene family was found to be significantly higher than for average genes. Here, after clarifying the relationships between the eight glue genes of *D. melanogaster*, we focus on the evolution of four glue genes: *Sgs1*, *Sgs3*, *Sgs7* and *Sgs8*. We use recently published high quality assemblies of closely related species of Drosophila flies [24] to reconstruct their evolutionary dynamics across approximately 30 million years of evolution. We observe that the rates of gene duplication, gene inversion and gene conversion vary between genes, and we explore the possible effect of genomic context on gene dynamics.

Results

Two families of glue genes in D. melanogaster

Alignments of the amino acid sequences encoded by the eight glue genes of *D. melanogaster* and their annotated orthologs from various Drosophila species [26] revealed that Drosophila glue genes form two distinct gene families and that there is no sequence match between them besides the signal peptide (Fig. 1, Fig. S1, Files S1-2). The first gene family comprises *Sgs1, Sgs3, Sgs7, Sgs8* and *Eig71Ee* (Fig. 1, File S2) whereas the second gene family contains *Sgs4, Sgs5* and *Sgs5bis* (Fig. S1, File S2). Genes of the first gene family are characterized by an IRXC[L/V]C motif in the encoded C-terminal domain and the presence of a phase 1 intron disrupting the signal peptide sequence whose position corresponds to

amino acid position 10 (Fig. 1A). The second family proteins display a PCXXXXK motif in the C-terminal region (Fig. S1A).

In a previous study [26], we found that for the group of *Sgs1, Sgs3, Sgs7* and *Sgs8* genes, the rate of gene gains and losses was significantly higher than for average genes. In order to examine further the evolutionary dynamics of gene copies for this glue gene family and the factors influencing their rate of evolution, we decided to take advantage of high quality genome assemblies that became available in 2021 [24]. We chose to focus on closely related species of Drosophila which diverged relatively recently, so that we were unlikely to interpret as gene copy stasis situations that resulted from rapid duplications followed by the elimination of one of the duplicated copies. In the present study, we did not analyze *Eig71Ee*, as it has a supplementary role in immune defense and is thus probably subjected to additional functional constraints compared to the other glue genes. Overall, we examined the evolutionary dynamics of four glue genes - *Sgs1, Sgs3, Sgs7* and *Sgs8* - across 25 Drosophila species.

Existing genome annotations are often incomplete for Sgs genes

Using BLAST [29], we identified and annotated all copies of the *Sgs* genes which are orthologs of *Sgs1, Sgs3, Sgs7* and *Sgs8* in high-quality genome reference sequences of *D. melanogaster* and 23 other Drosophila species (Table S1-3, File S1). Compared to previous studies of *Sgs* genes in diverse Drosophila species [26, 30], we analyzed here the genome sequence of 6 additional Drosophila species: *D. teissieri*, *D. triauraria*, *D. rufa*, *D. jambulina*, *D. obscura* and *D. subobscura*. Compared to Da Lage et al. previous study [26], which used only protein sequences from *D. melanogaster* as queries for BLAST searches, we used *Sgs* sequences from all species as BLAST queries and compared large genomic syntenic blocks between species. We thus identified 13 additional *Sgs* genes in the species examined by Da Lage et al. and annotated 13 new *Sgs* genes in genome sequences from four other species (Table S3). Furthermore, we corrected gene annotations for five *Sgs* genes in five species, where introns were absent or mislabeled (Table S3, File S1).

Da Lage et al. [26] annotated four *Sgs7* genes in *D. suzukii* based on a low-quality genome assembly [31]. Using a more recent Pacbio assembled genome [32] of the same strain, we found only one copy of *Sgs7*, located at the same position as in its closely related species *D. biarmipes*. This illustrates that determination of the number of gene copies is highly dependent on high quality genomes [20, 21]. In the present study we relied on PacBio- and Nanopore-based genome assemblies for all species, except for *D. eugracilis* and *D. takahashii* which had only Illumina-based genome sequences (Table S1).

A new nomenclature for Sgs3 genes

While *D. melanogaster* harbors a single *Sgs3* gene, multiple copies of this gene were previously found in several Drosophila species and were distinguished with letters *a, b, c* according to the number of copies per species and to the order of their discovery in each species [26]. Here, as we found even more *Sgs3* copies, we decided to change the gene nomenclature for better comparison between species. We define *Sgs3x* as the *Sgs3* ortholog that is deleted in the melanogaster subgroup and that is flanked in other

species by the *Parg* (*CG2864*) and *Mnt* (CG13316) genes in a large genomic syntenic block, which corresponds to position 3E2 on the X chromosome in *D. melanogaster*. All the other *Sgs3* copies are in a large genomic syntenic block corresponding to region 68C10-11 on chromosome 3L *in D. melanogaster*. We labeled them from 'b' to 'g' from 5' (near the *Chrb* gene) to 3' (near the *CG33489* gene) according to their respective positions within this genomic locus. We note that for serendipitous reasons there is no *Sgs3a* gene in this new nomenclature. *Sgs3* genes located at the same corresponding position in the genome of diverse species were labeled with the same letter.

Several Sgs genes incorrectly contained premature stop codons

The coding regions of *Sgs1* and *Sgs3* contain long internal repeats encoding motifs rich in proline, serine and threonine [25]. Premature stop codons were found in genome sequence assemblies within the repeated region of *Sgs1* in four species (*D. takahashii*, *D. rhopaloa*, *D. triauraria* and *D. ficusphila*) and of *Sgs3x* in *D. biarmipes.* Using a *D. takahashii* strain different from the genome sequence line, we PCR-amplified the region containing the presumptive premature stop codon and found an extra A nucleotide compared to the reference sequence of *Sgs1*, making up a stretch of 8 adenines instead of 7. The addition of this adenine removed the premature stop codon and gave a full length *Sgs1* coding region. In *D. triauraria* we found 6 premature stop codons dispersed throughout the 4212-bp repeated region of *Sgs1*, with frameshifts adjacent to each stop codon. The presence of repeats prevented us from amplifying the region by PCR, so we do not know whether these are genuine stop codons or sequence assembly artifacts. Analysis of raw reads from full genome sequencing projects suggests that D. rhopaloa *Sgs1* reference sequence may be corrected by adding an extra 'A' (supported by 21 reads compared to 42 reads harboring a deletion), that *D. ficusphila Sgs1* reference sequence should be corrected by removing a 'C' from a 6-bp stretch of C (supported by 45 reads harboring a deletion versus 10 reads an extra C) and that *D. biarmipes Sgs3x* reference sequence should be corrected by adding an extra 'C' (supported by 13 reads compared to 4 reads harboring a deletion) (Fig. S2, File S3). We therefore considered the modified sequences for these three species in our subsequent analysis.

In summary, we detected premature stop codons in five *Sgs* genes. Four of them likely correspond to sequence assembly errors. For *D. triauraria Sgs1*, it is not clear whether the 6 premature stop codons are real or artifactual.

The Sgs1, Sgs3, Sgs7 and Sgs8 genes form four subfamilies

The four genes *Sgs1, Sgs3, Sgs7* and *Sgs8* encode proteins with a signal peptide and conserved amino acid motif patterns in the N-terminal and C-terminal regions (Fig. 1A, File S4-5). They harbor two coding exons and a short phase 1 intron interrupting the signal peptide. They can be grouped into four subfamilies based on their genomic location and synteny: *Sgs1*, *Sgs3* (which includes *Sgs3b-g* genes but not *Sgs3x)*, *Sgs3x* and *Sgs7-8* (see below for a description of each subfamily). *Sgs* coding sequence length varies greatly between genes and species, with *Sgs1* being the longest gene (higher than 1,7 kb in all species) and *Sgs7-8* the smallest ones (between 222 and 240 bp in all species) (Fig. 2, File S5-6). The genes *Sgs7* and *Sgs8* are closely related to *Sgs3* and they can be distinguished from *Sgs3* by the length

of their coding sequence (Fig. 2) and the fact that they are located at other genomic locations (see below).

Sgs1 did not duplicate and was lost at least twice via gene deletions

In all the Drosophila species studied, *Sgs1* is composed of a first coding exon which is always 28 bp, a short phase 1 intron whose size varies between 50 bp and 71 bp, and a second exon which harbors a long repeat region and whose size varies from 1,758 bp in *D. takahashii* to 5,861 bp in *D. rufa* (Table S4). The synteny of *Sgs1* and its neighboring genes is conserved across all species (Fig. 3–5, Table S3). Using BLAST searches, *Sgs1* was not found in *D. erecta* and *D. kikkawai.* The loss of *Sgs1* in *D. erecta* and in *D. kikkawai* is associated with a 4-kb and a 3-kb deletion, respectively (according to *D. teissieri* and *D. jambulina* sequences, respectively), thus removing the full *Sgs1* coding region while preserving the two neighboring coding genes *hoe2* and *CG14044* (Fig. 4–5, File S7). We conclude that two recent *Sgs1* gene losses occurred, in association with gene-wide deletions.

In the outgroup species *D. pseudoobscura, D. obscura* and *D. subobscura*, and in further distantly related species, no *Sgs1* gene was found at the syntenic location (Fig. 5) nor across the whole genome via BLAST. This suggests that the *Sgs1* gene appeared after the divergence between the most recent common ancestor of these species and *D. melanogaster*, i.e. about 30 million years ago [33]. Our analysis reveals that since its appearance within the Drosophila genus, the *Sgs1* gene has maintained the same neighboring genes throughout all the Drosophila species we examined and that it did not duplicate.

Sgs3x did not duplicate and was lost at least three times via gene deletion

As for *Sgs1*, the first coding exon of *Sgs3x* is 28 bp in all the studied species and the second exon harbors repeats and varies in size, from 581 bp for *D. elegans* to 4,148 bp for *D. bipectinata*. In all species featuring an *Sgs3x* gene, the gene is located at the same corresponding genomic location, between genes *Parg* (*CG2864*) and *Mnt* (*CG13316*) (Fig. 3).

The most parsimonious scenario is that *Sgs3x* was already present in one copy in the ancestor of the species studied here. Based on our phylogenetic analysis and parsimony, we infer that *Sgs3x* has been lost three times: before the most recent common ancestor of *D. melanogaster* and *D. erecta* (*melanogaster* subgroup) (Fig. 6, via a 1-kb deletion when compared with *D. eugracilis)*, in the ancestor of *D. triauraria, D. rufa, D. jambulina* and *D. kikkawai* (*montium* group) (Fig. 7, via a 2-kb deletion compared to *D. bipectinata*) and in the ancestor of *D. ficusphila* (Fig. 7, via a 1-kb deletion compared to *D. elegans).* Overall, *Sgs3x* exhibits an evolutionary history like *Sgs1*: it did not change neighboring genes, did not duplicate and experienced deletions of its full gene coding sequence in a few species.

Two Sgs3 copies lost their internal repeats in the lineage leading to D. subobscura

We define *Sgs3, Sgs7* and *Sgs8* as copies of the *Sgs1-Sgs3-Sgs7-Sgs8* gene family that are present within a large genomic syntenic block corresponding to region 68C10-11 on chromosome 3L *in D. melanogaster.* The *Sgs3* genes are distinguished from *Sgs7* and *Sgs8* by the presence of repeats and by longer coding

regions (Fig. 2). However, in *D. obscura*, at the loci occupied by *Sgs3b* and *Sgs3d* in *D. subobscura*, we detected two *Sgs3* genes which are shorter (both 270 bp) than typical *Sgs3* genes (Fig. 2), do not present internal repeats but share similar N-terminal and C-terminal regions with their corresponding *Sgs3* copies in *D. subobscura* (Fig. 8). Dot plots suggest that the repeated sequences of *Sgs3b* and *Sgs3d* were lost in the lineage leading to *D. obscura* (Fig. 8–9). We named the resulting genes in *D. obscura Sgs3bshort* and *Sgs3dshort.* The coding sequence of these two genes are extremely similar (Fig. 1B), suggesting that they originate from a recent gene conversion event in the lineage leading to *D. obscura* (Fig. S3-4). In addition to *Sgs3bshort* and *Sgs3dshort*, *D. obscura* possesses a copy of *Sgs3e* harboring internal repeats (Fig. 8–9). Complete losses of internal repeats were not observed in *Sgs1* nor in *Sgs3x* (Table 1).

Sgs3 underwent several duplications, deletions, inversions and gene conversions

As opposed to *Sgs1* and *Sgs3x, Sgs3* first exon varies slightly in size, from 19 bp to 28 bp (Table S4). The second exon length varies from 356 bp in *D. jambulina Sgs3b* to 1967 bp in *D. bipectinata Sgs3e* (Table S4). The beginning of the second exon of *Sgs3* encodes for a relatively conserved amino acid sequence, ASILLI (Fig. 1A). Two *Sgs3* copies are found in most of the studied species: *Sgs3b* (which is located between genes *CG33272* and *CG7512*) and *Sgs3e* (which is located within an intron of the gene *Mob2*) (Fig. 9, S4). Parsimony suggests that both genes were present in the most recent common ancestor of all studied species (Table 1). Comparison of protein sequences (File S8) shows that *Sgs3c*, *Sgs3d*, *Sgs3f* and *Sgs3g* are duplicates of *Sgs3b* and that *Sgs3e* did not duplicate in the lineages studied here. The high similarity between the two *Sgs3* copies present in *D. pseudoobscura* is also indicative of gene conversion. Parsimony principle indicates that across the 24 studied species, *Sgs3e* underwent 2 gene losses and no duplications whereas *Sgs3b* experienced 2 gene losses and 4 gene duplications, all within the same syntenic block (Fig. 9, Table 1). Furthermore, inversions of the entire *Sgs3* coding sequence, together with adjacent regions, occurred in two instances (crosses in Fig. 9, S5). Such inversions were not observed for *Sgs1* nor for *Sgs3x* (Table 1).

Table 1
Summary of the sequence changes observed for the different *Sgs* gene subfamilies in the 24 studied species. Numbers indicate the number of genetic events inferred for each *Sgs* gene.

	Sgs1	*Sgs3x*	*Sgs3e*	*Sgs3b*	*Sgs7-Sgs8*
inferred number of copies in the common ancestor of all studied species	0 (appeared after the *D. melanogaster/D. pseudoobscura* divergence)	1	1	1	0 (appeared after the *D. melanogaster/D. pseudoobscura* divergence)
position and orientation relative to neighboring genes	constant	constant	constant	variable	variable
first coding exon size	constant (28bp)	constant (28bp)	variable (19-28bp)	variable (25-31bp)	constant (28bp)
internal repeats	present	present	typically present	typically present	typically absent
loss of all the internal repeats	0	0	0	2	not applicable
gene deletion	2	3	2	2	4
gene duplication	0	0	0	4	≥ 3
gene inversion	0	0	0	2	≥ 1
gene conversion	0	0	0	2	≥ 3

Sgs7 and Sgs8 underwent several duplications, gene losses and gene conversion

D. melanogaster possesses two glue genes near *Sgs3b* that are devoid of internal repeats, *Sgs7* and *Sgs8*. In the other 23 Drosophila species, we annotated in the corresponding syntenic region 0, 1, 2 or 3 *Sgs* genes with no repeats (Fig. 9). For all these *Sgs7* and *Sgs8* orthologs, the size of the first coding exon is 28 bp and the second coding exon size varies between 194 bp in *D. ananassae Sgs7* and 212 bp in *D. bipectinata Sgs7b*.

The two *Sgs8* copies in *D. eugracilis* exhibit very similar sequences (Fig. S6), suggesting that they originated from a recent duplication or from gene conversion in the branch leading to *D. eugracilis* (Fig. 9). Similarly, another recent duplication or gene conversion event seems to have occurred in the branch leading to *D. takahashii* (Fig. 9–10). In certain cases, it was impossible to determine with absolute confidence whether the different copies correspond to *Sgs7* or *Sgs8*, due to their short coding sequences, their rapid divergence and signs of gene conversion. For example, *D. erecta* and *D. teissieri* harbor *Sgs* genes at the exact genomic positions corresponding to *D. melanogaster Sgs7* and *Sgs8* genes (Fig. 10).

However, at the *Sgs7* position in *D. teissieri* is a coding region which is closer to *Sgs8* than *Sgs7*, and reciprocally at the *Sgs8* position (Fig. 1B). Dot plot analysis (Fig. S7) suggests that gene conversion occurred between *Sgs7* and *Sgs8* in the lineage leading to *D. teissieri*. Overall, our distinctions between the *Sgs7* and *Sgs8* genes are thus subject to caution.

In addition, synteny comparisons suggest that an inversion occurred between the group of *D. santomea*, *D. yakuba*, *D. teissieri* and *D. erecta*, and the *melanogaster* complex (*D. melanogaster*, *D. simulans*, *D. sechellia* and *D. mauritiana*), which inverted a pair of *Sgs7* and *Sgs8* genes together with their adjacent genes (Fig. 9–10, S8). And further gene conversion events blurred the relationships between *Sgs7* and *Sgs8* in these four species (Fig. 9–10, S8).

In summary, a single copy of *Sgs7*-8 was probably present in the common ancestor of *D. kikkawai* and *D. melanogaster*. It underwent at least 4 deletions, 3 duplications, one inversion and several gene conversion events (Table 1).

Genomic instability is associated with the presence of short "new glue" genes

Our analysis reveals two types of gene dynamics. A first group of genes, comprising *Sgs1*, *Sgs3x* and *Sgs3e*, experienced several gene losses but no duplication, no local inversion and no gene conversion across the 24 Drosophila species studied here. In contrast, the second category, involving *Sgs3b*, *Sgs7* and *Sgs8*, underwent multiple events of duplication, local inversion and gene conversion (Table 1, Fig. 9).

To test the potential involvement of repetitive elements, we looked for the presence of repeated sequences across 129-kb regions encompassing each *Sgs* gene in several Drosophila species (Fig. S9). We found that in *D. melanogaster* repeats are more frequent near the *Sgs3b/Sgs7/Sgs8* genes than around the *Sgs1* and *Sgs3x* genes. Furthermore, the recently duplicated genes *Sgs3c* and *Sgs3d* in *D. subobscura* and *Sgs3f* and *Sgs3g* in *D. teissieri* locate within regions dense in repeats. Interestingly, multiple genomic changes (duplications, inversions) were found at the *Sgs7-8-3b* and *Sgs3f-g* loci, and similar stretches of sequences were detected at both loci (Fig. S10). These sequences contain short (243–426 bp), intronless genes encoding for threonine-rich proteins with predicted signal peptides. These genes resemble four genes adjacent to *Sgs4* that were previously annotated in *D. melanogaster* as "nested genes" or "new glue genes", even though their putative role in glue production is unclear [35, 36] (Fig. S11). We thus decided to name the new sequences we identified as *new glue* (*ng)* genes.

In total, we annotated 154 such *ng* genes in the *Sgs3-7-8* genomic region of the 24 studied Drosophila species (Table 2, S3). We define *ng* genes as encoding for proteins displaying the following characteristics: (1) a protein shorter than 180 amino acids, (2) a signal peptide, (3) an internal region rich in alanines and containing stretches of at least three consecutive threonines, and (4) a C-terminal region rich in arginines and lysines (Fig. S11). The previously annotated *ng4* gene from *D. melanogaster* does not exhibit characteristics (2) to (4). The threonine stretch can attain up to 17 consecutive threonines, as

in *D. ananassae LOC6500299*. Noticeably, almost all the *Sgs7* and *Sgs8* genes are adjacent and tail-to-tail to an *ng* gene, with approximately 130–200 bp separating the stop codons of both genes (beige arrows in Fig. 9). *Sgs3f* and *Sgs3g* are distant of approximately 400bp from their tail-to-tail adjacent *ng* gene. Most duplications and inversion events appear to preserve the contiguity and distance between the *Sgs* gene and its adjacent *ng* gene (Fig. S12-S14).

We used BLAST to search for *ng* genes in other parts of the genome and we identified three additional loci, containing *ng* genes but no *Sgs* genes, in several of the 24 studied species (Table 2). In *D. melanogaster*, two of these three loci (87A1 and 88C3-4) are separated from each other by approximately 2Mb. No *ng* gene was found at the *Sgs1* and *Sgs3x* loci. Furthermore, no *ng* genes were detected by BLAST in the full genomes of *D. virilis* and *D. hydei*. This suggests that *ng* genes appeared after the divergence of *D. virilis* and *D. melanogaster*.

In summary, a family of new genes called "new glue" genes was detected near *Sgs* genes in highly dynamic regions (*Sgs7-8-3b* and *Sgs3f-g*), but not in less dynamic regions (*Sgs1* and *Sgs3x*).

Table 2
Number of *ng* genes identified in 7 representative species (*D. melanogaster, D.ananassae, D. obscura, D. subobscura, D. willistoni* and *D. virilis*). Each column corresponds to a genomic region. Note that the 87A1 locus is located 5Mb away from *Sgs5* and that the 3C11-12 locus is 500kb away from *Sgs1* in *D. melanogaster*. No *ng* gene was found near *Sgs1, Sgs3e* and *Sgs3x*.

Species	**3C11-12 (near *Sgs4, Notch* and *dnc*)**	**68C11 (near *Sgs3b, Sgs7, Sgs8*)**	**68C13 (near *Sgs3f, Sgs3g*)**	**28E6-28E7 (near *mon2, Bsg* and *CG8673)***	**87A1 (near *cad87A, CG6959* and *sad)***	**88C3-4 (near *Cystatin-like*, *Phosphodiesterase 6* and *stumps*)**
D. melanogaster	4	2	4	none	none	4
D. ananassae	none	8	4	none	10	4
D. obscura	6	none	2	none	none	1
D. subobscura	5	none	none	none	none	3
D. willistoni	none	none	none	2	none	2
D. virilis	none	none	none	none	none	none
D. hydei	none	none	none	none	none	none

A recent gene duplication and an inversion were probably mediated by new glue genes

To investigate whether these new glue genes may have played a role in the evolutionary dynamics of genomic regions, we examined whether they were present at the boundaries of three relatively recent genomic rearrangements. First, we found that the duplication leading to *Sgs3d* in *D. subobscura* (which likely occurred approximately 15 million years ago [33]) (Fig. 9) included 5' and 3' non-coding regions surrounding the *Sgs3b* gene, and that there were no *ng* genes in the region (Fig. S15). Second, for the inversion of the *Sgs7-Sgs8* region which occurred just before the divergence of *D. teissieri* and *D. santomea* (around 2–11 million years ago [33]) (Fig. 9), we noticed that one of the breakpoints perfectly corresponds to the coding region of a *ng* gene (Fig. 11). Third, for the recent duplication leading to *Sgs3g* in *D. teissieri* (which occurred about 0–2 million years ago [33]), both breakpoints corresponded to *ng* genes (Fig. 11). The older the event, the more likely sequences at the breakpoints may be lost or modified. Here, we found that two breakpoints of a recent gene duplication and one breakpoint of an older inversion match the coding regions of *ng* genes. Given that *ng* genes are found in multiple copies over the genome, we suggest that they may facilitate large-scale genomic modifications such as gene inversion, gene duplications and gene losses.

Discussion

We reconstructed the evolutionary history of 102 *Sgs* genes present in 24 Drosophila species, including 26 newly annotated *Sgs* genes. Compared to our previous Da Lage et al. 2019 study [26], we used higher quality genome assemblies, synteny comparisons and blast queries from multiple species. This strategy allowed us to identify 13 new *Sgs* genes not reported in Da Lage et al. The *Sgs* glue genes can be difficult to annotate because their coding region is mostly composed of large repetitive sequences (prone to sequence misassembly and frameshifts) and evolves rapidly [23, 26]. We propose here a new nomenclature for *Sgs* genes based on protein sequence conservation, genomic location and presence/absence of internal repeats.

Our analysis suggests that three *Sgs* genes (*Sgs3x*, *Sgs3b*, *Sgs3e*) were probably present in the most recent common ancestor of all studied species and that the *Sgs1* and *Sgs7/8* genes arose after the divergence between *D. pseudoobscura* and *D. melanogaster*, i.e. about 30 million years ago [33]. No clear homologs of *Sgs1* and *Sgs7/8* were detected in more distantly related species using BLAST or HMMER, so the origin of these genes remain unclear.

The Sgs1 proteins exhibit a highly conserved motif, PCPC-X(1)-PQPP (Fig. 1A) which is also found in an uncharacterized domain of Suppressor of cytokine signaling 7 protein in mouse and human according to Prosite searches. The conserved motif C-x(2)-CGPGG from *Sgs3/7/8/3X* is found in the hormone transporter neurophysin in several mammal species and one mollusc. Interestingly, part of this sequence is also found in the repeat motifs (GGX or GPGXX) present in several silk proteins from spiders [37]. Thesestretches of amino acids probably evolved by convergent evolution in these proteins and in glue proteins.

Our present analysis of 24 Drosophila species spanning approximately 30 million years of evolution reveals that the *Sgs1, Sgs3x* and *Sgs3e* genes have remained at the same exact genomic location relative to their neighboring genes and did not duplicate, whereas the other genes (*Sgs3b*, *Sgs7, Sgs8*) have experienced inversions, translocations and duplications. Our observations are in agreement with a 1986 study which compared sequences from 5 closely related species of Drosophila and detected a 6-kb region containing *Sgs3*, *Ssg7* and Sgs8 which evolved faster than neighboring regions, via point mutations, insertions, deletions, inversions and the gain and loss of repetitive sequences [38]. In our study we did not assess mutation rate within coding sequences nor intraspecific variation.

In *D. virilis*, which diverged about 43 millions years ago from *D. melanogaster* [33] and was not examined in this study, three glue genes have been identified: *Sgs3a/Lgp1*, *Sgs3b/Lgp3* and *Sgs5bis/Lgp2* [25]. *Sgs* gene sequence divergence is too large between *D. virilis* and the species analyzed in this study to rely on phylogenetic trees to infer the relationship between their glue genes. *Sgs3a/Lgp1* and *Sgs3b/Lgp3* are adjacent to each other and result from a recent duplication in the *D. virilis* lineage [26]. Both genes lie near *AstA-R1, Ilp7, Parg* and *Rala* genes, which are also located at the *Sgs3x* locus in the species studied here. This suggests that *Sgs3a/Lgp1* and *Sgs3b/Lgp3* in *D. virilis* correspond to *Sgs3x* orthologs and that a gene duplication affecting *Sgs3x* did occur in species outside of the range of Drosophila species studied here.

Studies of *D. melanogaster Sgs1*, *Sgs3*, *Sgs4, Sgs5, Sgs7* and *Sgs8* indicate that glue genes display short, compact cis-regulatory regions that directly flank their start codon (within less than 1–2 kb) [39–44]. Such a characteristic, as observed for odorant receptor genes in insects [45], may facilitate gene turnover as shuffling of genomic regions is less likely to disrupt gene regulation. The *Sgs* genes we studied here display comparable expression patterns and amino acid sequences [25], so their difference in gene turnover dynamics does not seem to be related to variation in their gene function. Here we investigated the possible role of genomic context on glue gene dynamics. We observed that regions with high *Sgs* gene turnover contain copies of short coding genes named *new glue* (*ng*) genes that are immediately adjacent to the *Sgs* genes, whereas regions with low *Sgs* gene turnover do not. Several pieces of evidence suggest that the presence of these *ng* genes may accelerate gene dynamics: they are usually found in multiple copies at specific genomic locations, they lie near glue genes with rapid gene dynamics but not near the ones with reduced gene dynamics, they locate at two breakpoints of a recent *Sgs* gene duplication (0–2 million years ago) and at one breakpoint of an older inversion. These *ng* genes provide regions of high sequence identity for homologous recombination and thus may trigger genomic instability, similarly to the indirect effect of transposable elements on genome dynamics [12].

The four *ng* genes near *Sgs4* were first named "*nested genes*" (abbreviated as "ng") because they are nested together with *Sgs4* within the intron of the unrelated phosphodiesterase gene *dunce* [35, 36]. Three of them were found to resemble *Sgs3*, except that the intron was missing and the internal repeat region was smaller [35]. In the following publications, their name became "ng glue" [46] and then "new glue" [47, 48], with no justification given. In this study, we follow the most recent nomenclature and name them "new glue" (*ng*) genes, even though we are aware that no functional study has been reported so far to test

the hypothesis that they are involved in glue production or adhesion. We identified 154 *ng* genes in 24 Drosophila species. 89 of them are newly annotated genes that were not identified previously. The *ng* genes can be difficult to annotate because they appear to evolve rapidly and they are small genes (and thus may not generate sufficiently significant E-values in BLAST searches).

Our study reveals that *ng* genes surround not only *Sgs4* but also *Sgs3b/f/g*, *Sgs7* and *Sgs8* (Table 1). It would be interesting to examine the evolutionary dynamics of *Sgs4* genes to test whether the presence of neighboring *ng* genes might also promote genome dynamics at the *Sgs4* locus. In *D. melanogaster, ng* genes are found in at least four genomic locations and the expression pattern of three *ng* genes (*ng-1*, *ng-2* and *ng-3*) has been thoroughly studied in the 1990s. These three genes are exclusively expressed within the larval salivary glands [36] and only during a short temporal window, from the beginning of the third larval instar until the early wandering stage [49]. Proteins encoded by some of the *ng* genes have also been detected in a proteomics study in the whole body of developing larvae [50]. The presence of a putative signal peptide and an internal region rich in threonines (putative glycosylation sites) indicate that they may encode proteins that participate in the production of the glue. The presence of active ecdysone-responsive elements detected with the coding regions of *ng-1*, *ng-2* and *ng-3* [51, 52] also suggest that part of their function might be related to the regulation of expression of the neighboring glue genes. Several RNAi lines are available for future work to assess the role of *ng* genes in glue production and glue adhesiveness.

During animal evolution various glands evolved to produce large amounts of very specific proteins with diverse functions, such as venom in snakes and frogs or silk in spiders [53, 54]. Recent evolutionary studies indicate that, similarly to Drosophila glue genes, the genes encoding these secreted proteins underwent multiple events of gene duplications, losses and conversions in snakes and spiders [55, 56]. Our work on Drosophila glue genes, in combination with studies of these other secretory fluids, may thus help to provide general insights on how secretory products rapidly adapt to biotic and abiotic factors.

Conclusions

In this study, we used comparative phylogenomic methods to identify and characterize glue genes that are rapidly evolving in Drosophila species to better understand their dynamics in terms of duplications, losses, inversions and gene conversions. We uncovered several "glue" and "new glue" genes that were not found in previous studies and we propose a new nomenclature for glue genes. Our work highlights two modes of evolution for glue genes, differing in rates of inversion, duplication, gene loss and conversion. The most dynamic genes (*Sgs3b*, *Sgs7* and *Sgs8*) are in a region containing multiple "new glue" genes. Our analysis suggests that the presence of these short genes may have contributed to the higher dynamics of glue genes in this region. Our results serve as a framework for future studies on glue genes and glue adhesion in Diptera flies. This work also reveals new avenues of research for understanding why certain genomic regions evolve faster than others.

Methods

Fly stocks and nucleic acid extraction

To amplify part of the *Sgs1* gene, we used the following stocks: *D. rhopaloa* (line BaVi067 from Vietnam, Hanoi Ba Vì, near Vân Hòa [21°04'N, 105°22'E], collected in March 2005, gift from N. Gompel, obtained from H. Takamori), *D. takahashii* (stock number 14022 – 0311.07, isofemale line from Ulu Temburong National Park, Brunei, 2003, gift from N. Gompel). Flies were cultured at 22°C in plastic vials on standard medium [4 liters: 83.5 g yeast, 335.0 g cornmeal, 40.0 g agar, 233.5 g saccharose, 67.0 ml Moldex, 6.0 ml propionic acid]. For both species, DNA was extracted from five adults (3 males and 2 females) using Omega Bio-tek E.Z.N.A. Insect DNA Isolation Kit following the manufacturer's instructions. RNA was extracted from five adults (3 males and 2 females) using a Nucleospin RNA kit from Macherey-Nagel following manufacturer's instructions.

PCR and RT-PCR

PCR and RT-PCR
For *D. rhopaloa*, Omega Bio-tek E.Z.N.A. Insect DNA Isolation Kit was used for genomic DNA extraction. We used the following primers within the *Sgs1* repeated region and framing the observed frameshift: forward 5' ACT TGC ACC CCT CCC CCT GT 3' and reverse 5' GGA GTG CAC CCC AAC GCG AT 3'. The primer set gave a smear or shorter fragments than expected at different PCR conditions using Phusion high fidelity DNA polymerase (New England Biolabs, M0530S). We conclude that the repeated region where the primers were designed in *D. rhopaloa Sgs1* region did not allow us to successfully amplify the region of interest. Primer sets outside of the repeated region could not be used for PCR since the repeated region is close to 5kb.

For *D. takahashii* and *D. rhopaloa*, RNA was extracted from three third instar wandering larvae with Macherey Nagel Nucleospin RNA kit. A reverse transcription was then performed with the SuperScript VILO cDNA synthesis kit from Invitrogen. 200 ng of RNA were used for a reaction of 20uL. The samples were then placed 10 minutes at 25°C, 60 minutes at 42°C and 5 minutes at 85°C. PCR was then performed with Gotaq from Promega. For *D. takahashii*, the following primers were used to amplify part of the *Sgs1* sequence: forward 5' CCC GAT CCA ATG GAG CCC TGT 3' and reverse 5' GTG TCG GTG GCT GTG TCT GTA 3'. Annealing was performed at 55°C. The primers amplified a 350-bp fragment which contains an extra 'A' nucleotide in the repeated region compared to the NCBI *D. takahashii* genome sequence (accession number GCA_000224235.2). For *D. rhopaloa*, the following primers were used : forward 5' CCA CTC CTA CCC CCA TAA CT 3' and reverse 5' GGG TAG GAG TGG ATG TAG GT 3'. We obtained a smear and made the same conclusion as with the PCR results. We performed a new PCR on cDNA of *D. rhopaloa* with primers: forward 5' ACT TGC ACC CCT CCC CCT GT 3' and reverse 5' GGA GTG CAC CCC AAC GCG AT 3' (same primers as we used at first), and purified highest PCR product among several, about 5000bp long using Nucleospin Gel and PCR cleanup kit from Macherey Nagel. We did not manage to clone and sequence the purified product.

Annotation of Sgs genes

Sequence databases were searched by blastn and tblastn in a recursive manner, using the Sgs sequences of various Drosophila species. BLAST searches were performed via the NCBI BLAST page (https://blast.ncbi.nlm.nih.gov/Blast.cgi), the SpottedWingFlybase website (http://spottedwingflybase.org/) for D. suzukii or using Megablast, a variation on blastn that is faster but only finds matches with high similarity, in Geneious Prime (2019.2.3 Build 2019-09-24 10:49, Java Version 11.0.3 + 7 (64 bit)) (https://www.geneious.com/) after uploading the genomes. The coding regions were annotated manually (Table S2), using sequence homology with closely related species, conserved intron-exon structure and conserved stretches of amino acids (Fig. 1A). Peptide signals were predicted using SignalP-6.0 website (last accessed on 2022/08/24, https://services.healthtech.dtu.dk/service.php?SignalP). Annotations were then verified based on alignments of the respective protein sequences using MUSCLE (3.8.425) [57] implemented in Geneious Prime (version 2019.2.3) (https://www.geneious.com/).

Our analysis allowed us to identify 13 additional *Sgs* genes in the species previously examined by Da Lage et al.: *Sgs1* in *D. ananassae* and *D. bipectinata*; Sgs3x in *D. pseudoobscura*, *D. eugracilis*, *D. suzukii* and *D. takahashii; Sgs3* orthologs in *D. suzukii*, *D. santomea*, *D. yakuba*, *D. bipectinata*, *D. ananassae, Sgs7* in *D. ananassae* and *Sgs8* in *D. mauritiana*. We also annotated a few *Sgs* coding sequences that were absent in NCBI annotated genomes: *Sgs3e* in *D. suzukii, D. ananassae, D. eugracilis, D. takahashii, D. biarmipes, D. ananassae, D. pseudoobscura, D. bipectinata, D. elegans, D. rhopaloa, Sgs3b* in *D. ficusphila, Sgs1* in *D. ananassae, D. bipectinata, D. pseudoobscura, D. takahashii, D. suzukii, D. simulans, Sgs3x* in *D. pseudoobscura, D. eugracilis, D. suzukii, Sgs7* in *D. suzukii, D. ananassae, D. jambulina, D. bipectinata* and *Sgs8* in *D. suzukii.* We corrected gene annotations for: *Sgs1* in *D. ficusphila*, which had an intron disrupting its second exon sequence, *Sgs3e* in *D. obscura* and *D. subobscura* as they were missing the first exon and the intron, Sgs3x in *D. biarmipes* as it was missing its first intron and had a long intron in *D. pseudoobscura*.

Analysis of premature stop codons

For *D. rhopaloa Sgs1, D. ficusphila Sgs1* and *D. biarmipes Sgs3x*, premature stop codons were identified in the reference genome sequences. To examine whether they could be due to misassembly, we first BLASTed the raw reads of the respective genome sequence projects to the regions of interest and identified possible sequence corrections. Raw reads were then mapped to the coding region of interest using minimap2 (v.2.17-r941) [58] with -x map-ont parameter for nanopore reads (SRR13070618, SRR13070620) and -x splice:hq for Pacbio reads (SRR8032920). For species for which insertions were added in the corrected sequence (*D. rhopaloa, D. biarmipes*), reads were mapped to the corrected sequence whereas for *D. ficusphila* (where the sequence was corrected by removing a 'C' from a 6-bp stretch of C) reads were mapped to the published genome sequence. SAM files were converted to BAM file using samtools (v.1.6) and visualized in IGV (v.2.16.0) [59].

Figure preparation

Figures were prepared using the online tool Weblogo (version 2.8.2 (2005-09-08)) (https://weblogo.berkeley.edu/logo.cgi) [60] (Fig. 1A, S11), Geneious Prime (version 2019.2.3) (https://www.geneious.com/) (Fig. 1B, 8, 11, S1, S3, S6-7, S12-15), R version 4.1.2 (2021-11-01) (https://www.r-project.org) (Fig. 2, 4–7, 10, S4-5, S8, S10), IGV (v.2.16.0) (Fig. S2) and Inkscape 1.2.1 (2022-07-14 version) (https://inkscape.org/) for all figures.

Protein alignments and their Weblogo graphical representation

Protein alignments were done using MUSCLE (3.8.425) [57] with default settings, implemented in Geneious Prime (version 2019.2.3) (https://www.geneious.com/) with the full protein sequences. Regions with at least 30% of identity were extracted and used as input sequences for the online tool Weblogo (version 2.8.2 (2005-09-08)) (https://weblogo.berkeley.edu/logo.cgi) [60] to generate sequence logos. For Fig. 1A, *Sgs3e* from *D. ananassae* and *D. bipectinata* were excluded from the alignments given the Glycine amino acid at the phase 1 intron position for *D. bipectinata* and three successive Valine amino acids in the first exon and at phase 1 intron position for *D. ananassae*. *Sgs3bshort* was included with *Sgs7/Sgs8* sequences and *Sgs3dshort* with *Sgs3* sequences.

Phylogenetic trees

For *Sgs3, Sgs3x* and *Sgs1* orthologs, the aligned region containing the repeats was removed. Maximum Likelihood (ML) protein trees were then computed using PhyML (version 3.3.20180621) with default settings [61]. Bootstrap support was computed with 100 replications. Phylogenetic trees were drawn on R with the *read.dendrogram* function from the 'ape' package [62].

Identification and annotation of Sgs neighboring genes

To examine synteny around the *Sgs* genes, we searched for neighboring genes that tended to remain within the same locus near the *Sgs* genes in *D. yakuba*, *D. pseudoobscura*, *D. persimilis* and *D. willistoni* according to the Genomicus synteny browser (v30.01, https://www.genomicus.biologie.ens.fr/genomicus-metazoa-30.01/cgi-bin/search.pl) [63]. For the *Sgs3-Sgs7-Sgs8* gene cluster we selected the following genes: *rt*, *CG32086*, *CG7394*, *Mob2*, *Fuca*, *CG7512*, *Vha16*, *CG7551* and *CG12289*. For the *Sgs1* locus we selected: *CG3036*, *CG2831*, *hoe1*, *hoe2*, *mRpL24*, *betaggt-1* and *jet*. For the *Sgs3x* locus we selected : *AstA-R1*, *Ilp7*, *Parg*, *Mnt* and *Rala*. Sequences from *D. melanogaster* were used as BLAST queries as above to identify their homologues in other Drosophila species. When available, the NCBI gene annotations (Table S1) were collected. When no gene annotation was available or when the annotations were partial, we aligned DNA or protein sequences by using MUSCLE (see above) with global and free end gaps alignment to help in the manual annotations of the genes (Table S3 and S4). For *D. suzukii*, genes were annotated by comparison with the gene annotations of the genome of the closely related species *D. biarmipes*. *Ng* genes were found by BLAST using *D.*

melanogaster CG33500, CG33272, CG33270, CG43390, CG43391 amino-acid sequences as queries and by screening regions of interest. They were manually annotated based on start and stop codons as they are intronless. We note that other genes not found by our BLAST searches are also annotated as 'protein new glue' in several Drosophila genomes. We did not consider them in this study. Their phylogenetic relationship with the new *ng* genes we identified remains to be investigated.

Visualization of genomic region alignments with Easyfig and Genoplot

We used Easyfig (version 2.2.2) (https://mjsull.github.io/Easyfig/) [64] to compare *Sgs* genomic regions between species. As input for the EasyFig software, we used annotated genomic regions. EasyFig performs blastn searches on a one-by-one species comparison, starting from the first species, so that each sequence is used as a blast query for the next species in the list. We used the following BLAST parameters: Min. length (minimum length of blast hits to be drawn) = 0, Max. e Value (Maximum expected value of blast hits to be drawn) = 0.001, Min identity Value (Minimum identity value of blast hits to be drawn) = 0. We collected the Easyfig output files (.out) and processed them through the Genoplot package [65] (R version 4.1.2 (2021-11-01) (https://www.r-project.org)) to generate figures of sequence alignments. Genbank files were read with the function read.dna_seg from the Genoplot package. Colors and text on the figures generated with Genoplot were added with Inkscape 1.2.1 (9c6d41e410, 2022-07-14).

Dotplots

Dotplot drawing program in Geneious Prime (version 2019.2.3) (https://www.geneious.com/) was used to compare two genomic regions. We used the following parameters: High Sensitivity/Slow: sliding window, Score Matrix: Probabilistic: Weighted Ambiguous Matches, window size: 50, threshold: 100.

Repeats analysis

We examined genomic regions of 129 kb with the *Sgs* genes of interest being in the middle of the region. On Geneious Prime (version 2019.2.3) (https://www.geneious.com/), we used the FindRepeats plugin to annotate regions that are repeated at least once within a given sequence. We used the following criteria: minimum repeat length: 20 bp, maximum mismatches: 5. The repeat annotations were then transformed into bar plots representing the number of base pairs harboring repeats in adjacent windows of 1 kb using a custom-made R script (File S9).

Protein motif scanning

We used ScanProsite tool [66] (Release 2022_04 of 12-Oct-2022) (https://prosite.expasy.org/scanprosite/) to search for the protein motifs obtained from Fig. 1 against the protein sequence database given by ScanProsite. We chose ‘Option 2 - Submit MOTIFS to scan them against a PROTEIN sequence database’ and used the default settings.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by CNRS as part of the MITI interdisciplinary action, “Défi Adaptation du vivant à son environnement” and from the European Research Council under the European Community’s Seventh Framework Program (FP7/2007–2013 Grant Agreement no. 337579) to VCO. MM was supported by a PhD fellowship from “Ministère de l’Education Nationale, de la Recherche et de la Technologie” (MENRT) obtained from the BioSPC doctoral school.

Authors' contributions

VCO and MM designed the study. IN and MM performed DNA extraction, PCR and RT-PCR. MM, VCO and FB analyzed data. MM prepared the figures. VCO and MM wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We acknowledge the Cornell National Drosophila Species Stock Center and Nicolas Gompel for flies, Steven Marygold at FlyBase for information about the new glue genes. We thank Jean-Noel Lorenzi for his help with the R software. We thank Stéphane Prigent and Pierre Kerner for comments on the manuscript.

Availability of data and materials

The genome sequence assembly and annotation data used in this study can be retrieved at NCBI with the hyperlinks indicated in Table S1. Raw data, alignments and scripts are available as supplementary files associated with this article GenBank accession number (*D. takahashii* partial *Sgs1* sequence): OP857324.

References

1. Demuth JP, Hahn MW. The life and death of gene families. BioEssays. 2009;31:29–39.
2. Kondrashov FA. Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc R Soc B Biol Sci. 2012;279:5048–57.
3. Courtier-Orgogozo V, Arnoult L, Prigent SR, Wiltgen S, Martin A. Gephebase, a Database of Genotype-Phenotype Relationships for natural and domesticated variation in Eukaryotes. BioRxiv. 2019;:618371.
4. Albalat R, Cañestro C. Evolution by gene loss. Nat Rev Genet. 2016;17:379–91.
5. Lye ZN, Purugganan MD. Copy number variation in domestication. Trends Plant Sci. 2019;24:352–65.
6. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature. 2006;444:444–54.
7. Hayden S, Bekaert M, Crider TA, Mariani S, Murphy WJ, Teeling EC. Ecological adaptation determines functional mammalian olfactory subgenomes. Genome Res. 2010;20:1–9.
8. Birchler JA, Veitia RA. The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell. 2007;19:395–402.
9. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell. 2008;134:25–36.
10. Defoort J, Van de Peer Y, Carretero-Paulet L. The evolution of gene duplicates in angiosperms and the impact of protein–protein interactions and the mechanism of duplication. Genome Biol Evol. 2019;11:2292–305.
11. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. Nat Rev Genet. 2009;10:551–64.
12. Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet. 2007;41:331.
13. Xie KT, Wang G, Thompson AC, Wucherpfennig JI, Reimchen TE, MacColl AD, et al. DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. Science. 2019;363:81–4.
14. El-Mabrouk N. Predicting the Evolution of Syntenies—An. Algorithmic Rev Algorithms. 2021;14:152.
15. Walker EL, Robbins TP, Bureau TE, Kermicle J, Dellaporta SL. Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize R-r complex. EMBO J. 1995;14:2350–63.

16. Menardo F, Praz CR, Wicker T, Keller B. Rapid turnover of effectors in grass powdery mildew (Blumeria graminis). BMC Evol Biol. 2017;17:1–14.
17. Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, Bogwitz MR, et al. Copy number variation and transposable elements feature in recent, ongoing adaptation at the Cyp6g1 locus. PLoS Genet. 2010;6:e1000998.
18. Pajic P, Pavlidis P, Dean K, Neznanova L, Romano R-A, Garneau D, et al. Independent amylase gene copy number bursts correlate with dietary preferences in mammals. eLife. 2019;8:e44628.
19. Ishikawa A, Kabeya N, Ikeya K, Kakioka R, Cech JN, Osada N, et al. A key metabolic gene for recurrent freshwater colonization and radiation in fishes. Science. 2019;364:886–9.
20. Denton JF, Lugo-Martinez J, Tucker AE, Schrider DR, Warren WC, Hahn MW. Extensive error in the number of genes inferred from draft genome assemblies. PLoS Comput Biol. 2014;10:e1003998.
21. Han MV, Thomas GW, Lugo-Martinez J, Hahn MW. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol. 2013;30:1987–97.
22. Hahn MW, Han MV, Han S-G. Gene family evolution across 12 Drosophila genomes. PLoS Genet. 2007;3:e197.
23. Borne F, Kulathinal RJ, Courtier-Orgogozo V. Glue genes are subjected to diverse selective forces during Drosophila development. Genome Biol Evol. 2021;13:evab248.
24. Kim BY, Wang JR, Miller DE, Barmina O, Delaney E, Thompson A, et al. Highly contiguous assemblies of 101 drosophilid genomes. Elife. 2021;10:e66405.
25. Monier M, Courtier-Orgogozo V. Drosophila Glue: A Promising Model for Bioadhesion. Insects. 2022;13:734.
26. Da Lage J-L, Thomas GW, Bonneau M, Courtier-Orgogozo V. Evolution of salivary glue genes in Drosophila species. BMC Evol Biol. 2019;19:1–22.
27. Syed ZA, Härd T, Uv A, van Dijk-Härd IF. A potential role for Drosophila mucins in development and physiology. PLoS ONE. 2008;3:e3041.
28. Korayem AM, Fabbri M, Takahashi K, Scherfer C, Lindgren M, Schmidt O, et al. A Drosophila salivary gland mucin is also expressed in immune tissues: evidence for a function in coagulation and the entrapment of bacteria. Insect Biochem Mol Biol. 2004;34:1297–304.
29. Altschul S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.
30. Farkaš R. The complex secretions of the salivary glands of Drosophila melanogaster, a model system. Extracellular composite matrices in Arthropods.Springer; 2016. 557–600.
31. Chiu JC, Jiang X, Zhao L, Hamm CA, Cridland JM, Saelao P, et al. Genome of Drosophila suzukii, the spotted wing drosophila. G3 Genes. Genomes Genet. 2013;3:2257–71.
32. Paris M, Boyer R, Jaenichen R, Wolf J, Karageorgi M, Green J, et al. Near-chromosome level genome assembly of the fruit pest Drosophila suzukii using long-read sequencing. Sci Rep. 2020;10:1–14.

33. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol. 2017;34:1812–9.
34. Suvorov A, Kim BY, Wang J, Armstrong EE, Peede D, D'agostino ER, et al. Widespread introgression across a phylogeny of 155 Drosophila genomes. Curr Biol. 2022;32:111–23.
35. Furia M, Digilio FA, Artiaco D, Giordano E, Polito LC. A new gene nested within the dunce genetic unit of Drosophila melanogaster. Nucleic Acids Res. 1990;18:5837–41.
36. Furia M, D'Avino PP, Crispi S, Artiaco D, Polito LC. Dense Cluster of Genes is Located at the Ecdysone-regulated 3C Puff of Drosphila melanogaster. J Mol Biol. 1993;231:531–8.
37. Tokareva O, Jacobsen M, Buehler M, Wong J, Kaplan DL. Structure–function–property–design interplay in biopolymers: Spider silk. Acta Biomater. 2014;10:1612–26.
38. Martin CH, Meyerowitz EM. Characterization of the boundaries between adjacent rapidly and slowly evolving genomic regions in Drosophila. Proc Natl Acad Sci. 1986;83:8654–8.
39. Lehmann M, Wattler F, Korge G. Two new regulatory elements controlling the Drosophila Sgs–3 gene are potential ecdysone receptor and fork head binding sites. Mech Dev. 1997;62:15–27.
40. Giangrande A, Mettling C, Richards G. Sps–3 transcript levels are determined by multiple remote sequence elements. EMBO J. 1987;6:3079–84.
41. Roth GE, Wattler S, Bornschein H, Lehmann M, Korge G. Structure and regulation of the salivary gland secretion protein gene Sgs–1 of Drosophila melanogaster. Genetics. 1999;153:753–62.
42. Biyasheva A, Do T-V, Lu Y, Vaskova M, Andres AJ. Glue secretion in the Drosophila salivary gland: a model for steroid-regulated exocytosis. Dev Biol. 2001;231:234–51.
43. Shore EM, Guild GM. Closely linked DNA elements control the expression of the Sgs–5 glue protein gene in Drosophila. Genes Dev. 1987;1:829–39.
44. Hofmann A, Garfinkel MD, Meyerowitz EM. cis-acting sequences required for expression of the divergently transcribed Drosophila melanogaster Sgs–7 and Sgs–8 glue protein genes. Mol Cell Biol. 1991;11:2971–9.
45. Benton R. Multigene family evolution: perspectives from insect chemoreceptors. Trends Ecol Evol. 2015;30:590–600.
46. Li T-R, White KP. Tissue-specific gene expression and ecdysone-regulated genomic networks in Drosophila. Dev Cell. 2003;5:59–72.
47. Liu Y, Lehmann M. Genes and biological processes controlled by the Drosophila FOXA orthologue Fork head. Insect Mol Biol. 2008;17:91–101.
48. Ryuda M, Shimada K, Koyanagi R, Azumi K, Tanimura T, Hayakawa Y. Analysis of hunger-driven gene expression in the Drosophila melanogaster larval central nervous system. Zoolog Sci. 2008;25:746–52.
49. D'Avino PP, Crispi S, Polito LC, Furia M. The role of the BR-C locus on the expression of genes located at the ecdysone-regulated 3C puff of Drosophila melanogaster. Mech Dev. 1995;49:161–71.

50. Casas-Vila N, Bluhm A, Sayols S, Dinges N, Dejung M, Altenhein T, et al. The developmental proteome of Drosophila melanogaster. Genome Res. 2017;27:1273–85.
51. Crispi S, Giordano E, D'Avino PP, Peluso I, Furia M. Functional analysis of regulatory elements controlling the expression of the ecdysone-regulated Drosophila ng–1 gene. Mech Dev. 2001;100:25–35.
52. Crispi S, Giordano E, D'Avino PP, Furia M. Cross-talking among Drosophila nuclear receptors at the promiscuous response element of the ng–1 and ng–2 intermolt genes. J Mol Biol. 1998;275:561–74.
53. Lewis RV. Spider silk: ancient ideas for new biomaterials. Chem Rev. 2006;106:3762–74.
54. Casewell NR, Jackson TN, Laustsen AH, Sunagar K. Causes and consequences of snake venom variation. Trends Pharmacol Sci. 2020;41:570–81.
55. Baker RH, Corvelo A, Hayashi CY. Rapid molecular diversification and homogenization of clustered major ampullate silk genes in Argiope garden spiders. PLoS Genet. 2022;18:e1010537.
56. Dowell NL, Giorgianni MW, Kassner VA, Selegue JE, Sanchez EE, Carroll SB. The deep origin and recent loss of venom toxin genes in rattlesnakes. Curr Biol. 2016;26:2434–45.
57. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
58. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.
59. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.
60. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14:1188–90.
61. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21.
62. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289–90.
63. Nguyen NTT, Vincens P, Roest Crollius H, Louis A. Genomicus 2018: karyotype evolutionary trees and on-the-fly synteny computing. Nucleic Acids Res. 2018;46:D816–22.
64. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27:1009–10.
65. Guy L, Roat Kultima J, Andersson SG. genoPlotR: comparative gene and genome visualization in R. Bioinformatics. 2010;26:2334–5.
66. De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. 2006;34 suppl2:W362–5.

List of abbreviations

Sgs salivary gland secretion genes

Ng new glue genes

Figures

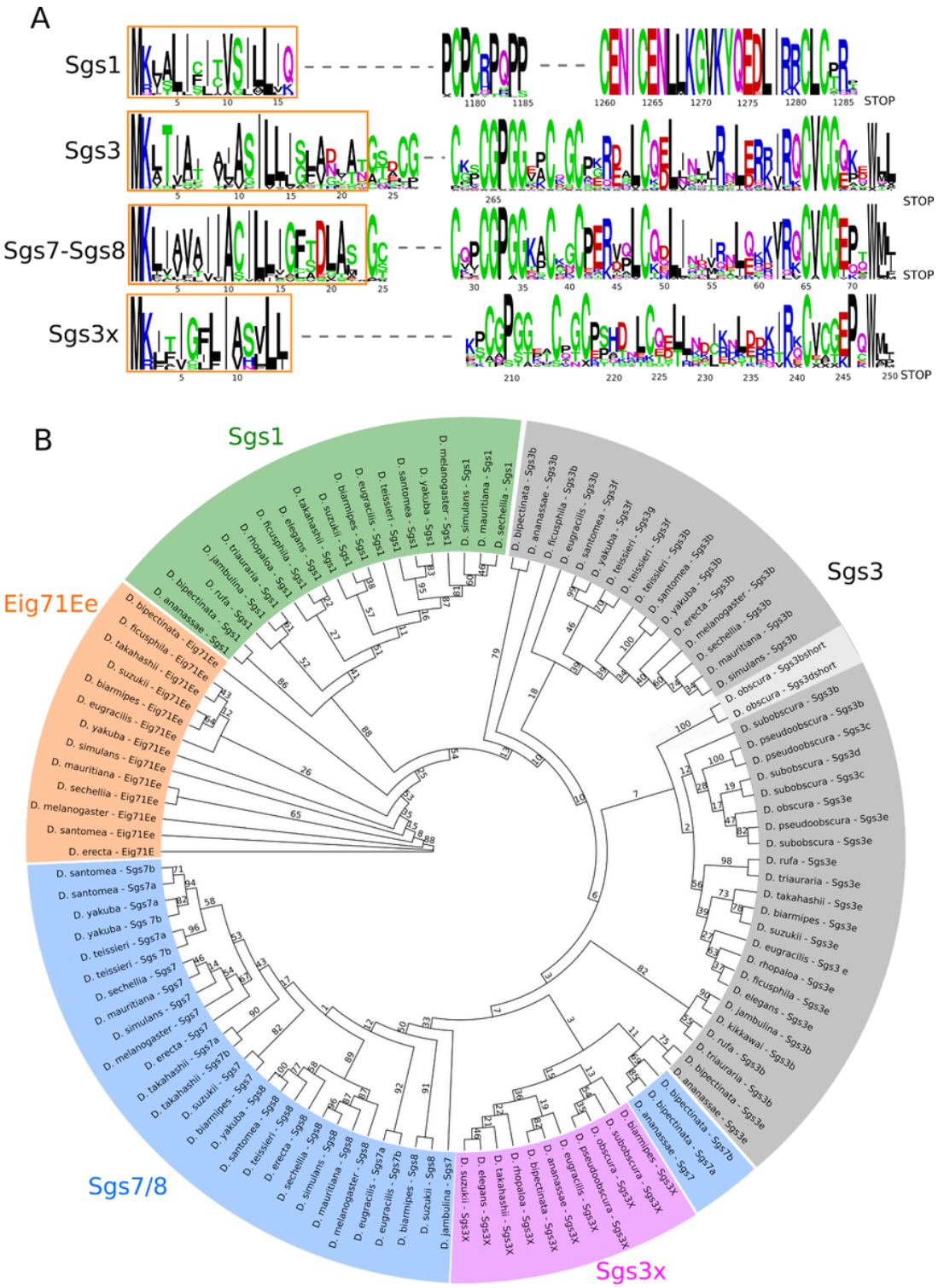


Figure 1

Overview of the Sgs1-Sgs3-Sgs7-Sgs8 protein family in Drosophila. (A) Conserved amino acid motifs in Sgs proteins. The column height indicates conservation of the sequence at that position while the height of the amino acids within the column shows relative frequency. Orange boxes indicate conserved sequences within signal peptides. Dotted lines indicate blocks of less conserved amino acid sequences. Numbers indicate the positions of the amino acid in the corresponding *D. melanogaster* protein, or in *D. suzukii* for Sgs3x as this protein is absent in *D. melanogaster*. All the *Sgs1-3-7-8* genes contain a phase 1 intron disrupting the signal peptide sequence whose position corresponds to amino acid position 10. (B) Maximum likelihood unrooted tree of Eig71Ee, Sgs1, Sgs3, Sgs3x, Sgs7 and Sgs8 amino acid sequences from all studied species. Gene names and colors were attributed based on synteny information (see text for details). Numbers on branches represent bootstrap values. Note that most bootstrap values are low, due in part to the small number of amino acids composing the Sgs7 and Sgs8 proteins.

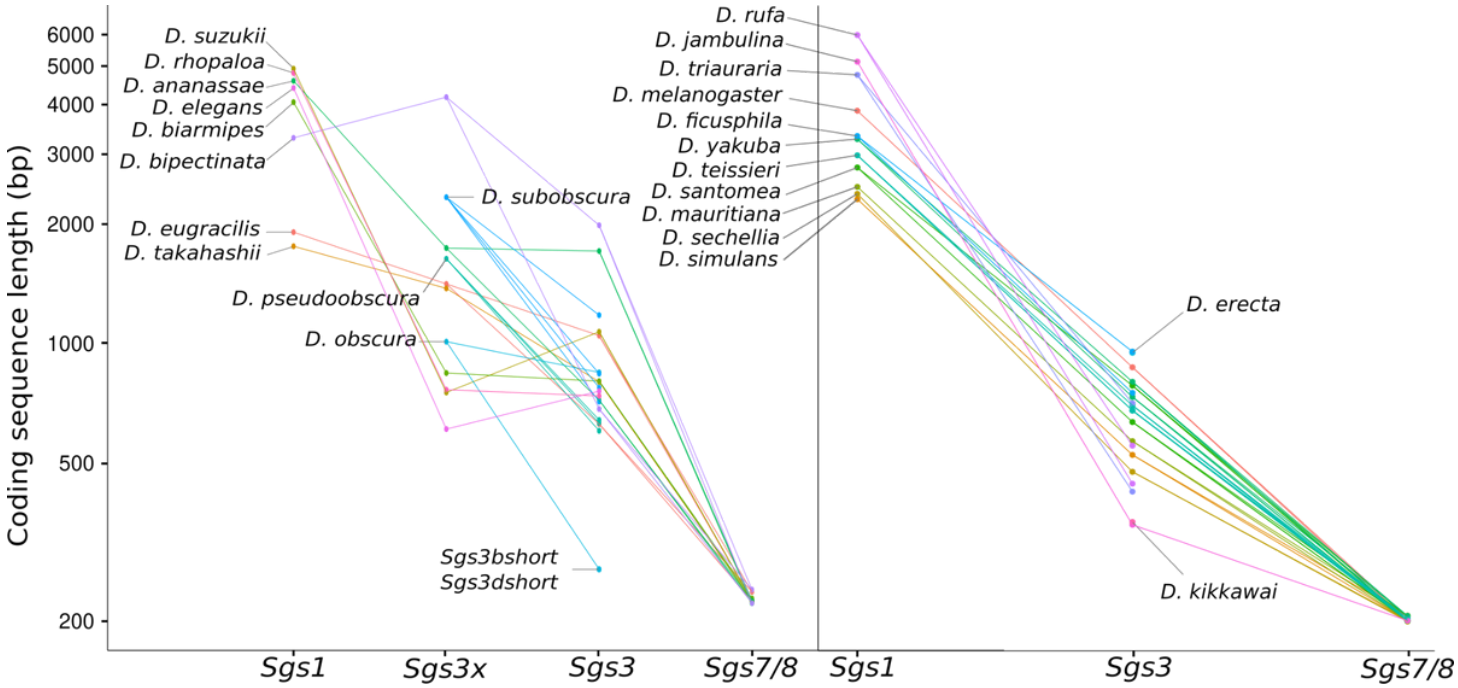


Figure 2

Length of *Sgs*coding sequences (with introns excluded). The y-axis is in log10 scale. Left: species which possess an *Sgs3x* gene. Right: species devoid of *Sgs3x*genes. All the 24 Drosophila species analyzed in this study are shown. For *Sgs1* in *D. triauraria*, *D. rhopaloa, D. ficusphila* and *D. takahashii* and *Sgs3x* in *D. biarmipes*, the length of the coding region was calculated as if the premature stop codons were artifacts.

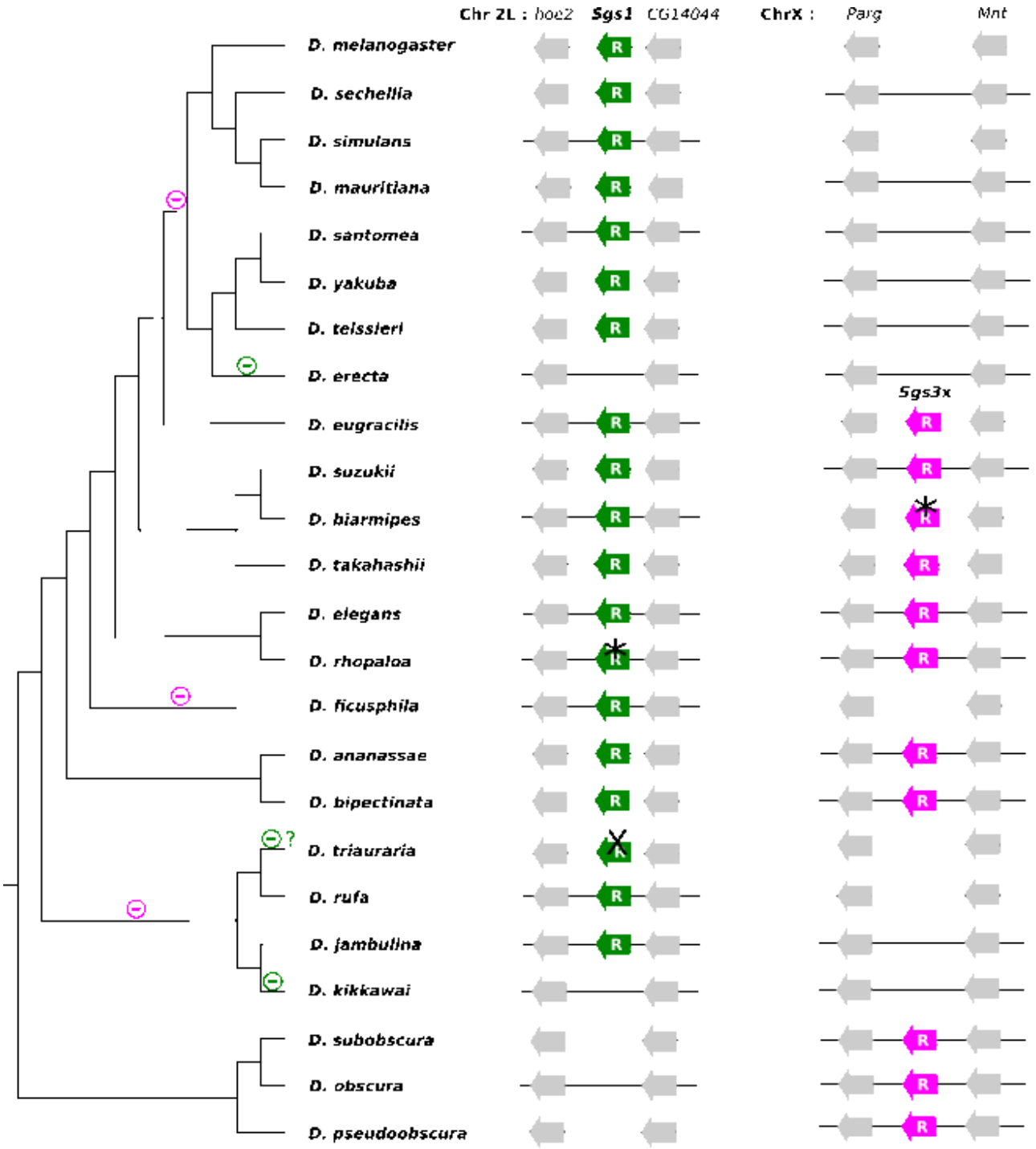


Figure 3

Distribution of the *Sgs1* and *Sgs3X* genes across the 24 studied Drosophila species and most parsimonious scenario for gene gains and losses. The species tree is from [32]. Branch distances are not on scale. Green, pink and gray arrows represent, respectively, *Sgs1, Sgs3X* and their adjacent neighboring genes. Gene lengths and intergenic distances are not to scale. “R” means that internal repeats are present. The cross 'X' on top of the *D. triauraria Sgs1* gene indicates the presence of six premature stop codons in the published genome sequence, which may be genuine stop codons or sequence assembly artifacts. * indicates a premature stop codon present in the published coding sequence of *D. rhopaloa*,

which we consider as an artifact (see text for details). Minus signs on tree branches indicate gene deletion events for *Sgs1* in green and for *Sgs3X* in pink. Minus sign followed by '?' indicates a presumed loss of a functional gene coding region that has not been confirmed by resequencing.

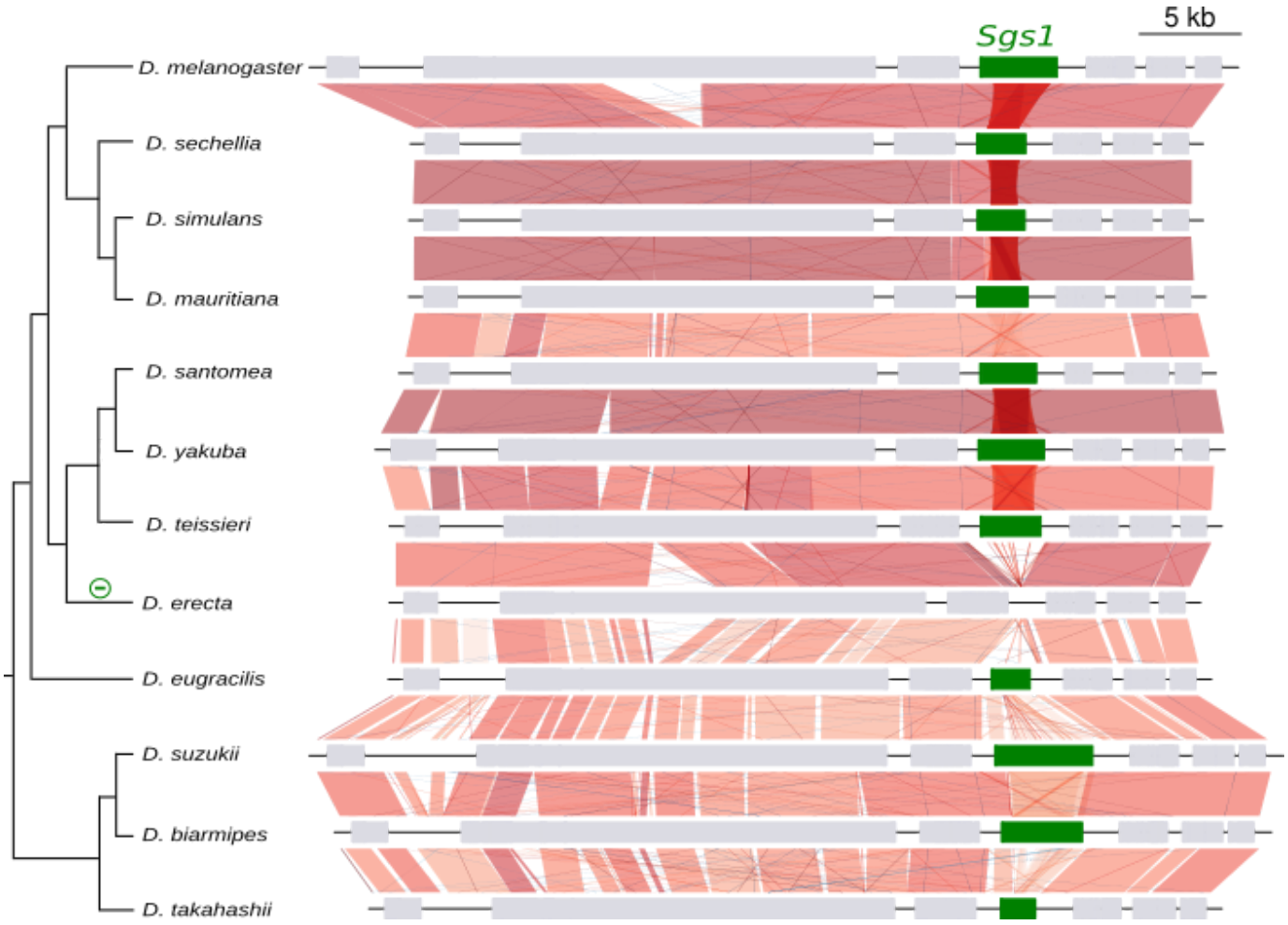


Figure 4

Comparison of the *Sgs1*gene region between Drosophila species closely related to *D. melanogaster*. The species tree is from [34]. Branch distances are not on scale. Boxes represent coding genes. *Sgs1* is in green and its neighboring genes in light gray. Introns and gene orientation are not shown. Vertical and diagonal lines between genomic sequences represent the pairwise similarity based on BLASTn analyses. They are red when BLASTn matches in the same direction and blue when BLASTn matches in the opposite direction. Shades of red and blue indicate the level of identity, with darker color for higher similarity. The minus sign on the tree branches indicates a gene deletion event.

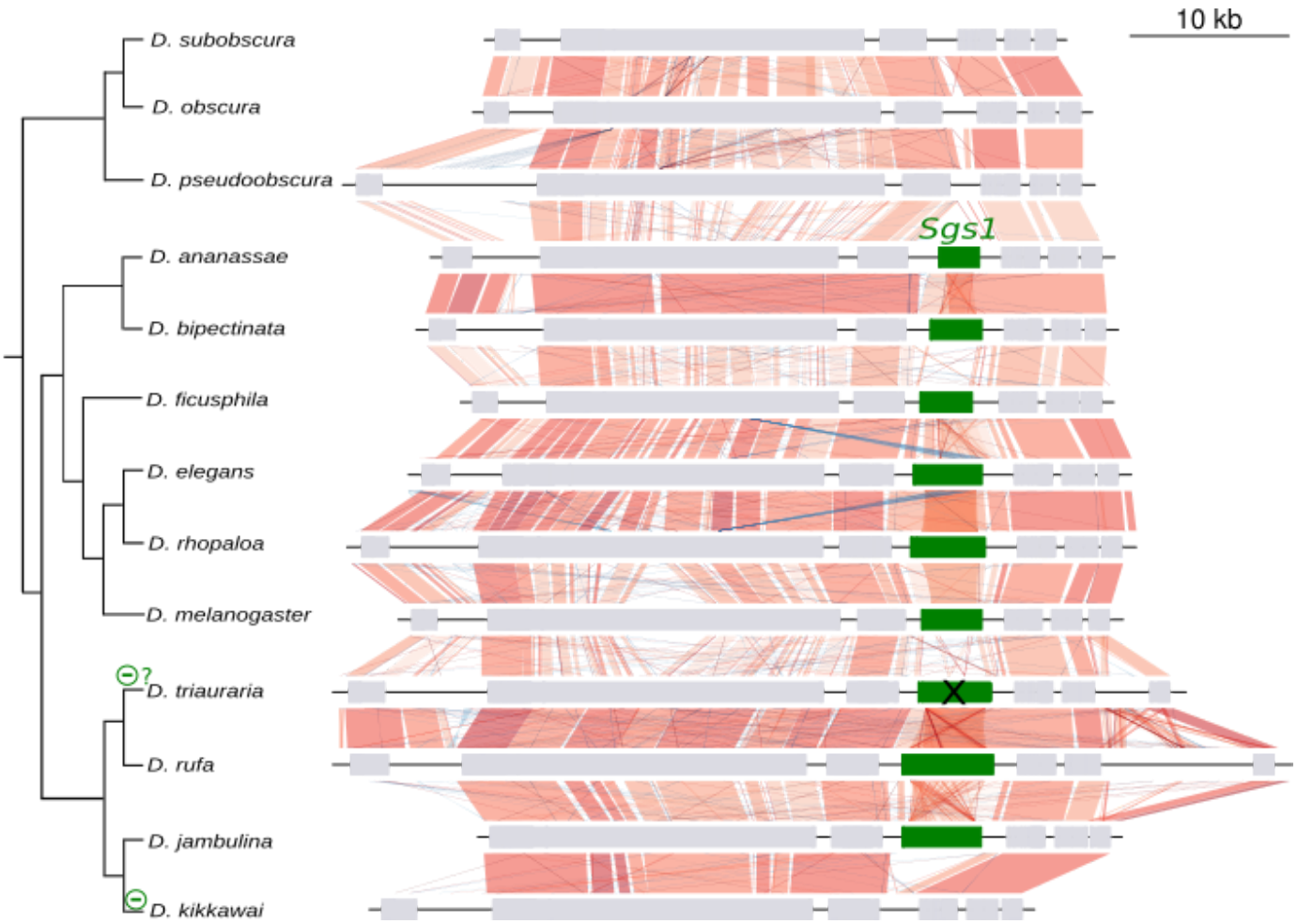


Figure 5

Comparison of the *Sgs1* gene region between Drosophila species. Same legend as in Fig. 4. The cross on top of the *D. triauraria Sgs1* gene indicates the presence of six premature stop codons and frameshifts in the published *Sgs1* gene sequence, which may be real or sequence assembly artifacts (see text for details).

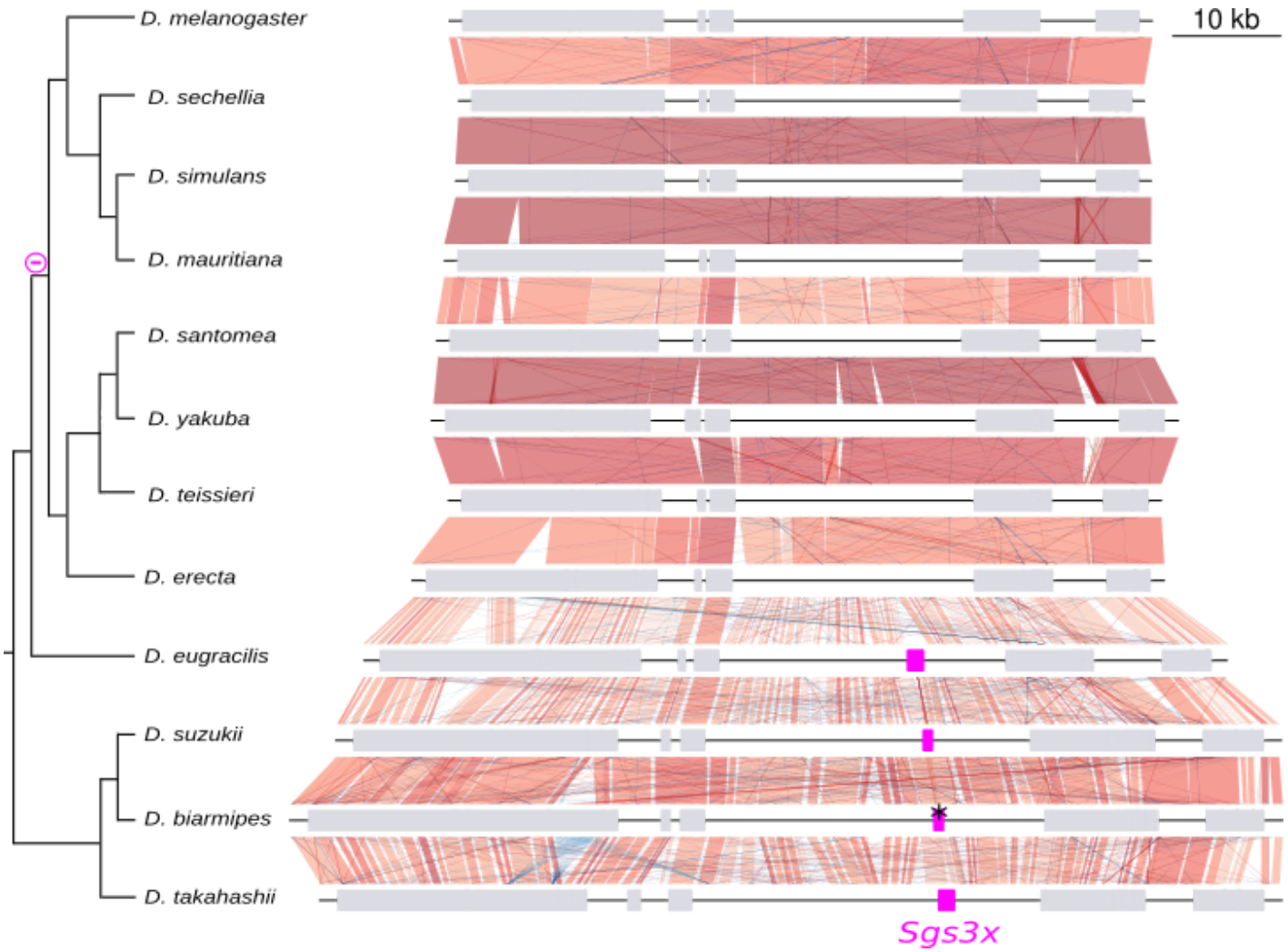


Figure 6

Comparison of the *Sgs3x* gene region between Drosophila species. Same legend as in Fig. 4. Pink boxes represent *Sgs3x*. * indicates a premature stop codon present in the published coding sequence of *D. biarmipes*, which we consider as an artifact (see text for details).

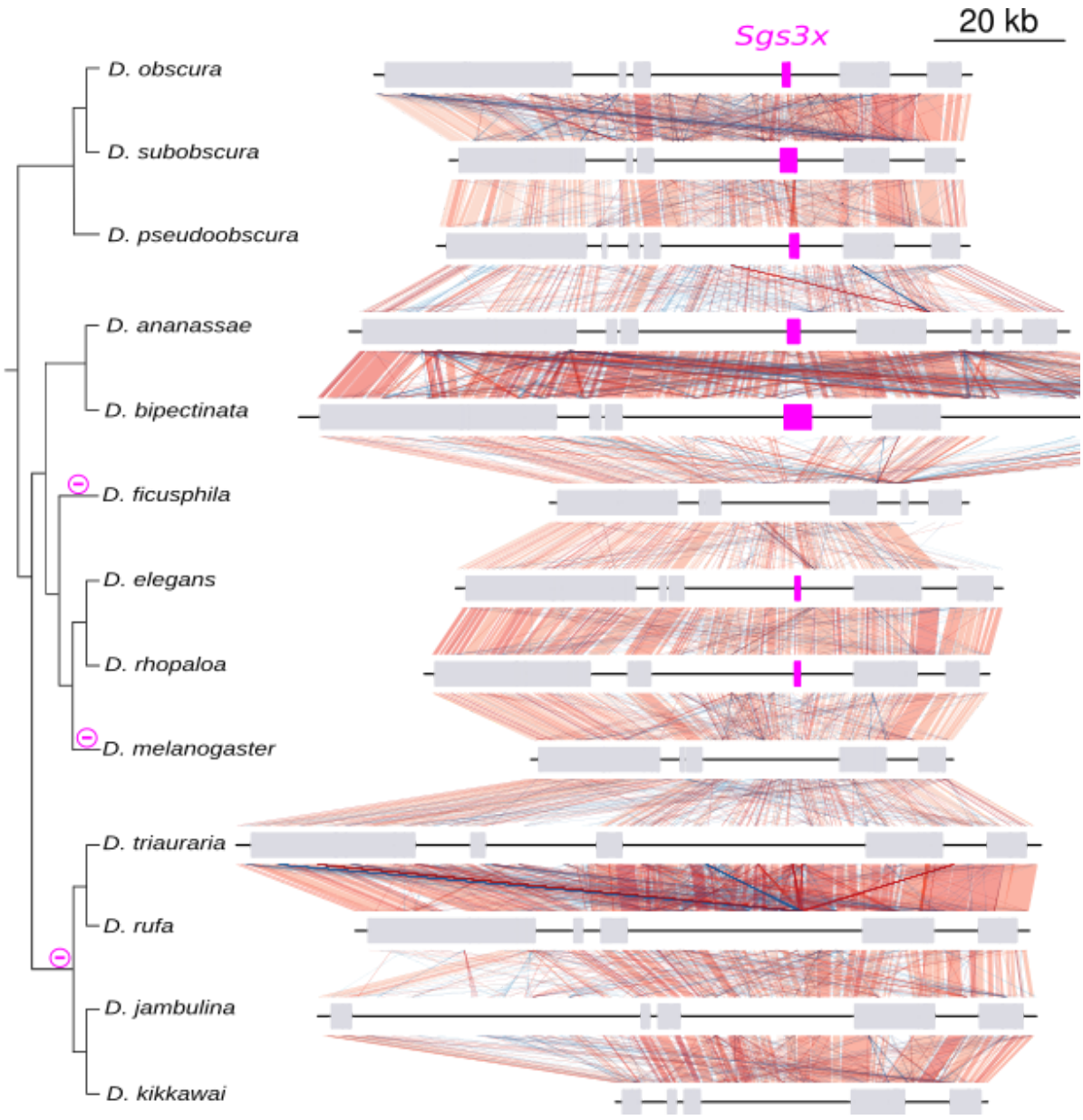


Figure 7

Comparison of the *Sgs3x* gene region between distantly related Drosophila species. Same legend as in Fig. 4. Pink boxes represent *Sgs3x.* Part of the genomic region of *D. bipectinata*(right) has been cut for clarity.

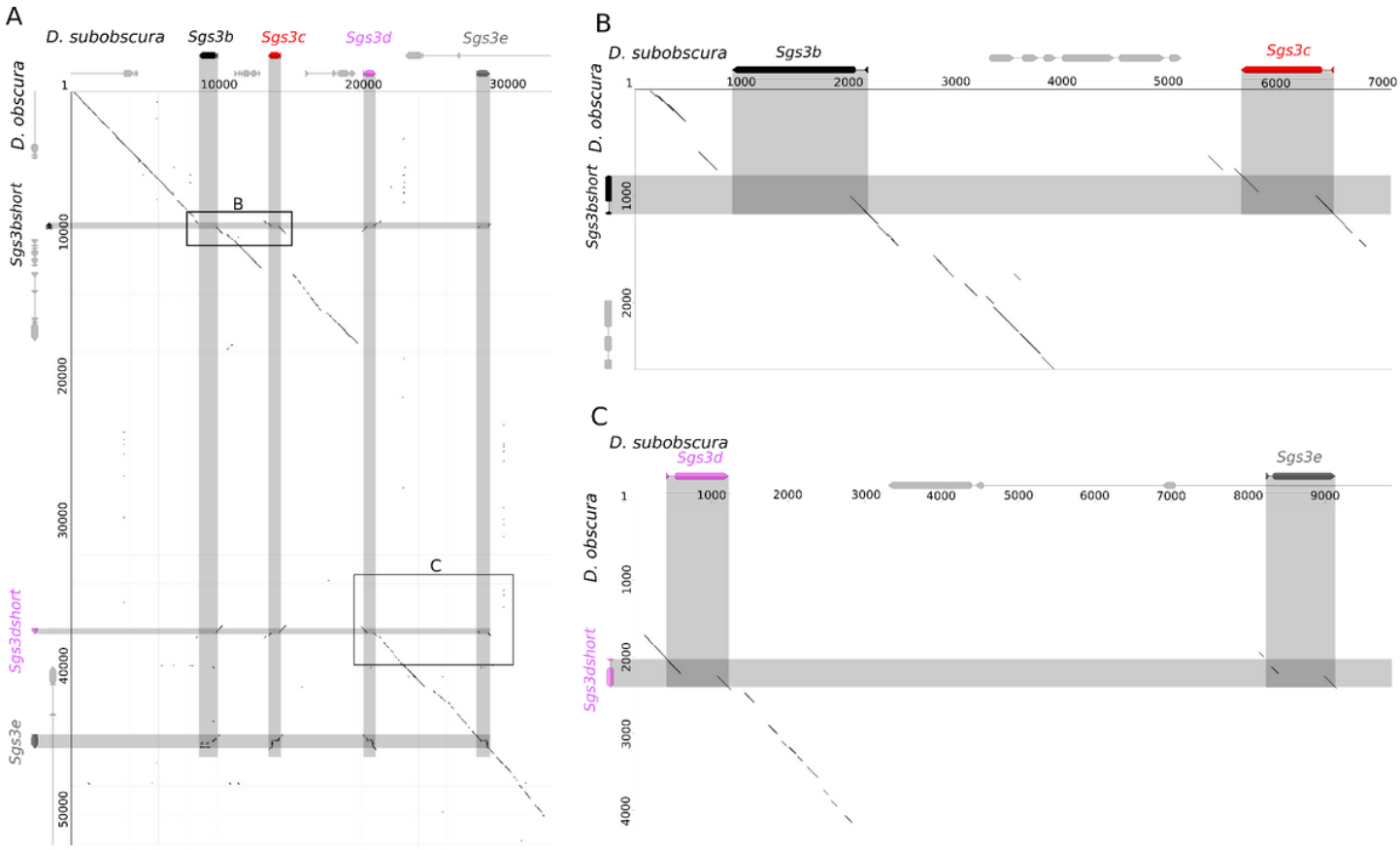


Figure 8

Dot plot comparing *D. subobscura* and *D. obscura Sgs3* genomic regions. (A) Main dot plot. (B-C) Magnifications of the regions of interest indicated in (A). Black diagonal lines indicate matching genomic regions. Black, red, pink, and dark gray arrows represent, respectively, *Sgs3b, Sgs3c, Sgs3d* and *Sgs3e.* Light gray arrows represent neighboring genes. Numbers indicate nucleotide positions in bp.

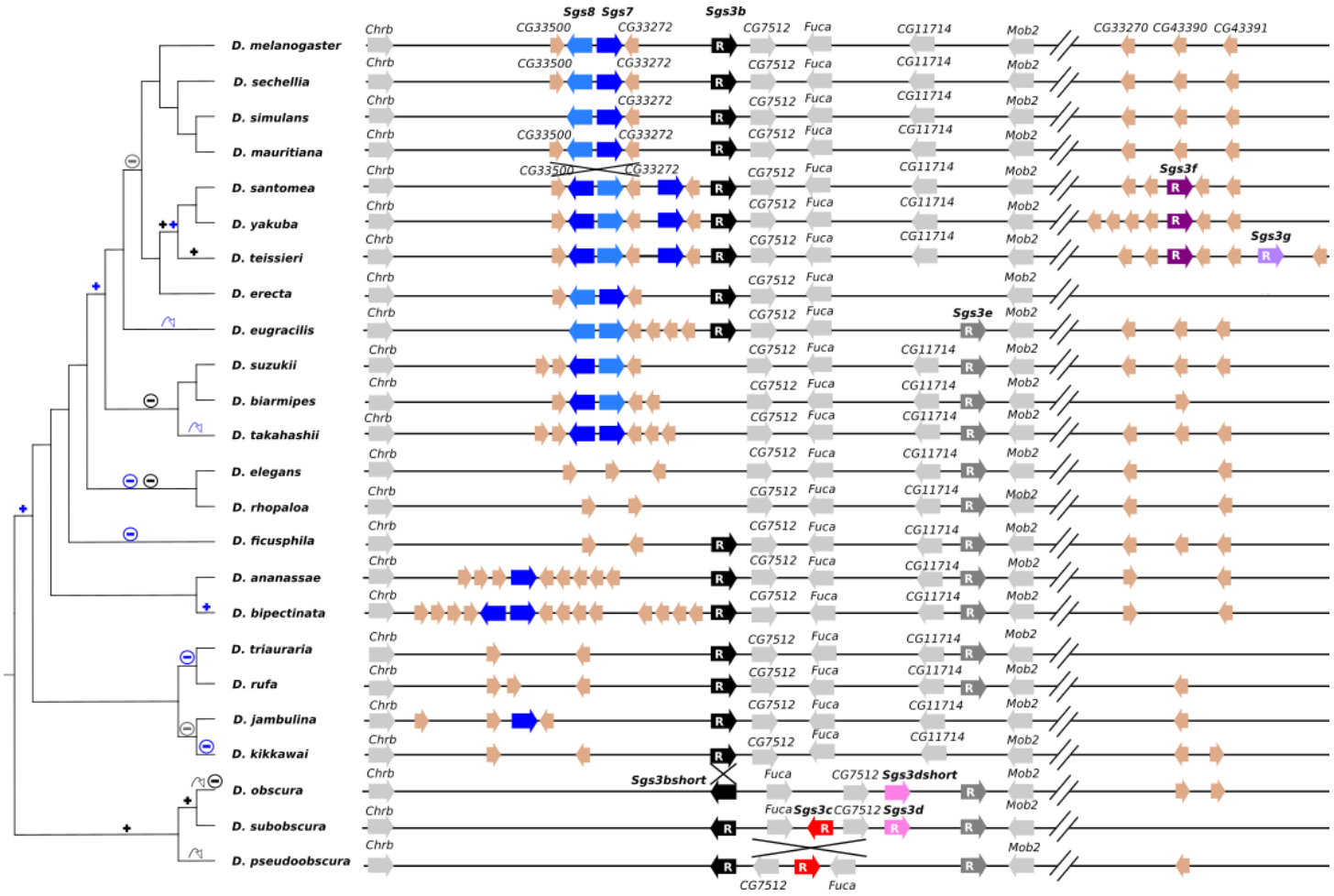


Figure 9

Distribution of the *Sgs3, Sgs7* and *Sgs8* ortholog genes across the 24 studied Drosophila species and the most parsimonious scenario for gene gains and losses. Same legend as in Fig. 3. Black, red, pink, dark gray, dark purple and light purple arrows represent different copies of *Sgs3* (respectively *Sgs3b, Sgs3c, Sgs3d, Sgs3e, Sgs3f, Sgs3g*). Dark blue and light blue arrows represent *Sgs7* and *Sgs8,* respectively. Here we present one proposition for the attribution of the names *Sgs7* and *Sgs8* to the short *Sgs* glue genes, but the distinction between *Sgs7* and *Sgs8* can be unclear. Beige arrows indicate genes encoding short threonine-rich proteins. Light gray arrows indicate other adjacent neighboring genes. The *Sgs3e* coding sequence is located within *Mob2* intron, but is represented near *Mob2* for simplicity. Also note that for clarity a few of the neighboring genes and their corresponding orthologs were omitted in this figure. Arrows, minus and plus signs on the tree branches indicate, respectively, gene conversion, gene deletion and duplication events for *Sgs3* in black and for *Sgs7* and *Sgs8* in blue. Crosses designate inversions. Double lines interrupting the genomic sequence indicate a gap of about 50 kb. Here we assumed that the most recent common ancestor of all represented species had two *Sgs3* copies, *Sgs3b* and *Sgs3e*.

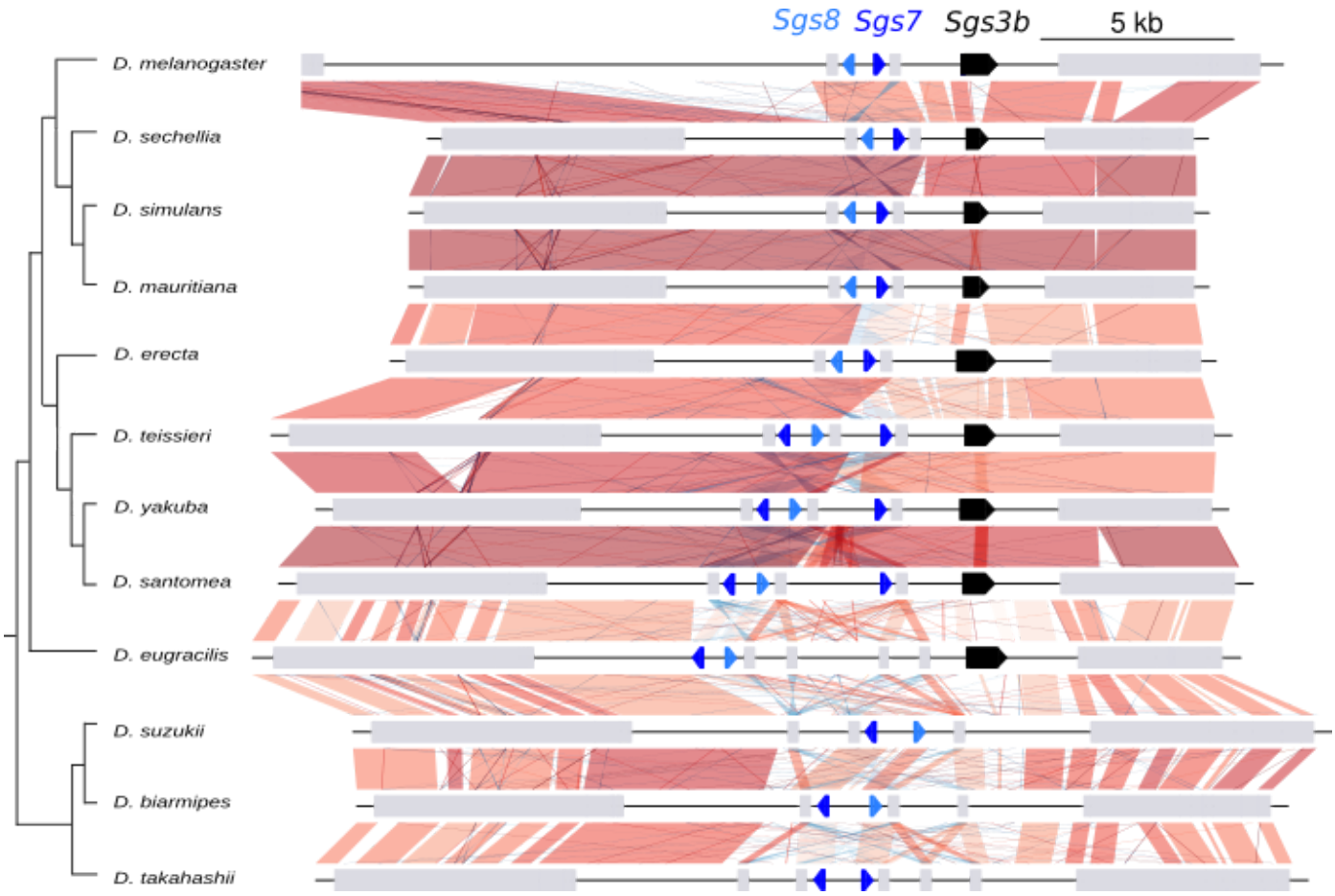


Figure 10

Closer view of the comparison of the Sgs3-7-8 gene region between Drosophila species. Same legend as in Fig. 5. Sgs7 copies are in dark blue, Sgs8 in light blue. Note that our distinction between Sgs7 and Sgs8 is subject to caution (see text for details). Sgs3b is represented in black. Sgs genes directions are given by arrows. Neighboring genes directions are not shown.

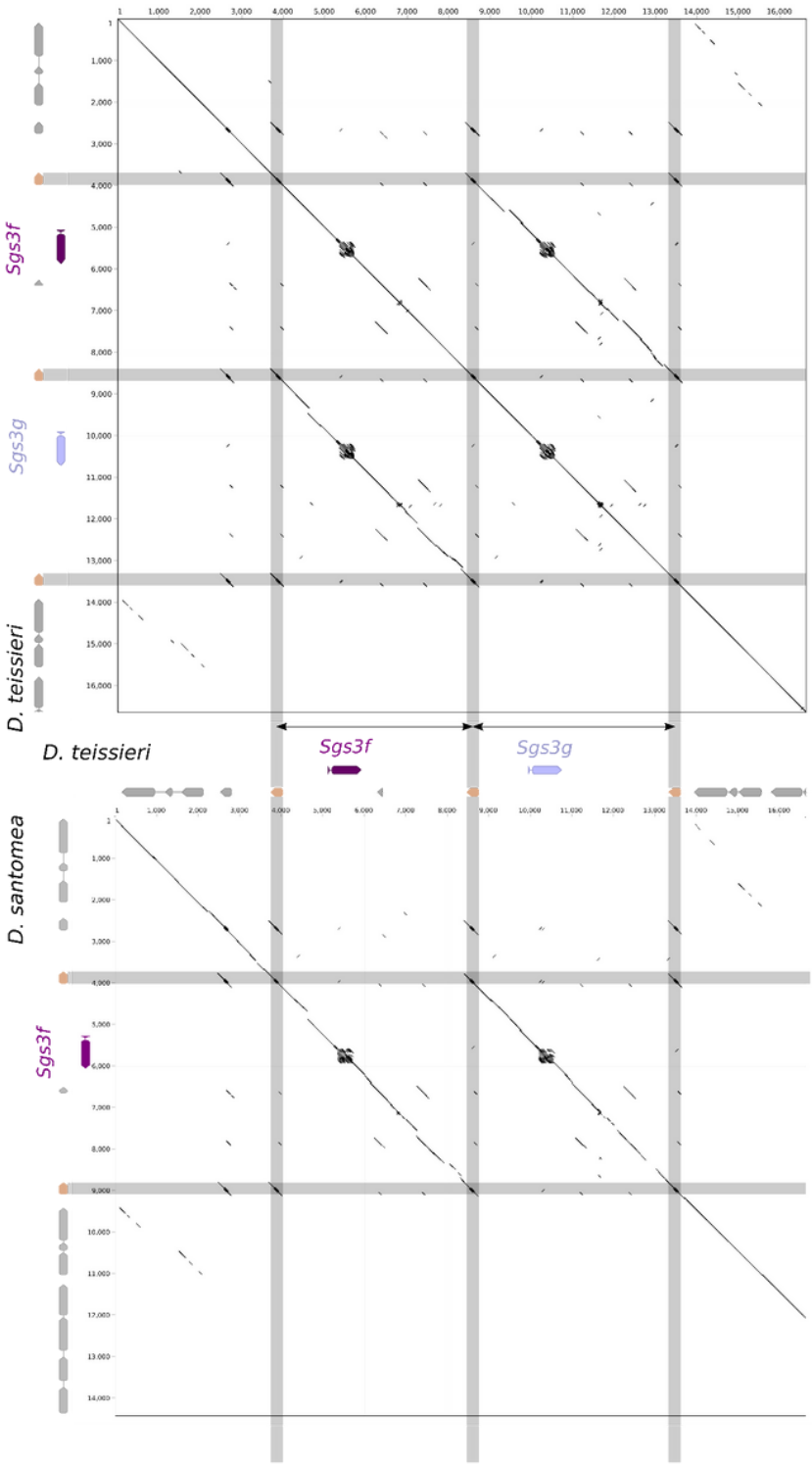


Figure 11

Dot plots of genomic regions from D. santomea and D. teissieri. In the upper dotplot, D.teissieri Sgs3f/Sgs3g genomic region is compared to itself. In the lower dotplot, D.teissieri Sgs3f/Sgs3g genomic region is compared to D. santomea Sgsf genomic region. Dark and light purple arrows represent Sgs3f and Sgs3g, respectively. Grey arrows represent neighboring genes. Beige arrows represent ng genes located at the duplication breakpoints. Double-headed black arrows indicate the duplicated region.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.csv
- TableS2.csv
- TableS3.zip
- manuscriptsupp.pdf
- Suppdatafiles114.zip

2.2.1 Supplementary figures

Higher evolutionary dynamics of gene copy number for Drosophila glue genes located near short repeat sequences

Supplementary Figures

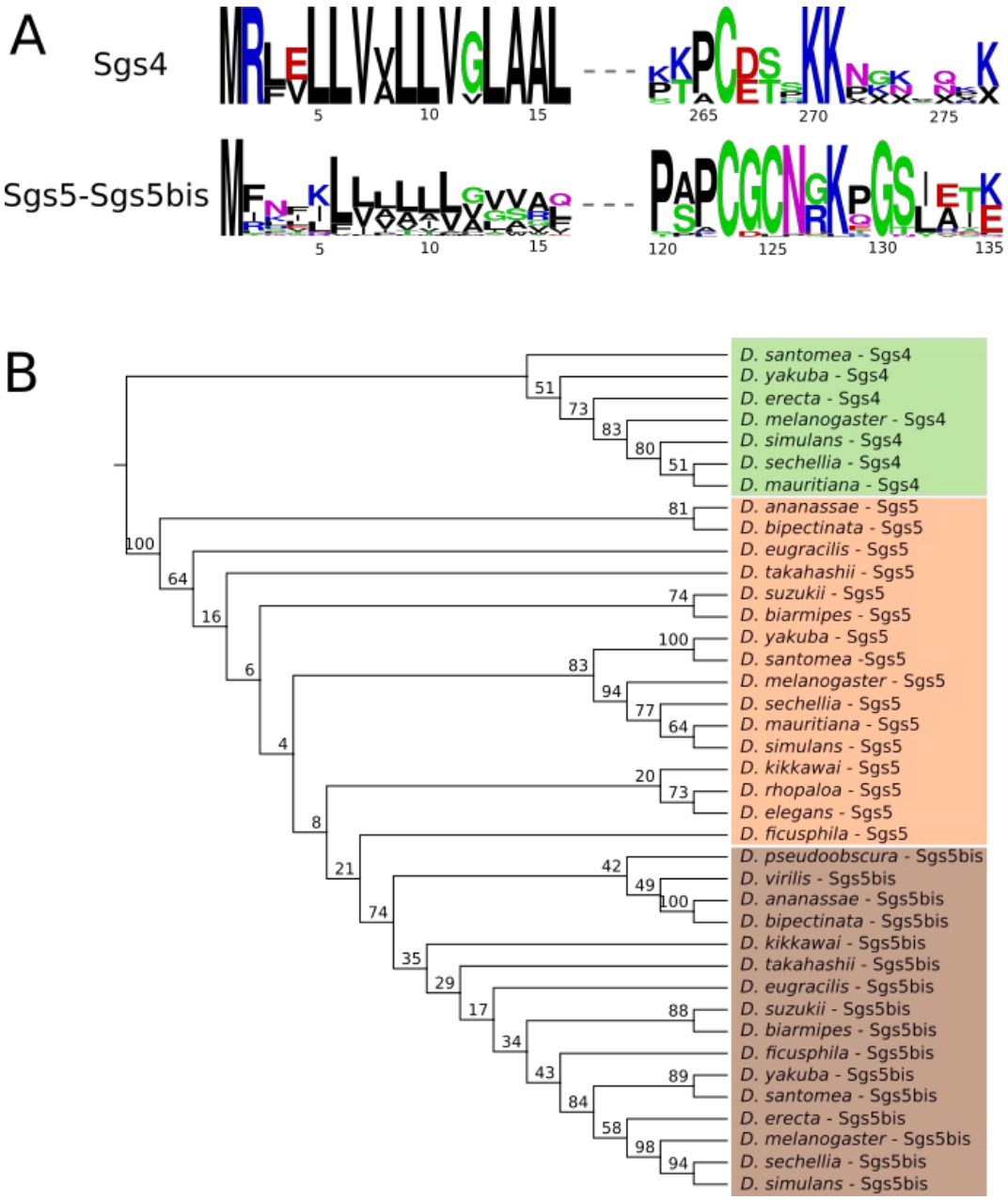


Fig. S1. Overview of the Sgs4, Sgs5 and Sgs5bis proteins in Drosophila. (A) Conserved amino acid motifs in Sgs4 and Sgs5-Sgs5bis proteins. Same legend as Fig. 1A. (B) Maximum likelihood (ML) tree of aligned, full Sgs4, Sgs5 and Sgs5bis amino acid sequences. Numbers along branches are bootstrap values. The tree was rooted between the Sgs4 cluster and the Sgs5-Sgs5bis cluster.

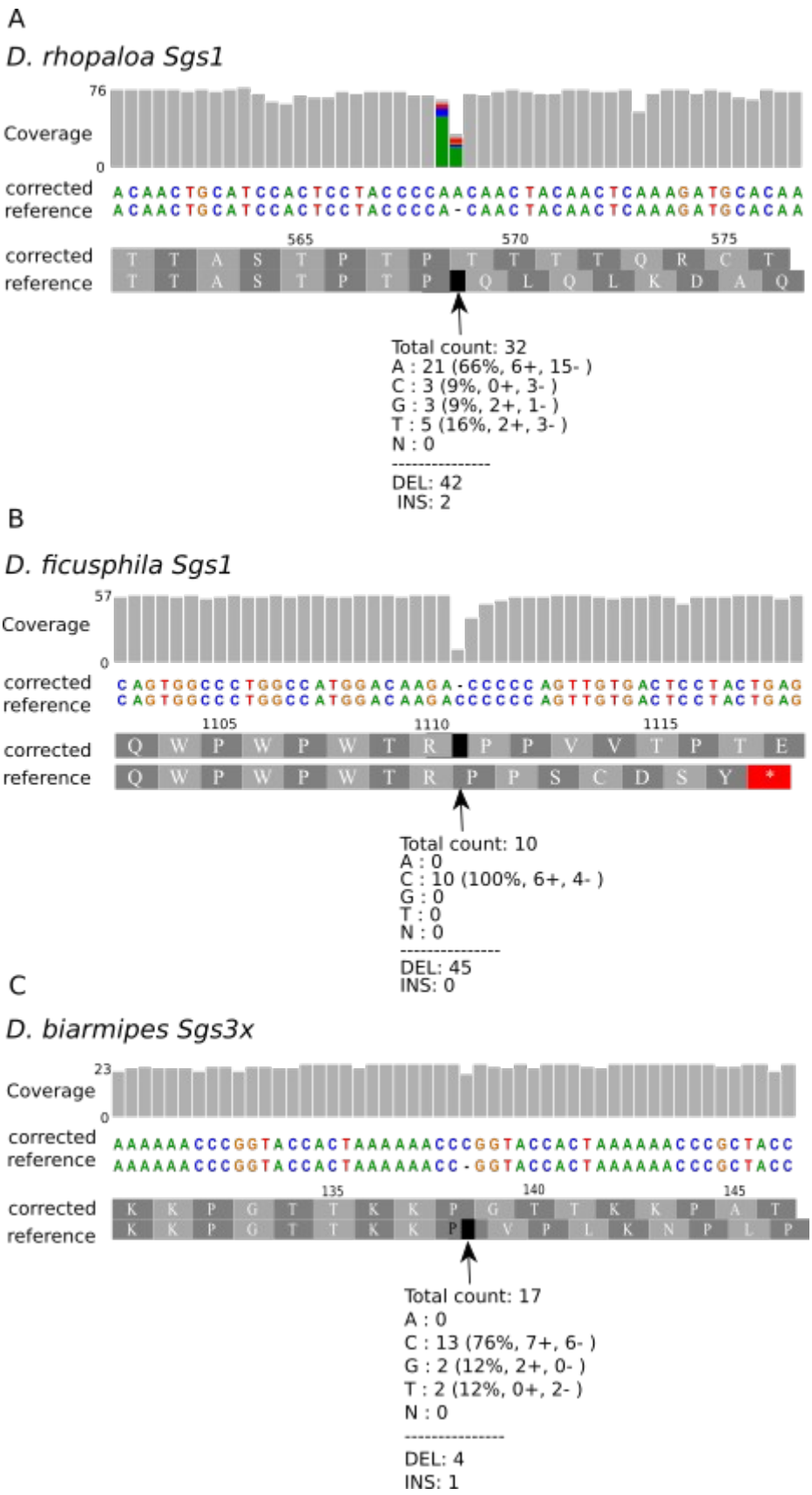


Fig. S2. Analysis of premature stop codons in three Sgs genes. Raw reads from respective full genome sequencing projects suggest that (A) D. rhopaloa Sgs1 nucleotide reference sequence should be corrected by adding an ‘A’ nucleotide, (B) D. ficusphila Sgs1 nucleotide sequence by deleting a ‘C’ and (C) D. biarmipes Sgs3x nucleotide sequence by adding a 'C'. In each panel, top gray bars represent the coverage of raw reads mapped to the corrected

(A,C) or reference (B) sequence. Reference and corrected sequences are indicated below. The distribution of nucleotides and indels at the site of interest is presented below. DEL: deletion; INS: insertion.

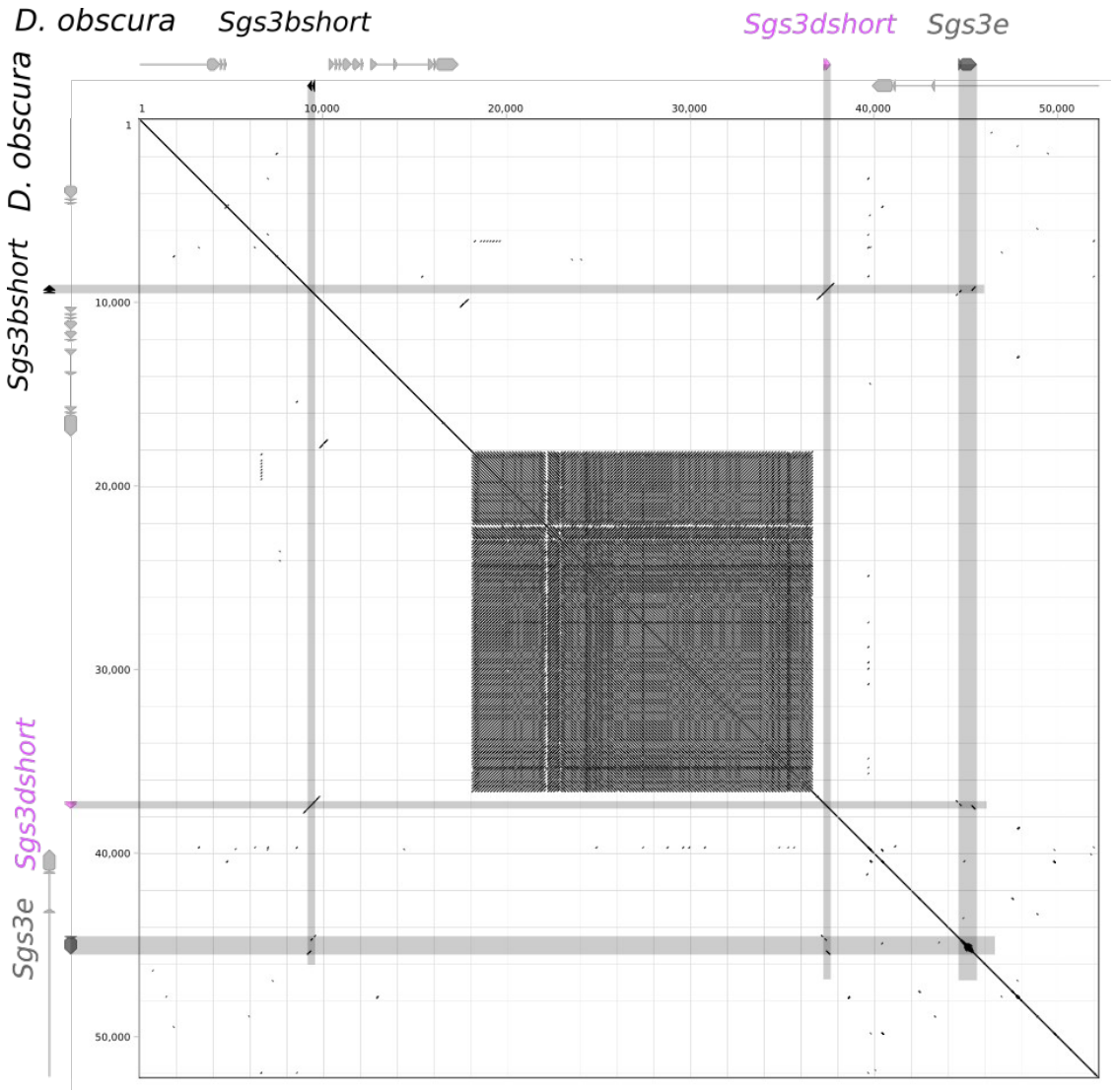


Fig. S3. Dot plot comparison of *D. obscura Sgs3* genomic region with itself. Black, light pink and dark grey arrows represent, respectively, *Sgs3bshort*, *Sgs3dshort* and *Sgs3e*. Light grey arrows represent neighboring genes. Numbers indicate nucleotide positions in bp.

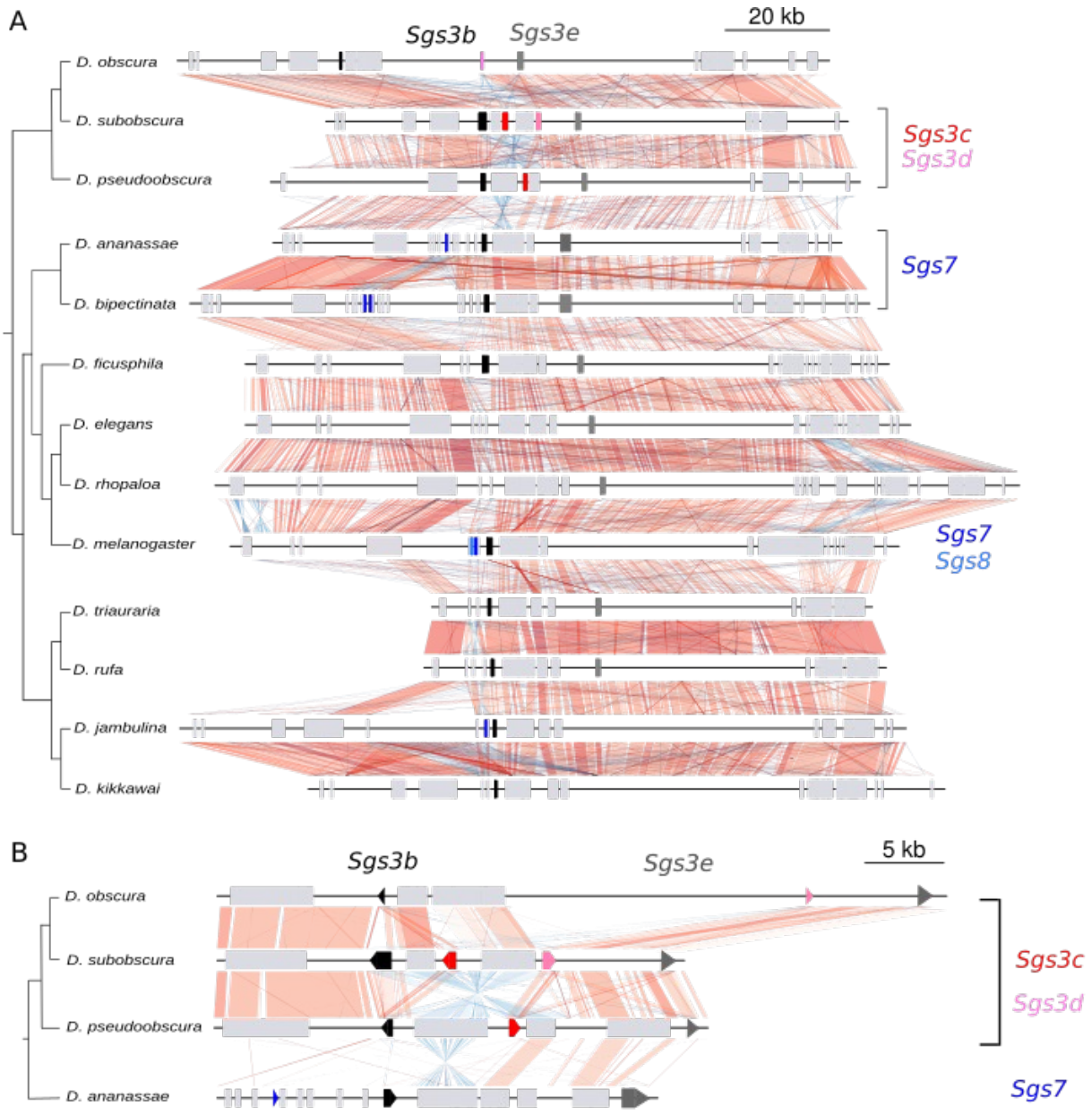


Fig. S4. Comparison of the *Sgs3-7-8* gene region between Drosophila species. (A) Entire locus comprising *Sgs3*, *Sgs7* and *Sgs8* genes. (B) Magnification of the *Sgs3b-Sgs3e* region. Same legend as in Fig. 5. *Sgs7* copies are in dark blue, *Sgs8* in light blue. *Sgs3b, Sgs3c, Sgs3d, Sgs3e* copies are respectively represented in black, red, light pink, dark gray. *Sgs3bshort* and *Sgs3dshort* are respectively in black and light pink and are only found in *D. obscura*. *Mob2* gene has been removed from this representation for the sake of clarity because it is superposed with *Sgs3e*.

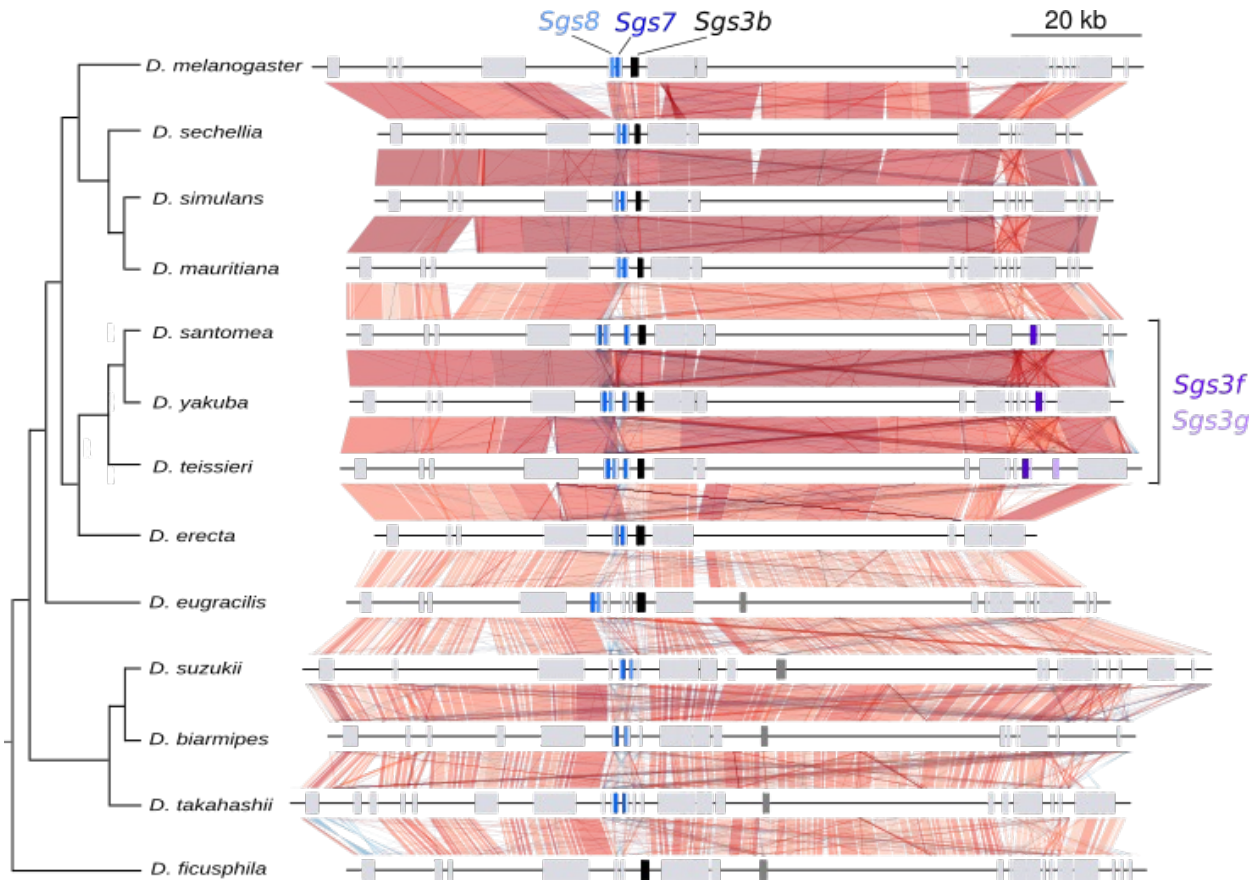


Fig. S5. Comparison of the *Sgs3-7-8* gene region between Drosophila species. Same legend as in Fig. 5. *Sgs7* copies are in dark blue, *Sgs8* in light blue. *Sgs3b*, *Sgs3e*, *Sgs3f* and *Sgs3g* copies are represented in black, dark grey, dark and light purple, respectively.

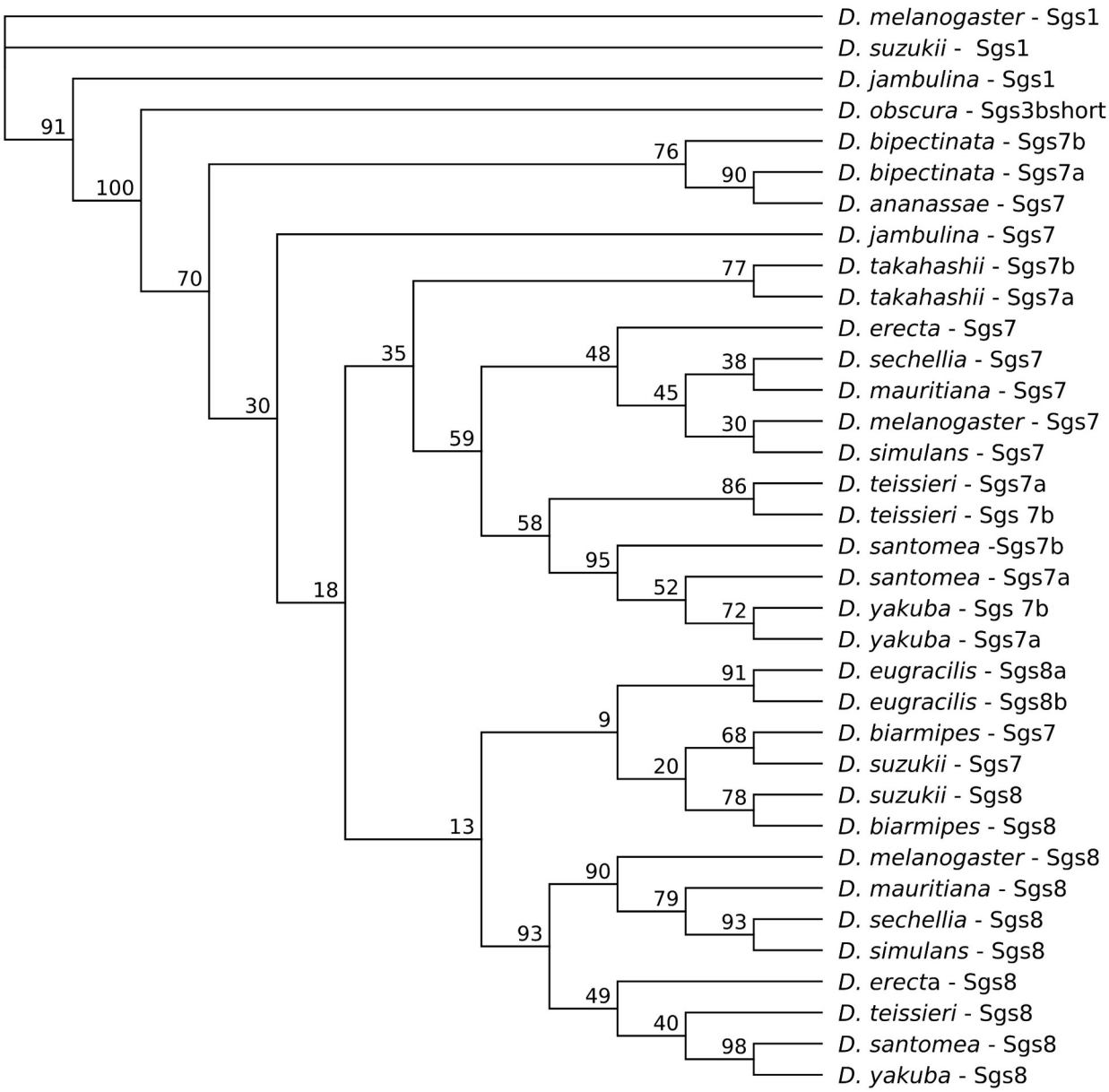


Fig. S6. Maximum likelihood (ML) unrooted tree (amino acid sequences) for *Sgs7* and *Sgs8* from all studied species and using a few outgroup sequences (*Sgs1* from *D. melanogaster, D. suzukii, D. jambulina and Sgs3bshort* from *D. obscura).* Numbers along branches indicate bootstrap values.

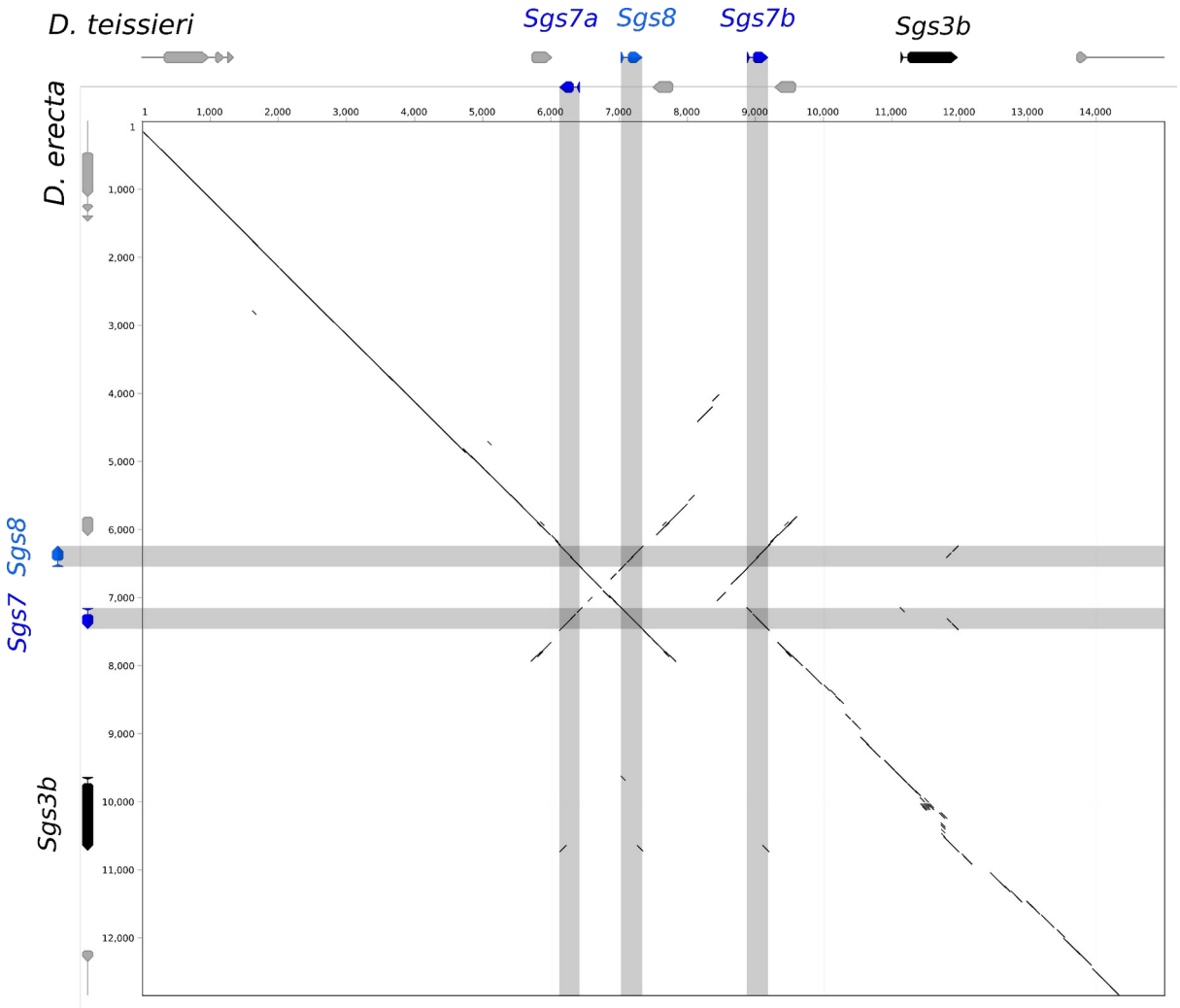


Fig. S7. Dot plot comparison of *Sgs3-7-8* genomic regions from *D. erecta* and *D. teissieri*. Black diagonal lines indicate matching genomic regions. Black, dark blue and light blue arrows represent, respectively, *Sgs3b, Sgs7* and *Sgs8* orthologs. Same legend as Fig. S1.

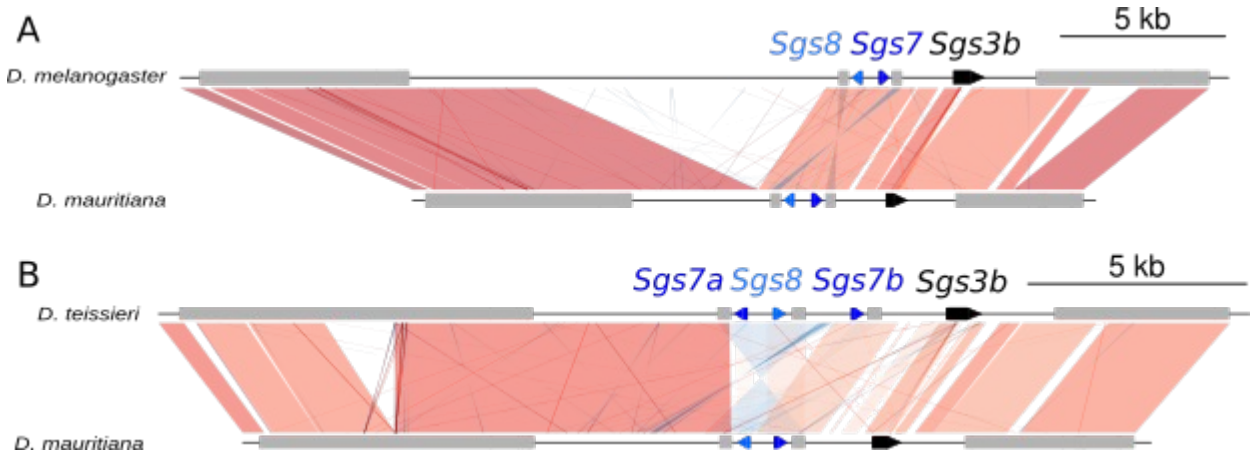


Fig. S8. Comparison of the *Sgs3-7-8* genomic region between *D. melanogaster, D. mauritiana* and *D. teissieri*. Same legend as in Fig. 4 (A) The orientation and position of *Sgs7, Sgs8 and Sgs3b is similar between D. melanogaster* and *D. mauritiana.* (B) The *Sgs3-7-8* genomic region experienced a gene duplication of *Sgs7* and an inversion affecting *Sgs7a* and *Sgs8* genes (light blue hourglass shape) in the lineage leading to *D. teissieri* (see also Fig. 9). One of the breakpoints of the inversion is a *ng* gene (gray box).

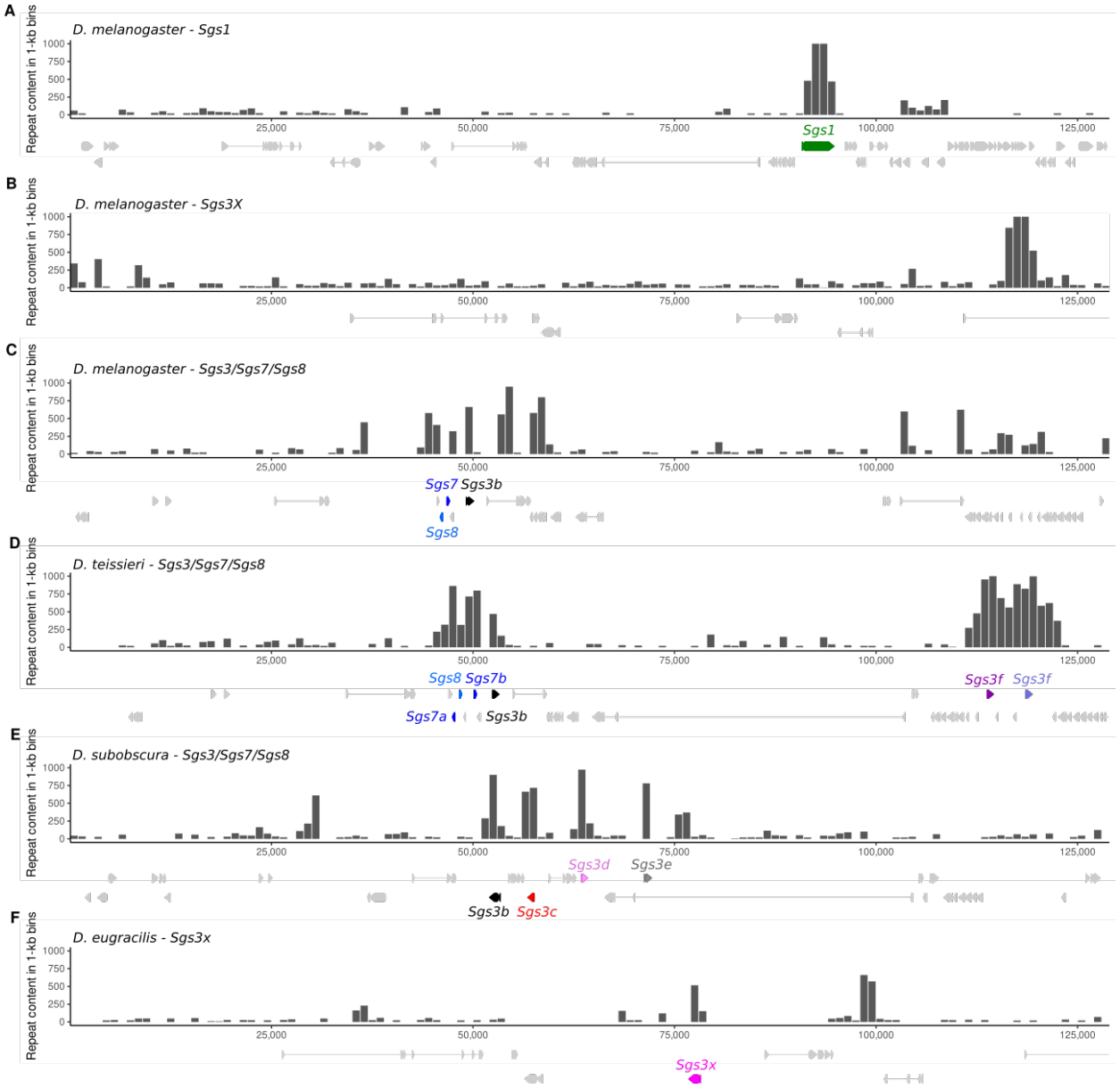


Fig. S9. Amount of repeats in 129-kb *Sgs* genomic sequences of several Drosophila species. Each bar represents the number of nucleotides within a 1-kb window that are annotated as repeats. The annotations of *Sgs* genes and neighboring genes are displayed with arrows. Sgs3x is absent in *D. melanogaster*. Note that internal repeats present within the coding regions of *Sgs* genes are annotated as repeats.

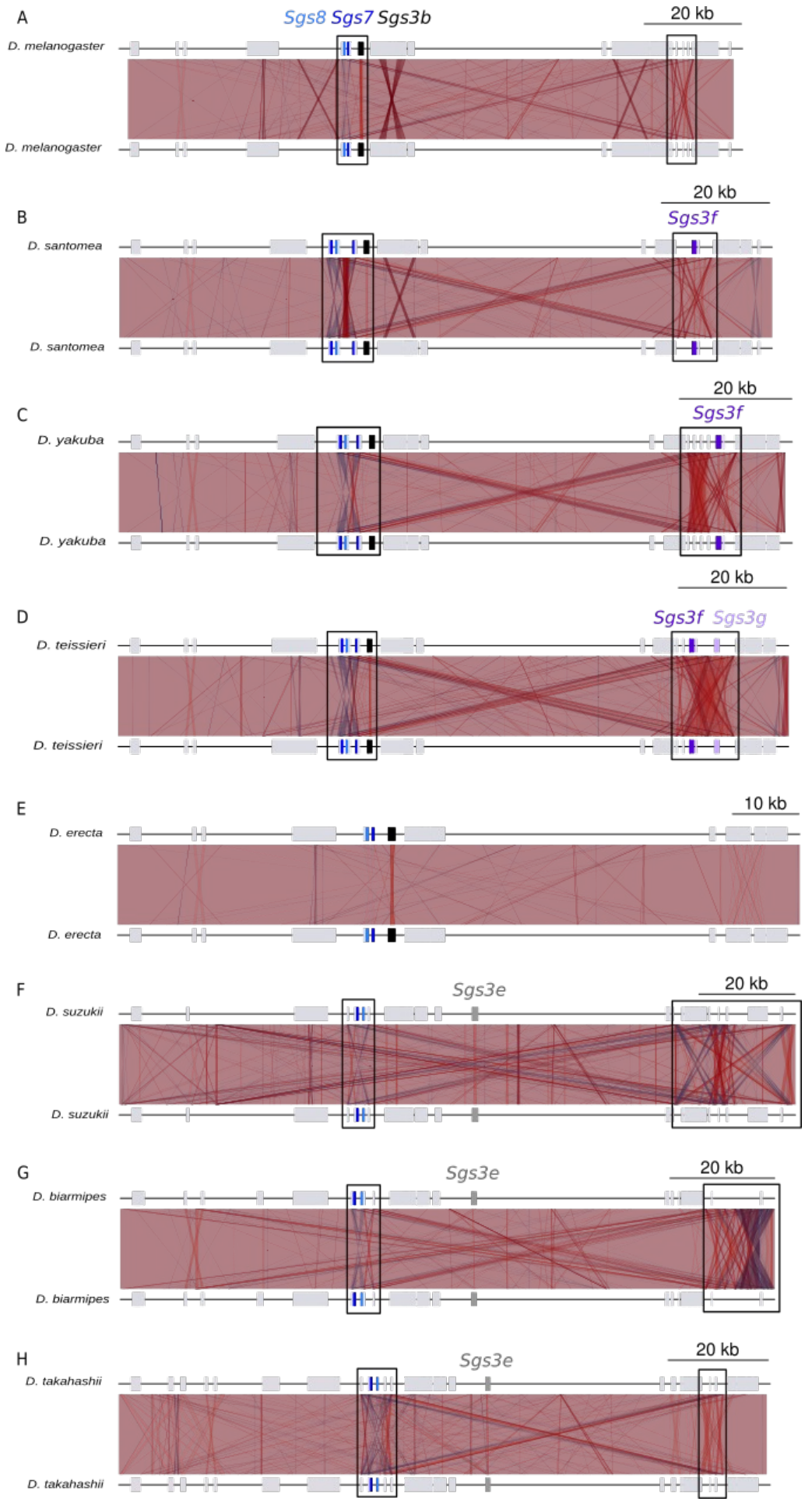

A
Sgs8 Sgs7 Sgs3b
20 kb
D. melanogaster
D. melanogaster
B
20 kb
Sgs3f
D. santomea
D. santomea
C
20 kb
Sgs3f
D. yakuba
D. yakuba
D
20 kb
Sgs3f Sgs3g
D. teissieri
D. teissieri
E
10 kb
D. erecta
D. erecta
F
Sgs3e
20 kb
D. suzukii
D. suzukii
G
Sgs3e
20 kb
D. biarmipes
D. biarmipes
H
Sgs3e
20 kb
D. takahashii
D. takahashii

Fig. S10. Comparison of the *Sgs3-7-8* genomic region from several Drosophila species with itself. Same legend as in Fig. S2. *Mob2* gene has been removed from this representation for the sake of readability because it is superposed with one of *Sgs3* copies. Black frames highlight two different genomic regions matching between each other, as we can see by the dark red lines forming a 'cross' pattern between the two loci. Comparison for *D. melanogaster* (A), *D. santomea* (B), *D. yakuba* (C), *D. teissieri* (D), *D. erecta* (E), *D. suzukii* (F), *D. biarmipes* (G) and *D. takahashii* (H).

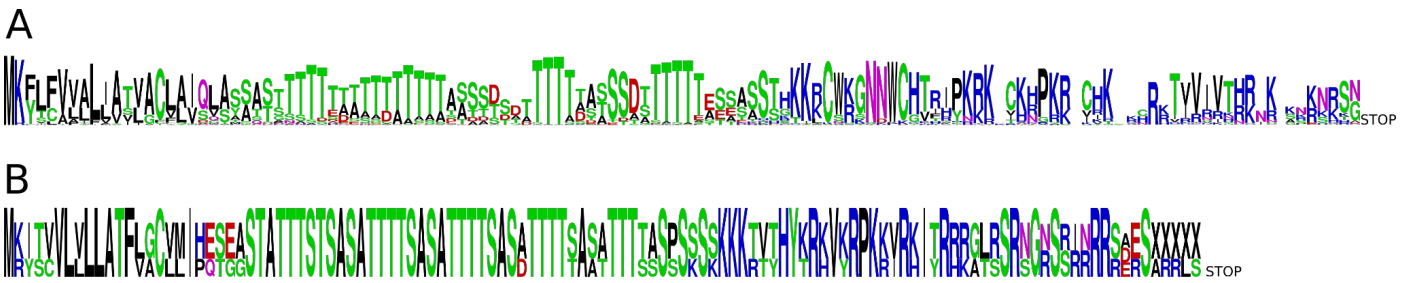


Fig. S11. Conserved amino acid sequence of ng proteins. Same legend as in Fig.1A. Numbers indicate the positions of the amino acid in *D. melanogaster CG33500* protein. (A) Conserved amino acid sequence based on an alignment of 154 ng proteins from all studied species. (B) Conserved amino acid sequence for three previously annotated *ng* genes at position 3C1 in *D. melanogaster (ng1*, *ng2*, *ng3*). We did not include previously annotated *ng4* gene because it is missing threonine-rich repeats.

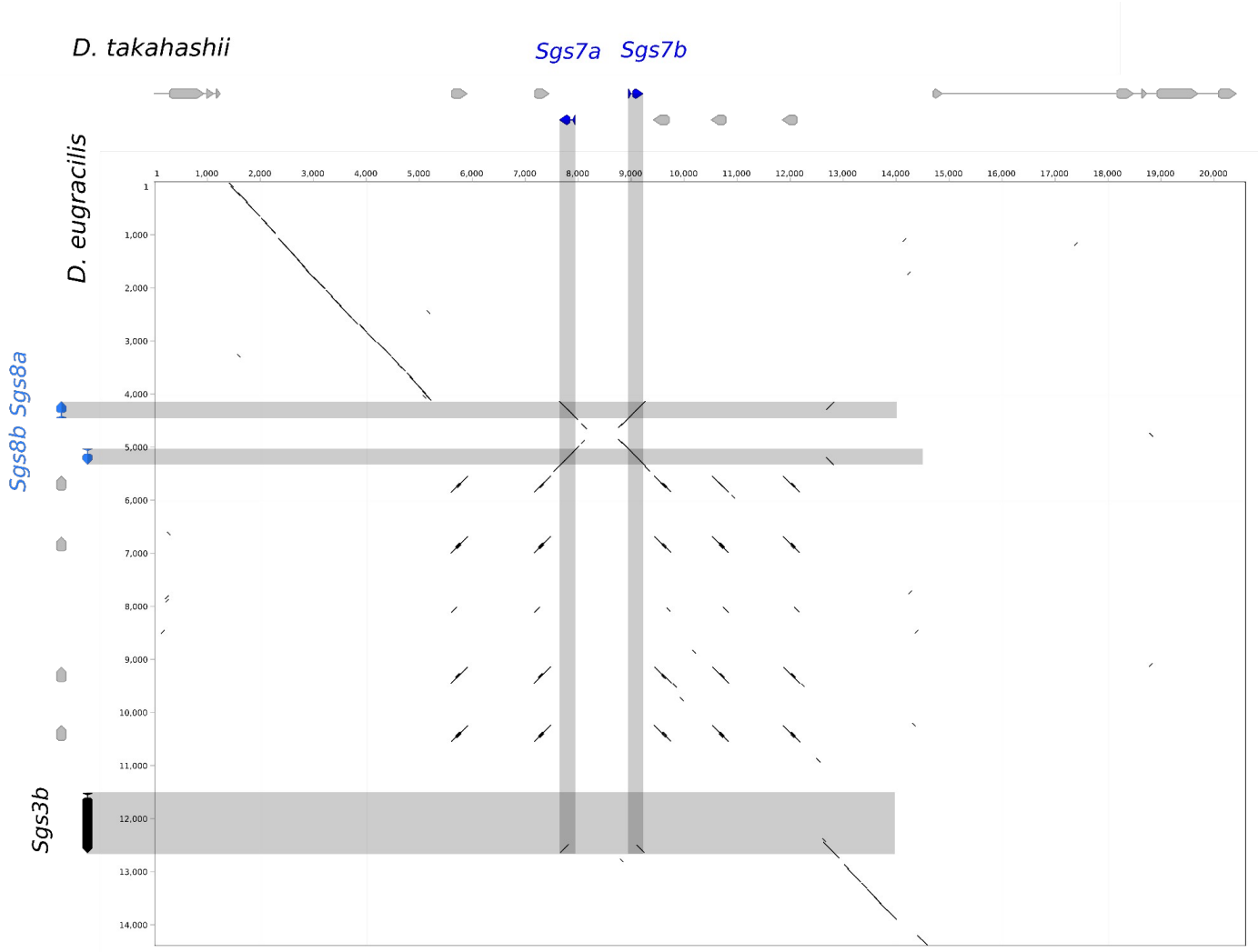


Fig S12. Dotplot comparison of *Sgs3-7-8* genomic regions from *D. eugracilis* and *D. takahashii*. Same legend as Fig. S1.

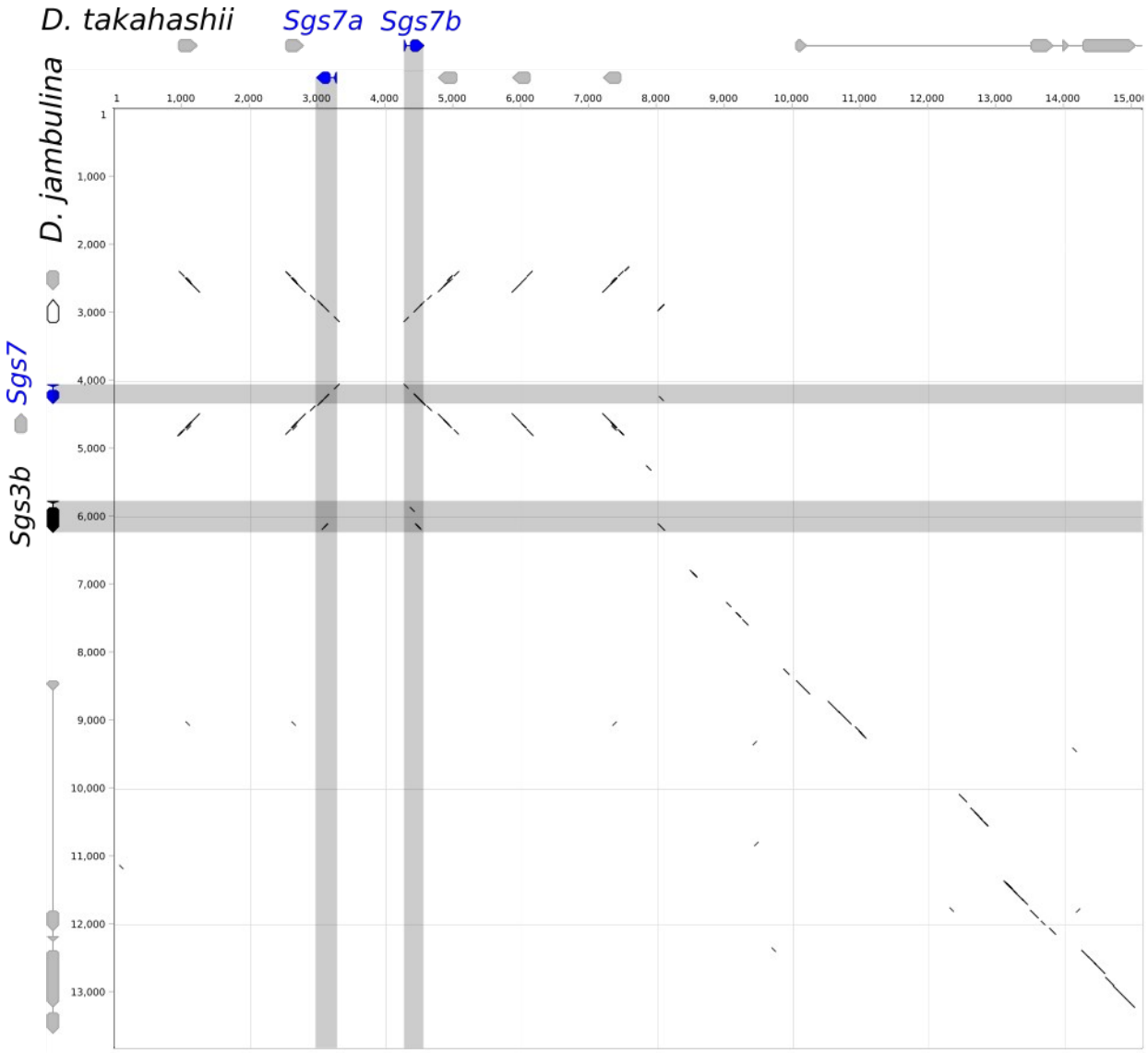


Fig S13. Dot plot comparison of *Sgs3-7-8* genomic regions from *D. takahashii* and *D. jambulina.* Same legend as Fig. S5. The white arrow indicates an *Sgs* pseudogene in *D. jambulina* which is missing the first coding exon and the start codon.

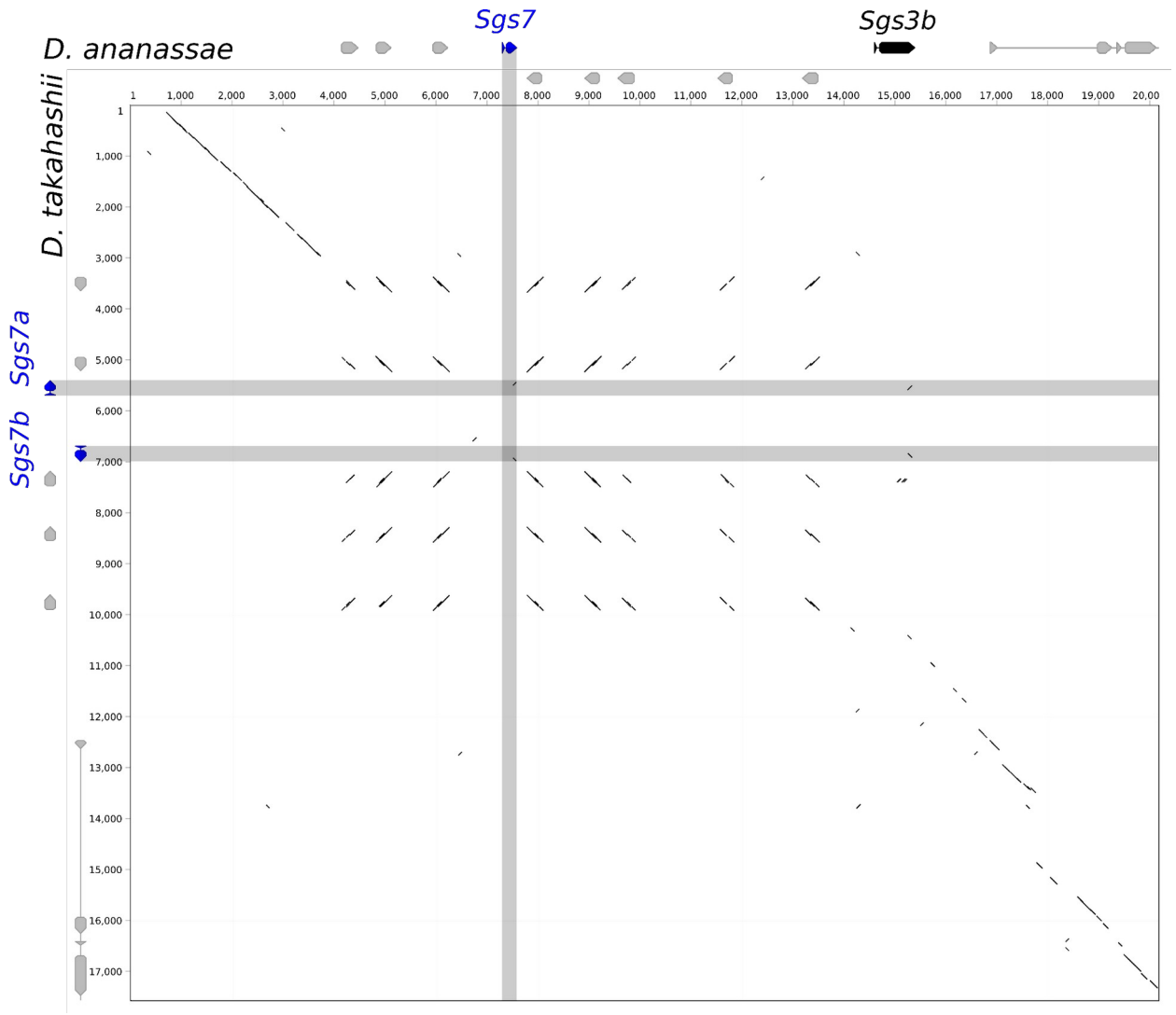


Fig S14. Dot plot comparison of *Sgs3-7-8* genomic regions from *D. ananassae* and *D. takahashii*. Same legend as Fig. S1.

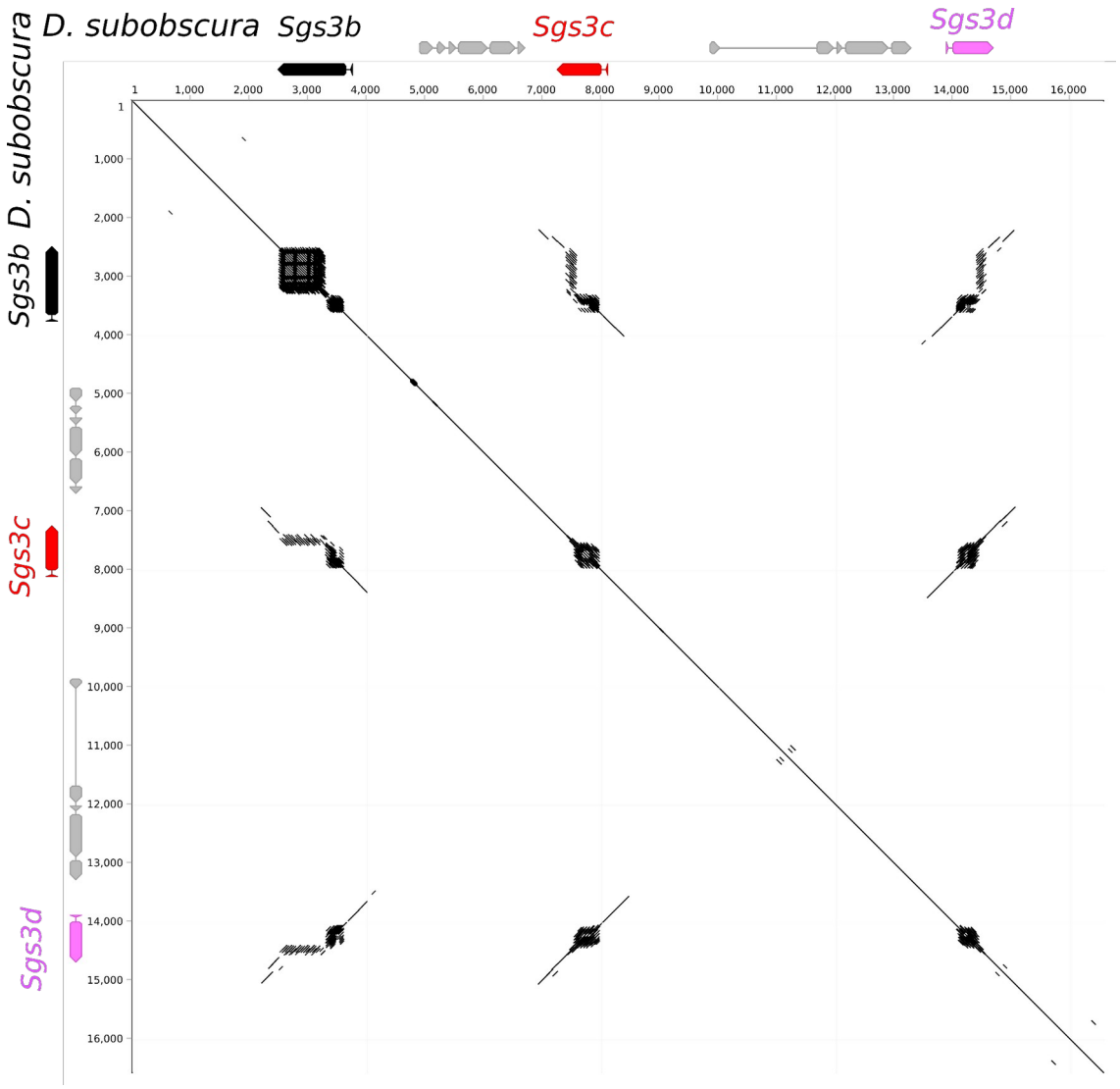


Fig S15. Dot plot comparison of *Sgs3b-d* genomic region from *D. subobscura* with itself. Same legend as Fig. S5. Black, red and light pink arrows represent *Sgs3b*, *Sgs3c* and *Sgs3d*, respectively. There are no ng genes in this region.

Supplementary Tables

Table S1. List of species and genome assemblies used in this study. All genome assemblies are PacBio-based or Nanopore-based, except the *D. eugracilis* and *D. takahashii* genome assemblies which relied on Illumina GAIIx data only. P: genome assembly based on PacBio and Illumina reads, N: genome assembly based on Nanopore and Illumina reads, I: genome assembly based on Illumina reads only.

Species	Strain	Genome Assembly	Gene annotations
D. melanogaster	iso-1	Release 6.32[P] The FlyBase Consortium/Berkeley Drosophila Genome Project/Celera Genomics https://www.ncbi.nlm.nih.gov/assembly/GCF_000001215.4/ GCA_000001215.4 P	FlyBase Release 6.32
D. simulans	w501	Princeton University https://www.ncbi.nlm.nih.gov/assembly/GCF_016746395.1 GCA_016746395.1 P	FlyBase Release 2.01
D. sechellia	sech25	University of California, Irvine https://www.ncbi.nlm.nih.gov/assembly/GCF_004382195.1 GCA_004382195.1 P	NCBI Release 101
D. mauritiania	mau12	University of California, Irvine https://www.ncbi.nlm.nih.gov/assembly/GCF_004382145.1/ GCA_004382145.1 P	NCBI Release 100
D. santomea	CAGO	Princeton University https://www.ncbi.nlm.nih.gov/assembly/GCF_016746245.2 GCA_016746245.2 P	This study. *Sgs* gene annotations from Da Lage et al., 2019.
D. teissieri	GT53w	Princeton University https://www.ncbi.nlm.nih.gov/assembly/GCF_016746235.2 GCA_016746235.2 P	This study.
D. yakuba	NY73PB	Princeton University https://www.ncbi.nlm.nih.gov/assembly/GCA_016746335.2 GCA_016746335.2 P	FlyBase Release 1.04
D. erecta	14021-0224.00,06,07	University of Arizona/ University of Chicago/ Cornell University https://www.ncbi.nlm.nih.gov/assembly/GCF_003286155.1 GCA_003286155.2 P	NCBI Release 101
D. eugracilis	14026-	The modENCODE Project	NCBI Release 101

	0451.10	https://www.ncbi.nlm.nih.gov/assembly/GCF_000236325.1 GCA_000236325.2 I	
D. takahashii	14022-0311.13	Baylor College of Medicine https://www.ncbi.nlm.nih.gov/assembly/GCF_000224235.1 GCA_000224235.2 I	NCBI Release 101
D. suzukii	WT3-2	Institut de Biologie du Developpement de Marseille https://www.ncbi.nlm.nih.gov/assembly/GCF_013340165.1 GCA_013340165.1 P	This study. *Sgs* gene annotations from Da Lage et al., 2019.
D. biarmipes	DSSC 14023-0361.11	University of Pennsylvania https://www.ncbi.nlm.nih.gov/assembly/GCA_005234255.1#/st GCA_005234255.1 P	NCBI Release 101
D. ananassae	14024-0371.16-18	University of Arizona/ University of Chiago/ Cornell University https://www.ncbi.nlm.nih.gov/assembly/GCF_003285975.2 GCA_003285975.3 P	NCBI Release 101
D. pseudoobscura	MV2-25 14011-0121.94	University of California, Irvine https://www.ncbi.nlm.nih.gov/assembly/GCF_009870125.1 GCA_009870125.2 P	NCBI Release 104
D. obscura	BZ-5 IFL	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018151105.1 GCA_018151105.1 N	This study
D. subobscura	Ksnacht	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCA_018903505.1 GCA_018903505.1 N	This study
D. rhopaloa	14029-0021.01	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018152115.1 GCA_018152115.1 N	This study
D. elegans	14027-0461.03	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018152505.1 GCA_018152505.1 N	This study
D. jambulina	14028-0671.01	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCA_018152175.1 GCA_018152175.1 N	This study

D. rufa	EH091 iso-CL_3	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCA_018153105.1 GCA_018153105.1 N	This study
D. kikkawai	14028-0561.14	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018152535.1 GCA_018152535.1 N	This study
D. triauraria	14028-0691.9	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCA_018151095.1 GCA_018151095.1 N	This study
D. bipectinata	14024-0381.04	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018153845.1 GCA_018153845.1 N	This study
D. ficusphila	14025-0441.05	Stanford University https://www.ncbi.nlm.nih.gov/assembly/GCF_018152265.1 GCA_018152265.1 N	This study

Table S2: (Table_S2.csv): Genomic coordinates of all the *Sgs* genes studied here in 24 Drosophila species.

Table S3: (.csv files compressed in a zip filev): Correspondence between NCBI gene names and the gene names used in this study, together with a description of the changes in gene annotations that have been made. ‘no change’ indicates that no modification was done on the annotations obtained from NCBI, ‘based on Borne et al, 2021 annotation’ means that the annotation was obtained from Borne et al. 2021 study. ‘annotations transferred from’ means that the gene annotation was done manually based on the existing annotation of the corresponding gene in a closely related species. There are four .csv files: (1) *Sgs1* and neighboring genes, (2) Sgs3x and neighboring genes, (3) *Sgs3/7/8* and neighboring genes, (4) *ng* genes annotated in 3C11-12, 87A1, 88C3-4 loci. In the third .csv file, ‘Newly annotated *ng* genes’ column indicates whether an *ng* gene was newly annotated in this study (‘Y’), already annotated (‘N’), or is not an *ng* (‘Not applicable’).

Table S4: (Table_S4.csv): *Sgs* exons and intron sizes for studied species. For each species, the size of the first coding exon (CDS1), intron and second coding exon (CDS2) are given in base pairs (bp). The amino acid encoded at the position of the unique phase 1 intron is also indicated.

Supplementary Files

File S1. Compressed zip file of the gene annotations (GenBank .gb files, inputs for Easyfig) of large genomic regions containing all the *Sgs* genes and their neighboring genes in the 24 studied species.

File S2. Fasta file of all the Sgs amino acid sequences used to create Figure 1B and Figure S1.

File S3. Compressed zip file of reference and corrected nucleotide sequences used to create Figure S2.

File S4. Compressed zip file of Sgs protein alignments (fasta.files) used to compute phylogenetic trees and make Weblogo figures.

File S5. *Sgs* coding sequence length in bp for species having an Sgs3x copy (.csv file, input for R script sgs_size.R).

File S6. *Sgs* coding sequence length in bp for species not having an Sgs3x copy (.csv file, input for R script sgs_size.R).

File S7. Compressed zip file of comparisons between pairs of large genomic regions (.out files obtained as outputs from Easyfig).

File S8. Table of pairwise percentage of identity between several Sgs1 and Sgs3 amino-acid sequences (.csv).

File S9. Compressed zip file of the repeats annotations (.csv files) obtained with FindRepeat in Geneious on large genomic regions for *D. melanogaster Sgs1, Sgs3/7/8, Sgs3x, D. teissieri Sgs3/7/8, D. subobscura Sgs3, D. eugracilis Sgs3.*

File S10. Compressed zip file of new glue protein alignments (.fasta files) used to make Fig. S9.

File S11. Fasta file of all the *Sgs* nucleotide sequences studied here.

File S12. Fasta file of the 154 *ng* nucleotide sequences found at loci 68C11 and 68C13.

File S13. Fasta file of the 41 *ng* nucleotide sequences found at loci 3C11-12, 28E6-28E7, 87A1 and 88C3-4.

File S14. Compressed zip file of all the R scripts (.R files) used to create the figures.

File S15. Bam file of raw reads mapped to D. rhopaloa *Sgs1* corrected nucleotide sequence, used to create Figure S2A.

File S16. Bam file of raw reads mapped to *D. ficusphila Sgs1* reference nucleotide sequence, used to create Figure S2B.

File S17. Bam file of raw reads mapped to *D. biarmipes Sgs3x* corrected nucleotide sequence, used to create Figure S2C.

2.3 Discussion

I annotated a total of 102 *Sgs* genes and identified several *new glue* gene copies not yet annotated. This study shows how highly accurate genome assemblies, using sequences from closely related species and synteny analysis, are of high importance to find orthologous genes. Indeed, the recently published genomes are more accurate and revealed fewer *Sgs7* copies in *D. suzukii* than the one used in the previous study (Da Lage et al., 2019). By using *Sgs* sequences from closely related species to search for orthologs, we were able to find more copies than in the previous study where only *D. melanogaster* sequences were used. Finally, studying *Sgs* genes synteny enabled us to find a new locus with *Sgs3* copies located 60 kb away from *Sgs3/7/8* cluster. Synteny analysis also revealed the presence of small genes named *new glue* (*ng*), neighbouring *Sgs* genes duplicates. These genes are short (243 to 426 bp), have a peptide signal and code for proteins rich in threonine. Despite their name, the precise function of *new glue* genes is unknown. Further experiments will decipher whether their activity is linked to pupa adhesion.

We observed two evolutionary dynamics. A first group of glue genes (*Sgs1*, *Sgs3X* and *Sgs3e*) has undergone a few deletions but no duplication, inversion or gene conversion. Genes from the second group (*Sgs3b*, *Sgs7* and *Sgs8*) duplicated several times and underwent multiple inversion and gene conversion events. We found that the second group of genes are located next to *new glue* genes. The presence of many *new glue* genes near *Sgs* genes may constitute repeated elements, that could promote a high rate of evolutionary events in *Sgs* genes.

Whether all the glue genes annotated in other species than *D. melanogaster* are expressed or not is still unknown. We can suppose that, if expressed, the *Sgs* duplicates found in this study play a role in glue properties and have evolved distinct protein functions. Our next objective will be to compare adhesion properties between these species, as previous work mostly analysed *D. melanogaster* adhesion (Borne et al., 2020, 2021a,b).

Chapter 3

Interspecific comparison of adhesive properties in Drosophila genus

3.1 Abstract

This section aims at characterizing adhesive properties of several Drosophila species and identifying species of interest to develop a future bioadhesive.

Previously in our lab, adhesion test experiments were developed on *D. melanogaster* using a dedicated machine. The force necessary to detach a pupa from a substrate was measured and a total of four species were tested. In this chapter, we tested 27 Drosophila species and assessed adhesive properties not included in previous analysis. We also measured pupa size and the surface of glue at the interface with the substrate to test their possible influence on adhesive properties.

We chose species distantly related from each other, they are spanning across 150 million years of evolution and have adapted to a wide diversity of pupation sites. Some species are found in various environments while others are specialized in a particular one. We hypothesized that this ecological diversity has led to different adhesive properties. By comparing pupa adhesion between different species, we aim at understanding better the role of adhesion in pupal adaption to its environment.

3.2 Introduction

Drosophila glue maintains the pupa attached to a substrate for several days as the animal undergoes metamorphosis (Monier and Courtier-Orgogozo, 2022). This adhesive is adapted to a great variety of surfaces in the wild and in laboratory conditions. As Drosophila species are widely used in the lab for research in genetics, they constitute an interesting model to study adhesion.

Pupa adhesion has been previously studied in our lab. Adhesion tests were performed with a force transducer, a sensor that converts a mechanical load input into an electrical output signal. Force-versus-distance curves, called force curves, were obtained. An automated pull-off adhesion test program was developed to measure the adhesion force, that is the force required to detach a pupa from a substrate (Borne et al., 2020).

Drosophila melanogaster adhesion force ranged between 151 mN and 269 mN for 1.1 mm^2, giving an adhesion stress of 137–244 kPa, which is similar to the one of strong commercial adhesive tapes (Borne et al., 2020; Du et al., 2021). Several substrates were tested (non-coated, Poly-L-lysine-coated (PLL-coated), poly-L-lysine–polyethylene glycol-coated (PLL–PEG- coated) and oxygen-activated glass slides) and displayed similar adhesion forces (Borne et al., 2020). As we are not yet able to extract the glue and conduct directly experiments on it, we are measuring adhesion on the pupa and its glue. As pupa size and shape or the amount of glue produced might differ between individuals and affect the measure, pictures of pupae before adhesion assay experiments are taken to analyse their morphology and the surface of glue in contact with the substrate.

Twelve different *D. melanogaster* strains, from geographically diverse regions, were previously tested for their adhesion force in the lab (Borne et al., 2021a). The variation of their adhesion force was not correlated to their glue surface of contact between the pupa and the substrate. These results suggest that the differences between genotypes are not due to the quantity of glue produced but to the glue properties.

Outside of *D. melanogaster*, adhesion force was tested for three other species, *D. simulans*, *D. suzukii* and *D. hydei* (Borne et al., 2021b). *D. simulans* adhesion force was similar to *D. melanogaster* (median adhesion force of 234 mN) but was lower in *D. suzukii* (median adhesion force of 78 mN) and was higher in *D. hydei* (median adhesion force of 482 mN). By dividing the adhesion force by the surface of glue between the pupa and its substrate, *D. hydei* displayed a higher adhesion stress than other species.

Here, our objective is to examine pupa adhesion in several fly species available in the lab (Borne et al., 2020). We used the experimental set up developed in the lab to conduct our experiments. We aimed to broaden our analysis by comparing other adhesion properties than adhesion force, in particular elasticity and hardness, given by the force-versus-distance curves (Butt et al., 2005).

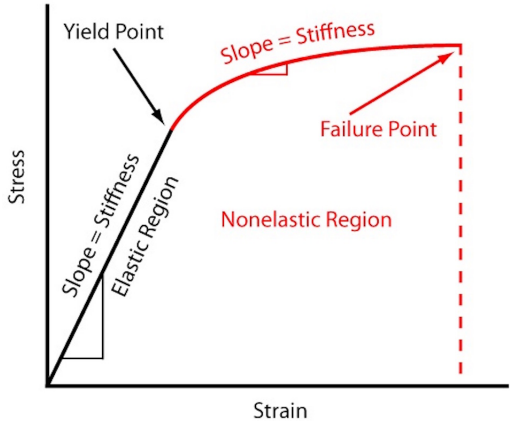


Figure 3.1: **Relationship between strain and stress for a given material.** Elastic and plastic regions of the material are respectively represented in black and red. Source: www.fictiv.com

When a stress is applied, a material is said to be elastic when it can regain its original shape after deformation (Figure 3.1). When the stress increases, the material will reach its elastic limit, the yield point, from which it will be permanently deformed, in which case, it is said to be plastic. Stiffness, also known as rigidity, is the resistance of an object to deformation of an applied force. For example, rubber has a low stiffness, it can undergo a lot of deformation and regain its original shape, while diamond has a high stiffness, its elastic region is very limited. The failure point is when the material has reached its plasticity limit and physically separates. Hardness is the measure of resistance of a material to a localised plastic deformation.

Our study aims to examine the glue properties for several Diptera species and determine the ones representing an interest to develop future adhesives.

3.3 Material and methods

3.3.1 Fly culture and stocks

Flies were cultured in plastic vials on standard medium (4 l: 83.5 g yeast, 335.0 g cornmeal, 40.0 g agar, 233.5 g saccharose, 67.0 ml Moldex, 6.0 ml propionic acid). For *D. suzukii*, *D. prostipennis*, *D. kurseongensis*, *D. rhopaloa*, *D. elegans* this medium was supplemented with 200 g of D-glucose anhydrous (VWR Chemicals, reference: 24379.294). Stocks, their origin and their temperature of culture are given in Table 3.1.

3.3.2 Adhesion assays

Third instar wandering larvae were washed in PBS to remove traces of food and microorganisms from their surface. They were then dried by putting them briefly on tissue

paper and transferred on glass slides (Menzel Superfrost microscope glass slide, Thermo-Scientific™ n° AGAB000080) with soft forceps within a closed box that was maintained on humid wet cotton. Between 15 and 21 hours after transfer, pupae naturally attached to the glass slides with their ventral part adhering to the glass slide were processed as follows. The pull-off force necessary to detach the pupa from the glass slide was measured using a universal test machine (LS1S/H/230 V Lloyd Instruments) with a 5 N force sensor (YLC-0005-A1 Lloyd Instruments), in a set-up similar to the one published earlier.

We used an automated program to detach pupae from glass slides, which controls the captor. We define a protocol of reference, called 'standard protocol'. For this protocol, pupae were detached between 15 and 21 hours after transfer, they were naturally attached to the glass slides with their own glue. Double-sided adhesive tape (Tesa, extra strong, n° 05681-00018) was attached to a cylindrical metal part in contact with the force sensor. The force was set to 0 before each run. The force sensor was moved down with a constant speed of 1 $mm.min^{-1}$ until it pressed the pupa with a force of 0.07 N. It was stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.2 $mm.s^{-1}$ until the pupa was detached. The force, time and position of the sensor were recorded using NEXYGENPlus software (Lloyd Instruments).

3.3.3 Alternative protocols for adhesion assays

To validate the standard protocol (1 tape ; glue) we used alternative protocols described below. In each of them, at least one parameter was modified compared to the standard protocol. These protocols are listed hereafter and in Table 3.2.

- 1 tape ; no glue: Before adhesion assays, pupae naturally attached on glass slides were manually detached with a paintbrush. Pupae damaged during this manual detachment were not kept. Detached pupae were then placed on a clean glass slide with their ventral side in contact with the slide. The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 1 $mm.min^{-1}$ until it pressed the pupa with a force of 0.07 N. It was stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.2 $mm.s^{-1}$ until the pupa was detached.

  This protocol enables to study pupa properties independently of the glue. Moreover, by comparing the properties of detached and attached pupae, we can infer the glue properties.

- 2 tapes ; no glue: Before adhesion assays, pupae naturally attached on glass slides were manually detached with a paintbrush. Pupae damaged during this manual detachment were not kept. Detached pupae were placed on a piece of Tesa double sided tape on a clean glass slide with their ventral side in contact with the slide. The

force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

- 1 strong tape ; glue: Compared to the standard protocol, we used Fermoflex n° 8616 9mm double sided tape on the sensor. The force sensor was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

  By changing the type of tape used, we assess the effect of the tape on our measurements. Note that this Fermoflex tape being stronger than the Tesa tape, we can detach pupae that do not detach with Tesa tape.

- no tape ; glue: No tape was placed on the force sensor. The force sensor was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.2 mm.s^{-1} until reaching the initial position.

  With this protocol, we can assess the influence of double sided tape on our measurements.

- 0.25 N: The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.25 N then stilled at a force of 0.21 N for 10 s, and finally moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

  In this protocol, the maximal force applied to the pupa is higher compared to the standard protocol. We want to determine up to which maximal force the compression curve remains linear.

- 3 d: Adhesion tests were performed 3 days after wandering larvae were deposited on glass slides.

  By using older pupae, we want to determine the evolution of adhesion properties along metamorphosis.

- 5 min: The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 5 min and finally moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

- 0 s: The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.07 N

and immediately moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

Both 5 min and 0 s protocols are used to determine the influence of experiment duration on the irreversibility of the pupa and glue deformation.

- Speed /3: The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 0.33 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.06 mm.s^{-1} until the pupa was detached.

- Speed x3: The force sensor, covered with Tesa double sided tape, was moved down with a constant speed of 3 mm.min^{-1} until it pressed the pupa with a force of 0.07 N then stilled at a force of 0.03 N for 10 s and finally moved up with a constant speed of 0.6 mm.s^{-1} until the pupa was detached.

  Speed variation deciphers the role of plasticity and viscosity in the irreversibility of the pupa and glue deformation. When speed affects the irreversibility, the system is viscous, if not, it is plastic.

- 1 strong tape ; glue ; 0.25 N: The force sensor, covered with Fermoflex double sided tape, was moved down with a constant speed of 1 mm.min^{-1} until it pressed the pupa with a force of 0.25 N then stilled at a force of 0.21 N for 10 s and finally moved up with a constant speed of 0.2 mm.s^{-1} until the pupa was detached.

  This protocol was used to detach large pupae that could not be detached with a standard protocol or with the "1 strong tape ; glue" protocol.

3.3.4 Adhesion tests data analysis

The adhesion test data were analysed with scripts developed on R version 4.1.2 (2021-11-01) https://www.r-project.org. Pipeline of the script is in the README (GitHub link : https://github.com/manonmonier2/Adhesion/tree/manon_branch).

Figures were prepared by scripts developed on R version 4.1.2 (2021-11-01) https://www.r-project.org and modified on Inkscape 1.2.1 (2022-07-14 version) https://inkscape.org/.

Adhesion curves obtained from the adhesion assays were segmented into six regions (Figures 3.2 and 3.9 below) defined by seven landmarks.

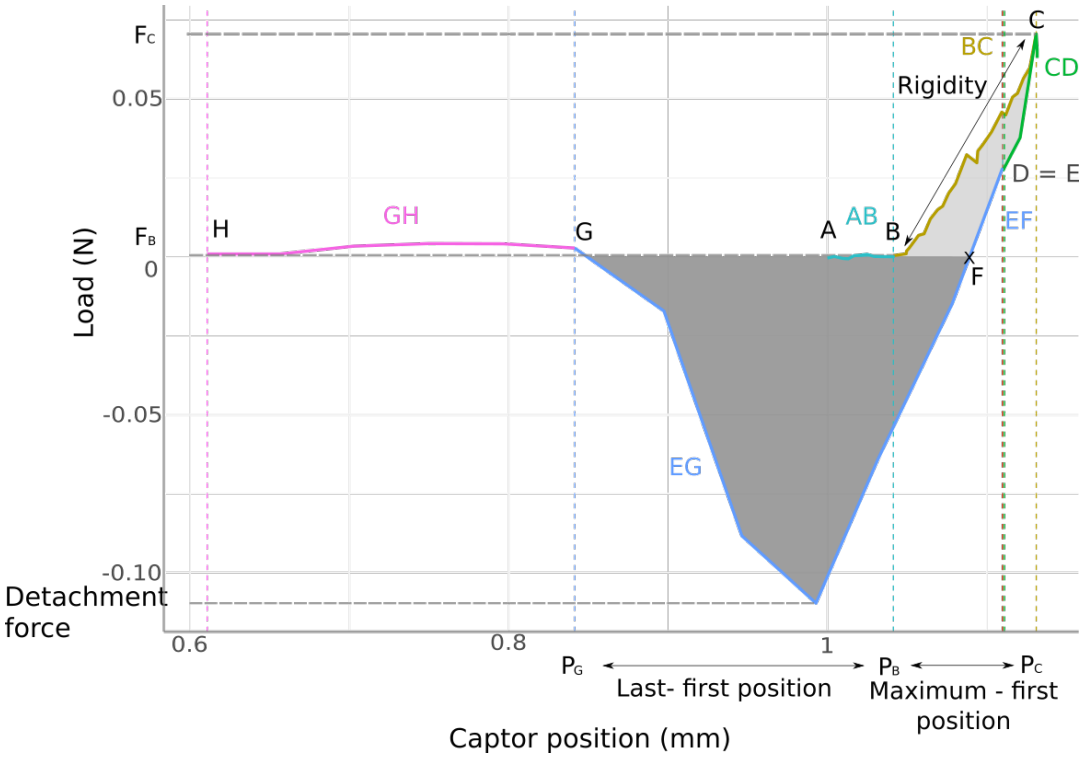


Figure 3.2: **One example of force-versus-distance curve obtained with our pupa adhesion assay.** The curve represents the detachment of a pupa during a standard protocol assay. It is separated in six regions distinguished by colors (AB, BC, CD, DE, EG, GH) and separated by colored dashed lines at their respective landmarks (A, B, C, D, E, G, H). Landmarks D and E being at the same position, their dashed lines are overlapping. The light blue region DE is not visible on this curve because the captor position is stable at 0.03 N for 10 s. Forces F_B, F_C, detachment force and positions P_B, P_C and P_G are indicated by grey dashed lines.

Landmark A corresponds to the beginning of the experiment. We calculated the standard deviation of the force over the 20 first values. When the force exceeds the standard deviation, landmark B is defined at the position P_B and force F_B. The region between this two landmarks, AB, corresponds to the descent of the captor. Landmark C is defined when the maximal positive force F_C is reached. In the second region of the curve (BC), the captor is pressing into the pupa until reaching a given force (0.07 N for standard protocol). Landmark D is reached when the force reaches for the first time a given value (0.21 N for "0.25 N" and "1 strong tape ; glue ; 0.25 N" protocols and 0.03 N for other protocols). In the third region (CD), the captor goes back to the pause force (0.03 N for standard protocol). Landmark E corresponds to the end of the pause. In the fourth region (DE), the captor is stable (at 0.03 N for 10 s for standard protocol). Starting from landmark E, the force decreases, is null at landmark F, reaching the maximal negative force representing the detachment force, and then increases again until reaching landmark G when it is null again. The region (EG) corresponds to the pupa detachment. Landmark H is the end of the experiment. Finally, the captor is going

back to a defined position in region GH.

The compression curve is comprised between landmarks A and C and decompression curve between landmarks C and G. The maximal force measured in this zone is the detachment force.

3.3.5 Definition and calculation of adhesion properties

By analysing the adhesion assay curves for each pupa, we were able to estimate six quantities relating to the pupa and/or its glue: deformation reversibility, the maximal difference of extension, the rigidity, the detachment force and one way and two ways detachment energies, as explained below:

A deformation is said to be reversible when the object keeps its shape after a force is applied. Then, the object is elastic. When the object shape has changed, the deformation is irreversible. The deformation reversibility of the animal and its glue is obtained by the difference of position between the first contact with the pupa (P_B) and the last one (P_G). It is calculated by the formula below and is expressed in mm.

$$Last\ -\ first\,position = P_B - P_G$$

The difference of position between the position where the maximal force is applied (P_C) and the contact with the pupa (P_B) is expressed by the formula below in mm.

$$Maximum\ -\ first\,position = P_B - P_C$$

The rigidity represents the resistance of an object to deformation of an applied force. It is obtained by dividing the force applied by the deformation of the pupa:

$$Rigidity = \frac{F_C - F_B}{P_C - P_B}$$

F_B is the force applied at the first contact with the pupa at a position P_B and F_C is the maximal force applied at a position P_C. Rigidity is expressed in N/mm.

The detachment force is the maximal force required to detach the pupa. It corresponds to the maximal force applied on the decompression curve. It is the unique measure that was estimated based on previous adhesion tests (Borne et al., 2020).

One way detachment energy represents the energy given by an external force to detach the pupa. This energy corresponds to the one exerted by a predator in the wild to detach the pupa. It is the area of the negative part of the decompression curve.

$$One\,way\,detachment\,energy = \int_F^G f(x)\,dx$$

Two ways detachment energy corresponds to the energy given by an external force to

detach a pupa after pressing into it. It is calculated by the difference between the area under the compression curve and the decompression curve.

$$Two\,ways\,detachment\,energy = \int_B^C f(x)\,dx - \int_C^G f(x)\,dx$$

Both two ways detachment energy and one way detachment energy are expressed in N.mm.

In Figures 3.7, 3.14, 3.13, 3.15 and 3.16 decimal log is used and statistics are calculated on the decimal log values.

3.3.6 Phylogenetic tree

Phylogenetic tree representing the evolution of adhesion properties through evolution were computed using Mesquite software (version 3.8 http://www.mesquiteproject.org/). Adhesion force median values for each species are given to the software, which infers the ancestral state for each branch, based on a parsimony reconstruction method. Adhesion force is illustrated by a color gradient affected to branches. Phylogenetic tree topology is adapted from (Suvorov et al., 2022).

3.3.7 Pupae pictures measurements

Before adhesion tests, three images of the pupa on glass substrate were taken with a Keyence microscope VHX-2000 Z20 ×20: dorsal view, ventral view and side view.

Pictures of the pupa dorsal view are analysed automatically. First, an ImageJ (version 1.53c (https://imagej.nih.gov/ij/) macro converts the pictures into a .tif format in order to be processed by Ilastik (version 1.3.2 https://www.ilastik.org/) software. Second, these converted pictures are used for machine learning pixel classification in Ilastik. We used the pixel classification function to allocate pixels in two object classes: pixels belonging to the pupa and pixels belonging to the rest of the picture (glass slide). The features used to distinguish these two types of objects are the colour, edge and texture. Ilastik uses the machine learning algorithm Random Forest classifier, that is trained from user annotations interactively. The experimenter manually draws lines on the picture with two different colours: one to label the pupa and another to label the rest of the picture. Based on these lines manually drawn, the Random Forest assigns a probability for each pixel to belong to the pupa or the rest of the picture. This step is a training for the software to automatically recognize pupae. It is done for 100 pictures (6% of the total dataset) randomly chosen. Ilastik automatically classifies pixels for the rest of the pictures dataset. Ilastik output pictures are segmented into black and white regions, white belonging to the pupa and black to the rest of the picture. Third, an ImageJ macro segments the white pixels from the others, based on these binary pictures. It calculate its area and the Feret

diameter that corresponds to pupa length. Feret diameter is the longest distance between any two points along the selection boundary. Pupa shape being oval, its length is the longest distance.

Pictures of the ventral view were used to measure glue secretion area. Pictures of pupae were anonymized for manual contour acquisition so that the digitizer did not know the genotype. Romane Petit, intern in the lab, measured the areas of the print corresponding to the pupa–substrate interfaces manually using ImageJ.

Pictures of the side view were used to measure manually the thickness of the animal. It is measured as the distance between the ventral side and the the dorsal side, perpendicular to the glass slide.

3.4 Results

3.4.1 Our adhesion measure is not sensitive to protocol parameters

To determine the effect of the adhesion assay parameters on our measurements, we conducted control adhesion assays on *D. melanogaster* (CantonS strain) with different protocols. For each of these protocols (Table 3.2), at least one parameter differs from our reference protocol (standard protocol). We found that the measured adhesion force (also named detachment force) was not affected by the type of double-sided tape used, the maximal force at which pupae were pressed (0.07 N or 0.25 N), the time during which the pupa was pressed (0 s ; 10 s ; 5 min), the speed at which the captor was moved during compression and decompresion (speed x3 or speed/3). This indicates that our measure of adhesion force is robust and does not change with these experimental parameters.

The force-versus-distance curves allowed us to infer the elasticity, plasticity and rigidity properties of the pupa, when attached or detached. We compared these values between standard and alternative protocols (Figures 3.10 and 3.11).

- Last - first position : The difference of position between the last contact with the pupa and the first one (Last - first position) represents the deformation reversibility of the pupa and its glue. When the pupa is detached from its glue and placed on glass slides (1 tape ; no glue), pupa shape remains the same. We can infer that pupa deformation is reversible and that the pupa is elastic. In comparison, the system pupa and glue (standard) is deformed after adhesion assay, the deformation is said irreversible. When the irreversible deformation is sensitive to speed variation, the system is said to be viscous, if not it is plastic. The deformation observed does not depend on the compression and decompression speed (speed x3 and speed /3). We can infer that the pupa with glue is plastic.

When the pupa is detached from its glue and placed on Tesa tape (2 tapes ; no glue), the deformation is irreversible. This irreversibility is then due to the Tesa tape.

When the experiment duration is of 5 minutes, we observe a slight difference with the standard protocol, but not significant enough to undertake further study.

- Rigidity : The rigidity of the pupa/glue system increases when the pupae is older, meaning the pupa and/or its glue is more rigid when ageing. When the maximal force applied increases, the rigidity increases, which is due to the non linearity of compression curves (Figure 3.12). The standard force of 0.07 N is compatible with a linear behaviour.

  Detached pupae are softer than pupae attached with their own glue. We can suppose that the pupa is damaged when it is manually detached from its glue and that the animal cuticle gets softer. We can also suppose that the glue makes the attached pupae more rigid by attaching it to the slide.

- Detachment force and energy : For the different protocols used in control experiments, the detachment force, one and two ways energies are positively correlated together. We can thus use the detachment force as a proxy for the two other variables.

  No difference in detachment force and both energies are observed between pupae detached 3 days and between 15 to 21 hours after larvae transfer on glass slides. This suggests that pupa adhesion force is constant during the first three days after pupa formation.

These control experiments show that our standard experiment is adequate to study glue adhesion.

3.4.2 Pupae from several species do not detach after adhesion assay

After adhesion assay, three scenari are observed for the pupa. The first scenario is when the pupa detaches from the glass slide, and remains on the piece of double sided tape on the captor. The second is when the pupa is broken in two parts, the body remains on the double sided tape on the captor and the head is still attached onto the glass slide with the glue. The third is when the pupa is not detached after adhesion assay: the full body remains on the glass slide and is not damaged.

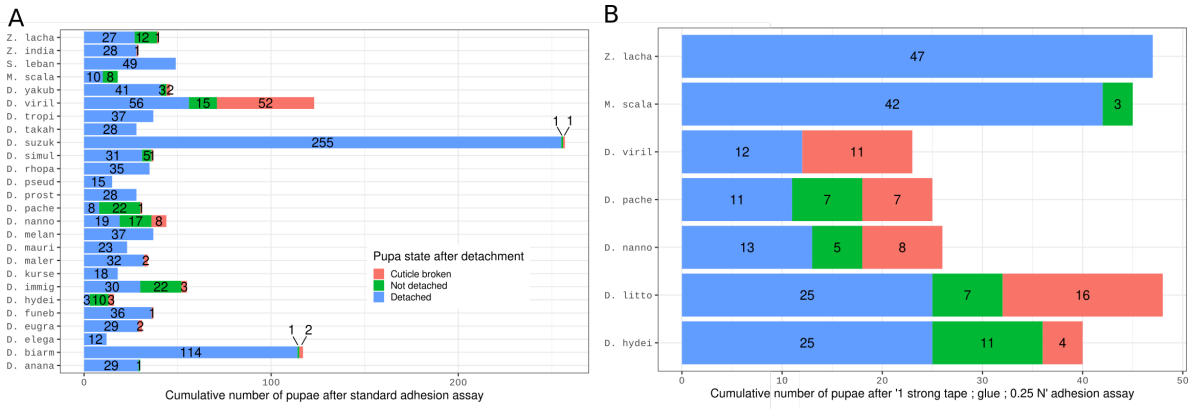


Figure 3.3: **Cumulative number of pupae after adhesion assays.** For each species, the number of pupae either broken, not detached or detached after adhesion assay are given for the standard protocol (A) and for the '1 strong tape ; glue ; 0.25 N' protocol (B).

Under standard protocol conditions where Tesa tape is used, *D. virilis*, *D. pachea*, *D. nannoptera*, *D. immigrans*, *D. hydei* display a high proportion of pupae broken or not detached compared to the proportion of detached pupae (Figure 3.3A). We tested whether using a stronger tape (Fermoflex) can help detaching more pupae from these species (Figure 3.3B). We noted that some of these species are larger, such as *D. virilis*, *D. littoralis*, *D. hydei* and we also applied a stronger force (0.25 N) as the standard force (0.07 N) might not enable a sufficient contact with the tape.

The protocol '1 strong tape ; glue ; 0.25 N' is efficient regarding *Z. lachaisei* and *M. scalaris* with most of the pupae detached from glass slides. For *D. virilis*, pupae are either detached or broken. For *D. pachea*, *D. nannoptera*, *D. littoralis* and *D. hydei*, broken and not detached pupae are still observed. For *M. scalaris*, the proportion of broken and not detached pupa did not exceed 7 %. In the following sections, only detached pupae will be used for our analysis.

3.4.3 Adhesion has no visible correlation with phylogeny

We conducted adhesion assays on pupae from 27 Drosophila species. We calculated adhesion properties for each pupa and grouped them by species (Figures 3.4 and 3.5).

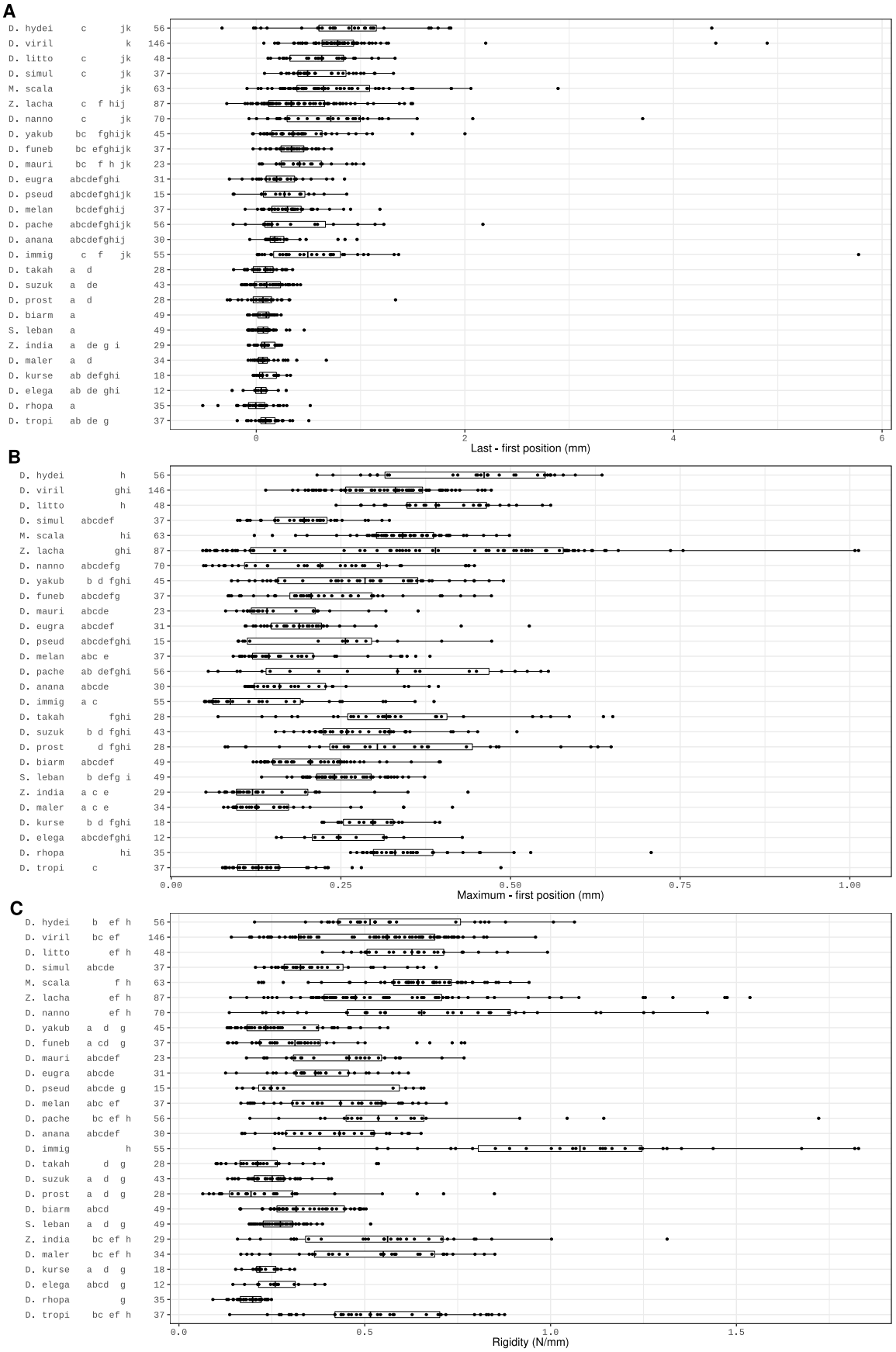


Figure 3.4: **Adhesion properties for Drosophila species.**

See legend below

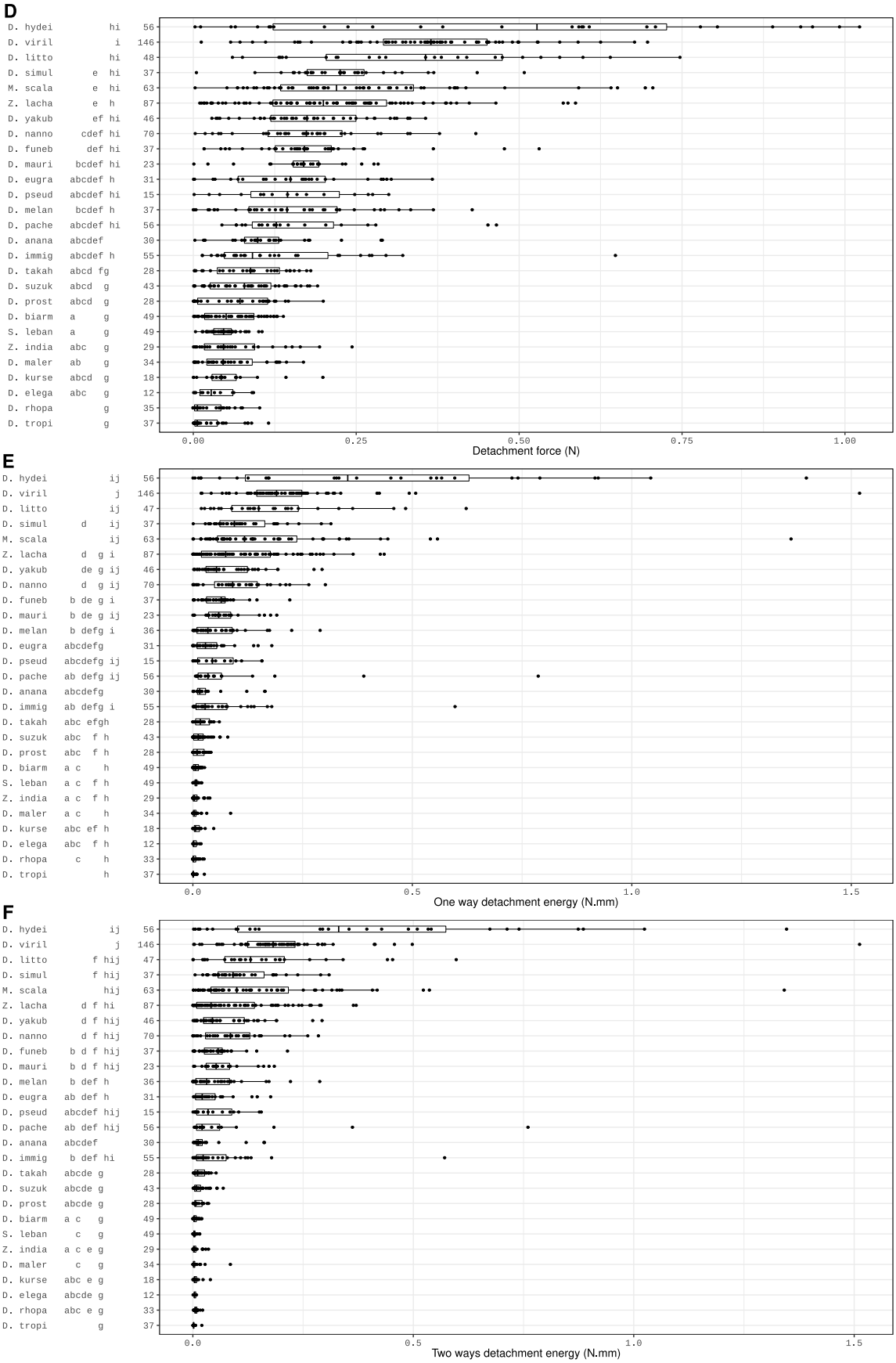


Figure 3.5: **Adhesion properties for Drosophila species.**

Last - first position (mm) (A), maximum - first position (mm) (B), rigidity (N/mm) (C) detachment force (N) (D), one way detachment energy (N.mm) (E) two ways detachment energy (N.mm) (F). Two ways detachment energy is the difference between the energy given during compression and the decompression energy. One way detachment energy corresponds to the area of the decompression curve negative part. First vs last position corresponds to the difference of position between the first contact with the pupa and the last one (detachment). First vs maximum position corresponds to the difference of position between the first contact with the pupa and the position where the maximal force is applied. Each dot corresponds to a single pupa. Ends of the boxes define the first and third quartiles. The black vertical line represents the median. The horizontal line on the right side of the box extends to the largest value no further than 1.5 × IQR from the right side of the box. Inter-quartile range (IQR) is the distance between the first and the third quartiles. The horizontal line on the left side of the box extends to the smallest value at most 1.5 × IQR of the box. Data beyond the end of these lines are 'outlying' points represented in grey. An ANOVA followed by all pairwise comparisons after Tukey correction ($P < 0.05$) was performed on the set of protocol groups. Groups that are not significantly different from each other share a letter. Figures on the right side of statistical letters indicates the total number of pupae tested for each condition.

Last - first position : The deformation of the pupa and its glue varies from 0 mm for *D. rhopaloa* up to 0.9 mm for *D. hydei*. It is correlated with the detachment force (Figures 3.4 and 3.13). Species with a high detachment force (*D. hydei* and *D. virilis*) are more deformed after the adhesion assay than species with a low detachment force (*D. tropicalis* and *D. rhopaloa*).

Rigidity : The rigidity of the pupa and its glue is very variable between species, from 0.19 N/mm in *D. prostipennis* up to 1 N/mm in *D. immigrans*. It is not correlated with other adhesion properties (data not shown).

Detachment force : The detachment force is positively correlated with one way and two way detachment energies (Figure 3.14). We will further use detachment force as a proxy for one way and two way detachment energies.

The median detachment force varies from close to 0 N for *D. tropicalis* up to 0.52 N for *D. hydei* while *D. melanogaster* has a median detachment force of 0.14 N. We can distinguish three groups of species according to their detachment force. A first group of low detachment force is composed of *D. tropicalis*, *D. rhopaloa*, *D. elegans*, *D. kurseongensis*, *D. malerkotliana*, *Z. indianus*, *S. lebanonensis*, *D. biarmipes*, *D. prostipennis*, *D. suzukii* and *D. takahashii* for which the force ranges from 0.005 N to 0.08 N. A second group of medium detachment force ranging between 0.09 N and 0.17 N comprises *D. immigrans*, *D. ananassae*, *D. pachea*, *D. melanogaster*, *D. pseudoobscura*, *D. eugracilis*, *D. mauritiana*, *D. funebris*, *D. nannoptera*, *D. yakuba*. Finally, a third group is composed of species with

a high detachment force comprised between 0.20 N and 0.52 N: *Z. lachaisei*, *M. scalaris*, *D. simulans*, *D. littoralis*, *D. virilis*, *D. hydei*.

Adhesion force has evolved rapidly between species in both directions. We did not identify any specific taxonomic groups sharing the same adhesion features. In the Figure 3.6), the adhesion force of two closely related species (*Z. indianus* and *Z. lachaisei* for example) diverged rapidly. The reconstruction of the ancestral state highlights the strong adhesion of *D. hydei* and *D. virilis* group. However, as adhesion seems to have evolved very quickly through evolution, we can not infer the common ancestral state of the whole tree with precision.

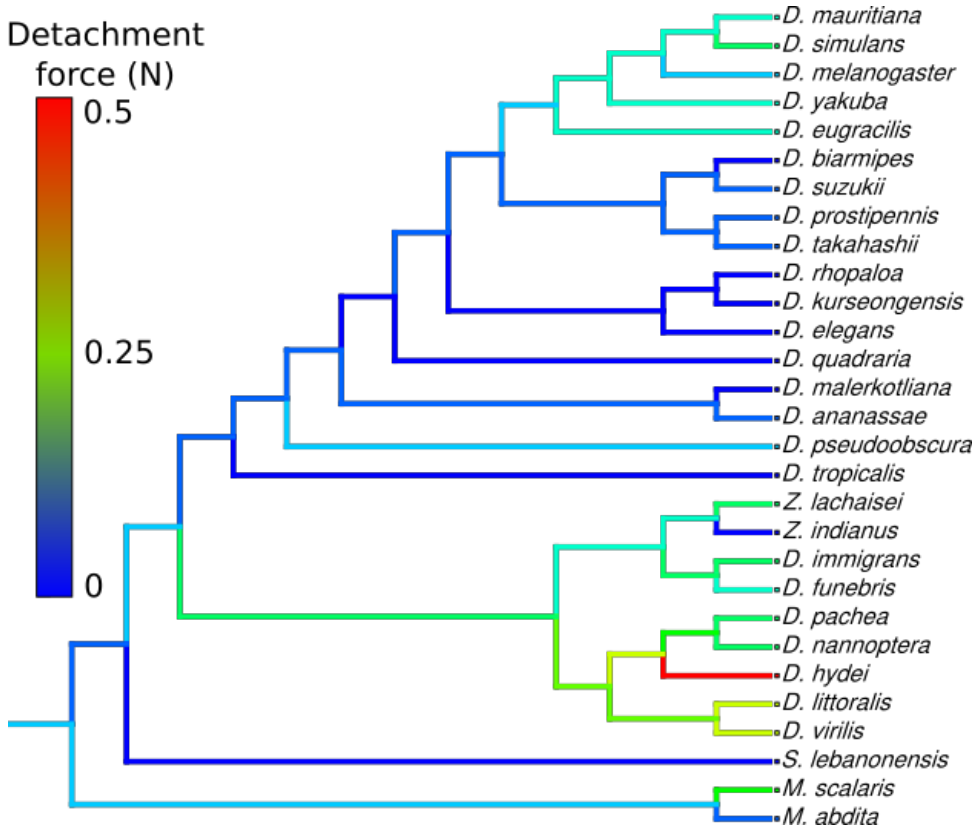


Figure 3.6: **Evolution of adhesion force through phylogeny.**
Species branch are colored according to their detachment force median value, from blue to red respectively corresponding to low and high detachment force. Ancestral branches are colored based on a parsimony algorithm. For example, *Z. indianus* is not adhesive, its branch is colored in dark blue and *Z. lachaisei* has a medium adhesion, its branch is colored in green. The most parsimonious scenario is that the common ancestral branch for both species has a low adhesion, its branch is colored in light blue. Phylogenetic tree is adapted from (Suvorov et al., 2022) and tree branch lengths do not represent real distances.

3.4.4 Interspecific adhesion properties variation is correlated with area of glue produced

Pupa size and length are positively correlated together (Figure 3.15), *D. malerkotliana* and *M. scalaris* being respectively the smallest and largest pupae among studied species. Interestingly, *D. virilis*, *D. hydei* and *D. littoralis* stand aside the correlation, they display a longer pupal case compared to their area. In my personal experience, *D. novamexicana* and *D. americana* display the same feature, although no measurements were done for these species. *D. virilis*, *D. littoralis*, *D. hydei*, *D. novamexicana* and *D. americana* are closely related, we can suppose that this elongated shape appeared in the common ancestor of these species.

Pupa area and pupa length are poorly correlated to the detachment force and other adhesion properties (Figure 3.16).

The glue surface of contact between the pupa and the glass slide varies between species, their median is ranging from 0 mm^2 for *D. biarmipes* up to 2 mm^2 for *D. virilis*. The detachment force and both one way and two detachment energies, increase with the glue surface of contact (Figure 3.7).

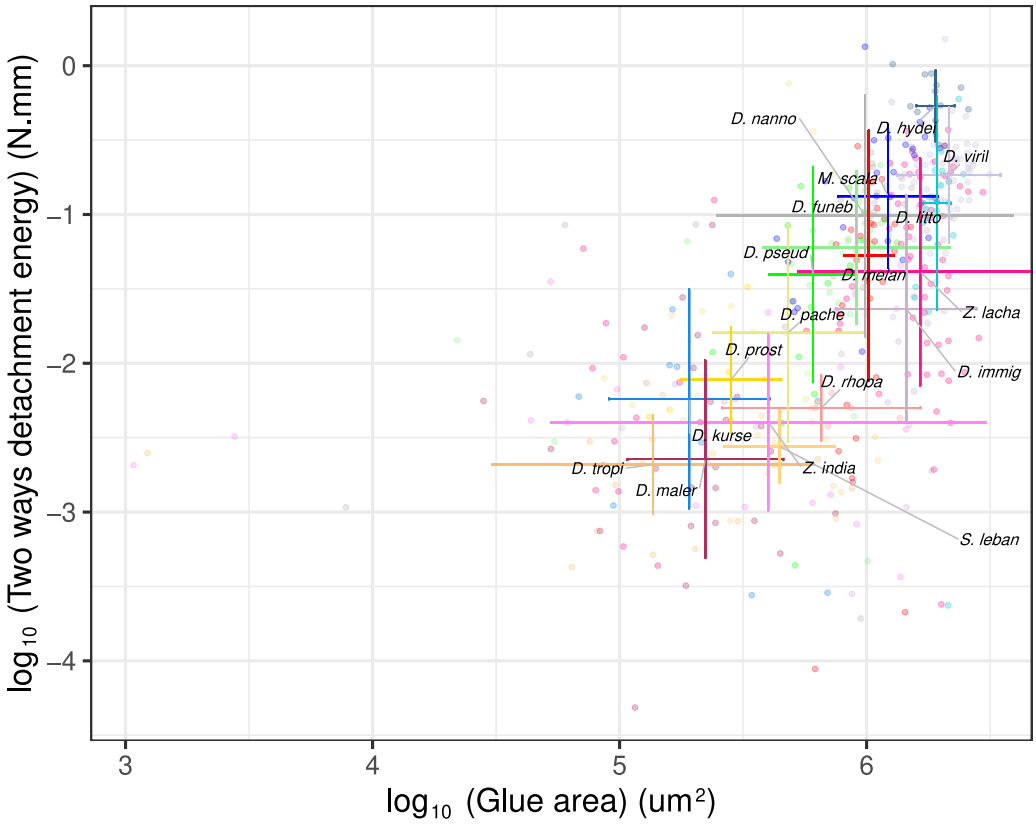


Figure 3.7: **Positive correlation between two ways detachment energy and the glue area in contact between the pupa and the glass slide.**
Each dot corresponds to a single pupa. Center of crosses correspond to the median of each protocol. Horizontal and vertical bars respectively represent standard deviation for x and

y axis.

The detachment force per unit area of glue, or detachment stress, is equivalent between species (when considering only pupae detached after adhesion assay) (Figure 3.8). *D. hydei* presents the highest detachment stress (0.37 N/mm^2) while it is null for *D. rhopaloa*. The same observation is made when considering only pupae that did not detach or broke during adhesion assay (Figure 3.18).

Taken together, these results suggest that the glue surface of contact between the pupa and the substrate is responsible for the adhesion force interspecific variation observed.

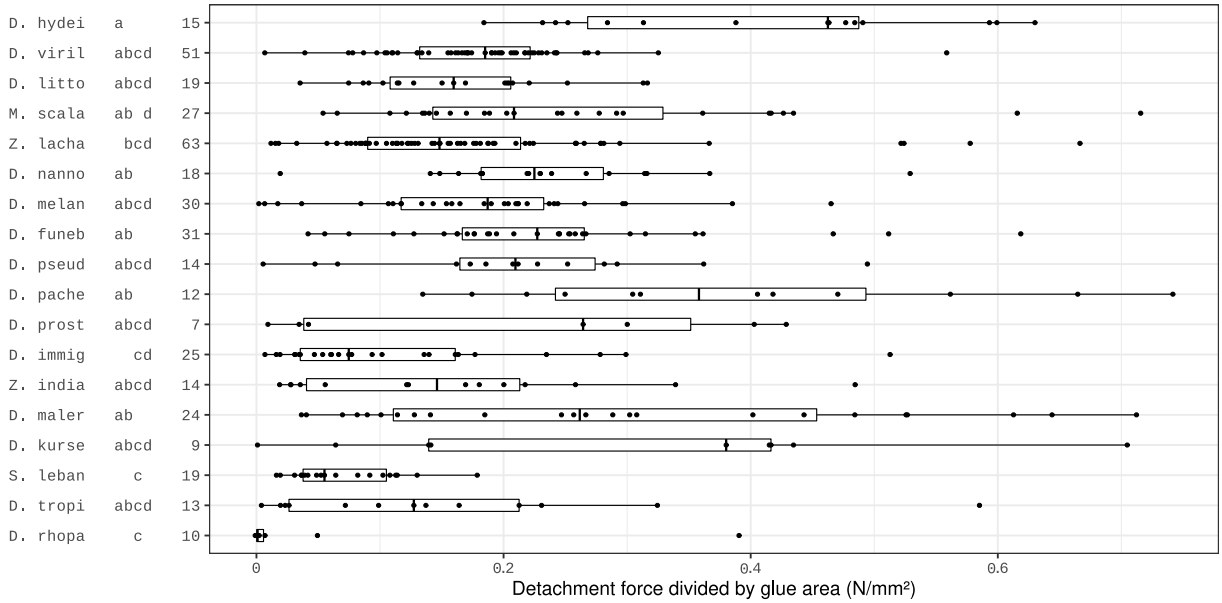


Figure 3.8: **Detachment force per unit area of glue for several Drosophila species.**
Same legend as Figure 3.5. For detached pupae, the detachment force is divided by the surface of the glue area in contact between the pupa and the glass slide. For pupae where no glue was observed on the glass slide, the glue area is null, therefore the detachment force per unit area can not be displayed. In this figure, 5% of the data being outliers are not displayed for better visualization.

3.4.5 Intraspecific adhesion properties variation is explained by experiment uncertainty

We compared adhesion properties between several strains for *D. suzukii*, *D. biarmipes* and *D. simulans* (Figure 3.19). For *D. suzukii*, the seven strains tested do not have significantly different detachment forces, they range between a median adhesion force of 0.05 N and 0.12 N. For *D. biarmipes*, two of the studied strains (Iso001 and Iso003) have significantly different detachment forces (median adhesion force below 0.025 N) from the third one (G224 median adhesion force is of 0.05 N). We conducted adhesion tests on detached pupae, for which the adhesion force should be null. However, the machine

recorded a maximal force for these pupae of 0.03 N, which thus represents the uncertainty of our measure. Given the low range of adhesion forces for Iso001 and Iso003 (both adhesion force median are lower than 0.025 N), we can assume that the determination of the adhesion force is within our experiment and algorithm uncertainty. For *D. simulans*, two of the three strains (vincennes and md221) have significantly different detachment forces between each other (median adhesion force of 0.22 N and 0.16 N respectively).

3.5 Discussion and perspectives

In this chapter, we analysed the force-versus-distance curves obtained from an adhesion assay protocol already used in the lab and were able to compare adhesion properties between pupae. We validated that this protocol is adapted to our measurements and found that the pupa is elastic while the pupa and the glue taken together are plastic. We can infer that plasticity is due to the glue as this property increases with the glue surface. We found that pupa adhesion in constant until 3 days after the L3 wandering stage and that the animal is more rigid when it ages. We plan to include pupa thickness measures to our analysis to normalize the rigidity of the pupa and its glue.

We found that the adhesion properties vary greatly between the 27 Diptera species tested, without correlation with the phylogeny. Measurements of pupa pictures revealed that this variation is not due to pupa morphology. We found that a higher surface of glue between the pupa and the substrate correlates with higher adhesion. Normalization of the detachment force per unit area indicates that the glue adhesion properties of the species tested are similar. We wonder if the presence/absence of *Sgs* genes in the genome of these species can be linked to the adhesion properties variation. We will compare the detachment force per unit area with the presence/absence of genes, annotated in the genomes in Da Lage et al. (2019) and in our study in Chapter 2. As *Sgs* genes were not annotated in the genome of some of the species used in this study, we plan to annotate them. To understand how *Sgs* genes can affect adhesion, we can inhibit their expression in *D. melanogaster* and test their adhesion force (see Chapter 4).

We can suppose that adhesion variation among species is an adaptation to their ecological site of pupation. Pupation site varies greatly between species, as discussed in section 1.3.3. As the pupa is vulnerable during metamorphosis, being able to stick to its substrate of pupation is of high importance for pupa survival. However, little information about the substrates on which Drosophila species pupate is available. We plan to include the available information and to try to link it to adhesion variation.

Our results reveal that *D. virilis* and *D. hydei* display the strongest adhesion force and the higher surface of glue, suggesting their salivary glands produce larger amounts of glue than other species. *D. virilis* and *D. hydei* salivary glands present a relatively low number of cells (respectively 110 and 82 per gland) compared to species from *D.*

melanogaster group (approximately 130 per gland) that present a medium adhesion force (Babišová et al., 2023). Among the species tested in this study, we could not find a link between the salivary glands number of cells and pupa adhesion force. *D. virilis* genome presents only three *Sgs* gene copies, *Sgs3a*, *Sgs3b* and *Sgs5bis* (Da Lage et al., 2019). *Sgs* genes were not annotated in *D. hydei* genome.

This study also reveals that Diptera species are good models to study adhesion, as these organisms are easily raised in laboratory conditions, are widely studied in genetics, and secrete a strong, biodegradable and biocompatible adhesive lasting for several days. We envisage using Diptera glue for future applications in industry to develop a new adhesive safe for human health and the environment. Based on our results, is seems that *D. virilis* and *D. hydei* glues are promising: they display the strongest adhesion force and *D. virilis* has relatively few *Sgs* genes.

3.6 Supplementary figures

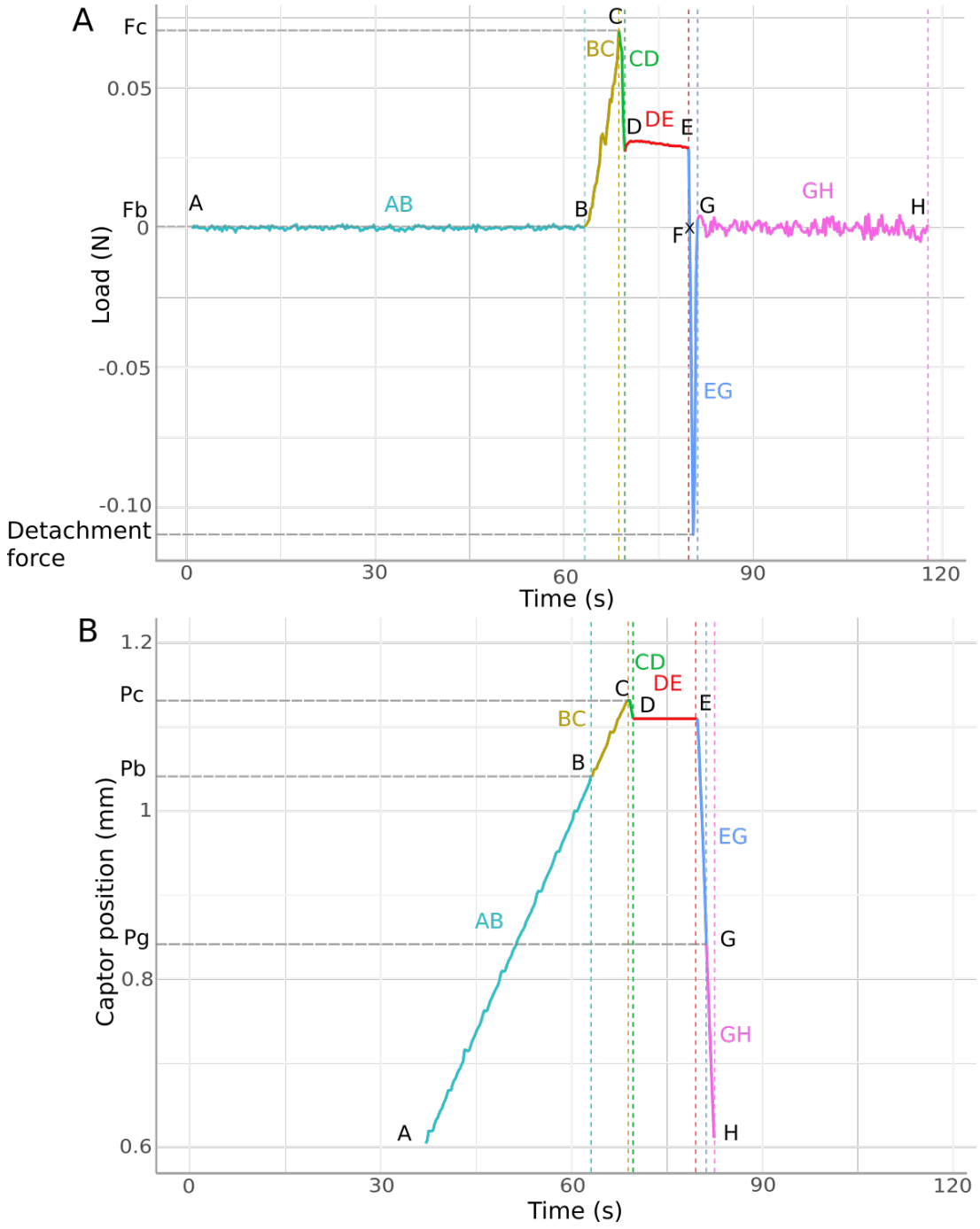


Figure 3.9: **Adhesion test curves of the force and position of the captor during the experiment.**

Same legend as 3.2. (A) Curve of the force recorded by the captor during the detachment experiment. (B) Curve of the captor position during the detachment experiment.

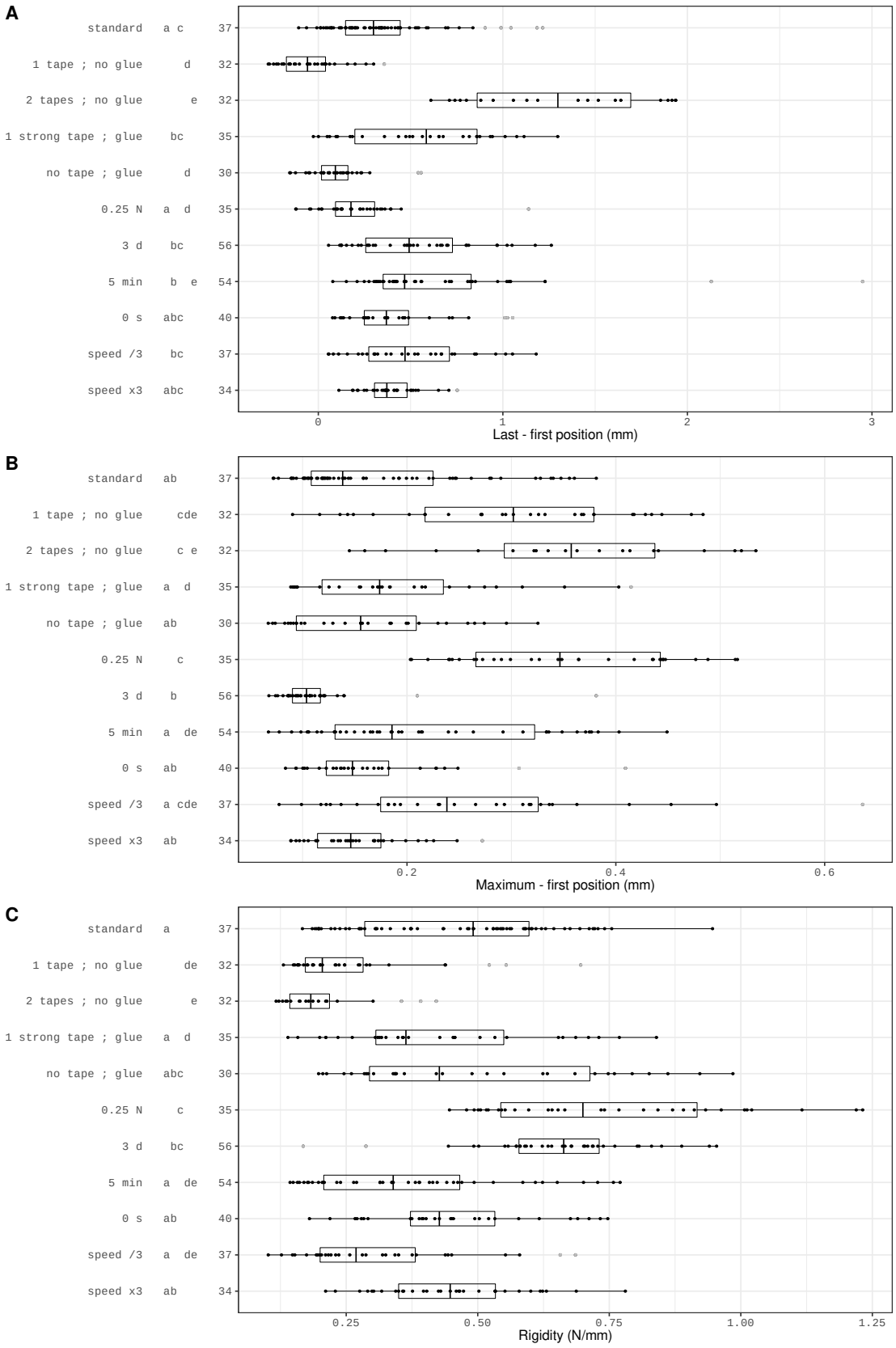


Figure 3.10: **Adhesion properties for control adhesion tests for *D. melanogaster*.**

See legend below

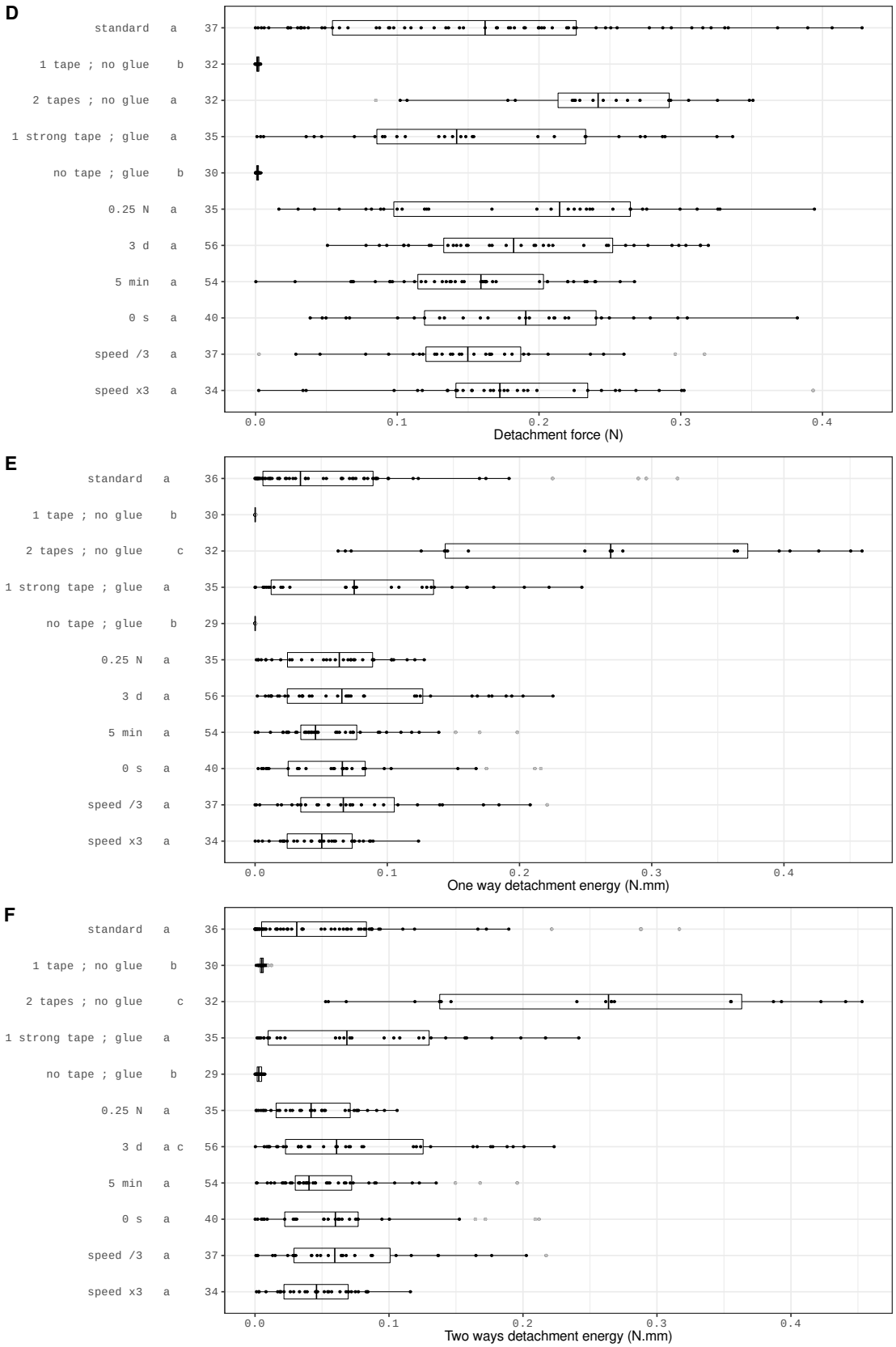


Figure 3.11: **Adhesion properties for control adhesion tests for *D. melanogaster*.**

Same legend as Figure 3.5

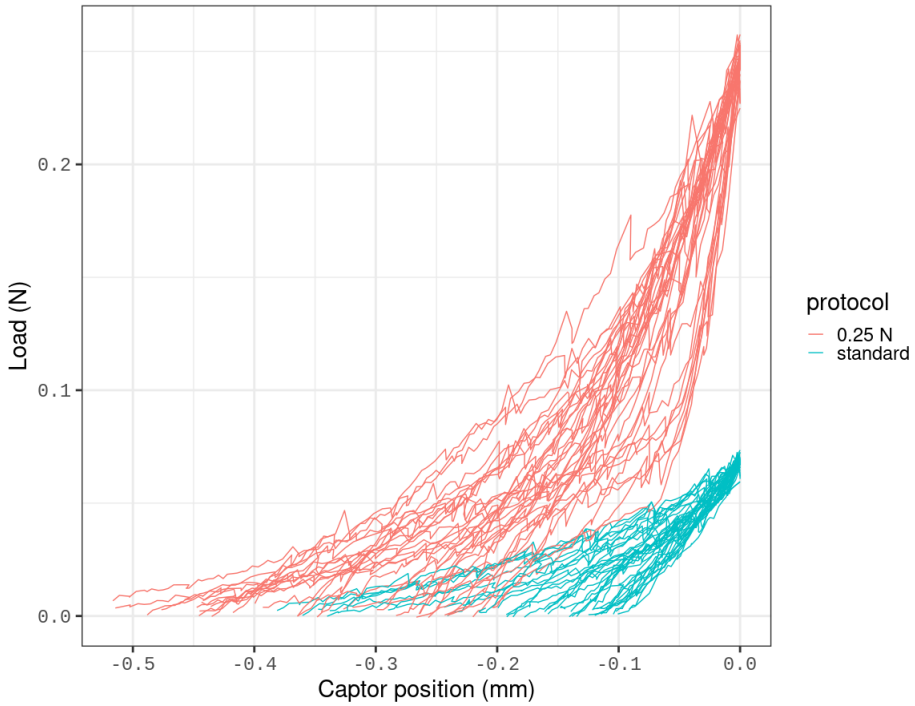


Figure 3.12: **Superposition of *D. melanogaster* detachment curves for standard and 0.25 N protocols.**

The force measured (load in N) during captor movement (captor position in mm) is given between landmarks B and C. Curves of pupae detached with a standard and 0.25 N protocol are coloured in blue and red respectively. Curves corresponding to standard protocols are linear while the ones corresponding to 0.25 N display two slopes, with a breakpoint between 0.07 N and 0.1 N.

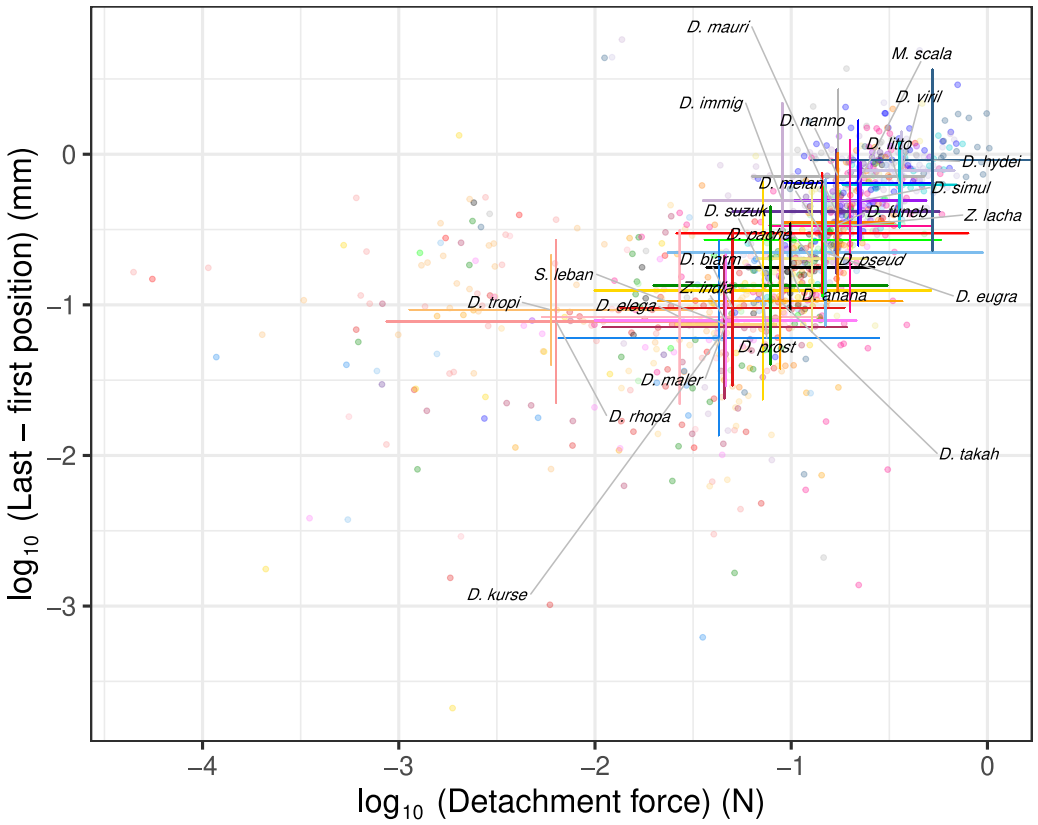


Figure 3.13: **Positive correlation between the detachment force and the difference of position between the first and the last contact for several species.** Each dot corresponds to a single pupa. Each species is represented by a different colour and its median is tagged. Center of crosses correspond to the median of each species. Horizontal and vertical bars respectively represent standard deviation for horizontal and vertical axes.

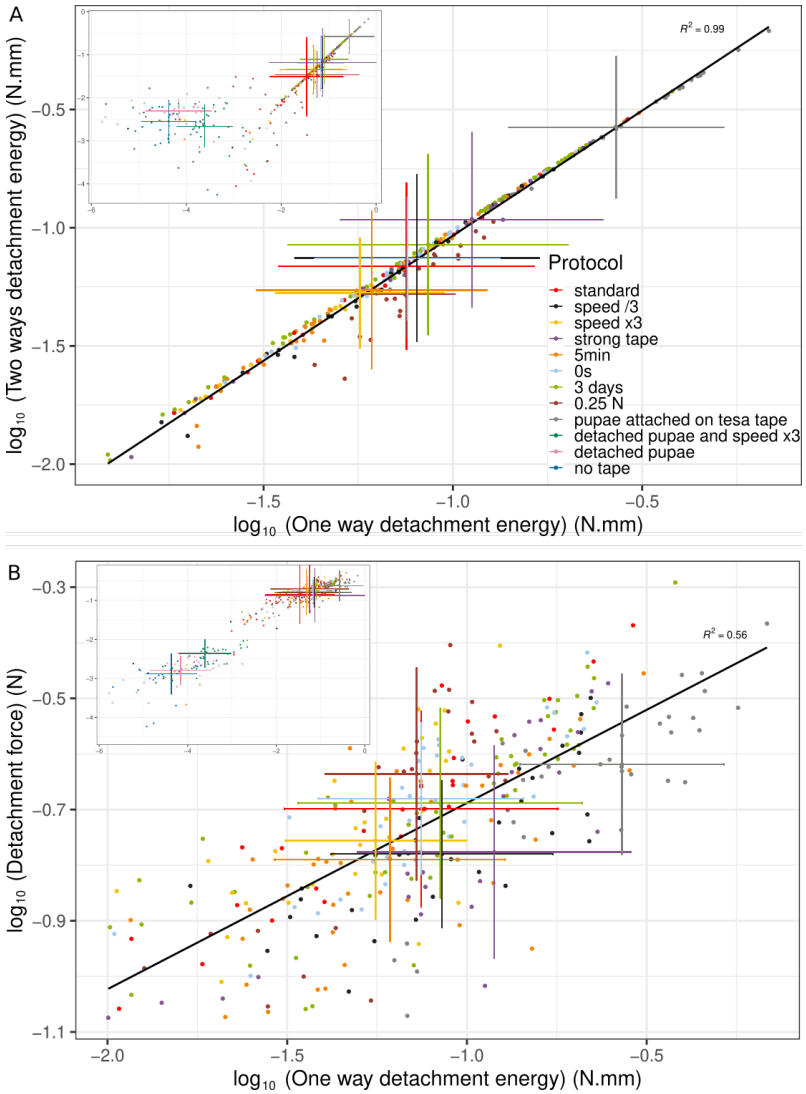


Figure 3.14: **Positive correlation between the detachment force, two ways detachment energy and one way detachment energy for control experiments conducted on *D. melanogaster*.**

(A) Correlation between log_{10} (two ways detachment energy) and log_{10} (one way detachment energy) (N.mm), (B) Correlation between log_{10} (detachment force) (N) and log_{10} (one way detachment energy) (N.mm). Main graphs are zooms of the inserted graph. Regions under the experiment uncertainty (0.03 N) are not displayed on zoomed graphs. Regression lines are drawn in black and regression coefficients are given. Protocols used for control experiments are distinguished with colours. Each dot corresponds to a single pupa. Center of crosses correspond to the median of each protocol. Horizontal and vertical bars respectively represent standard deviation for horizontal and vertical axes.

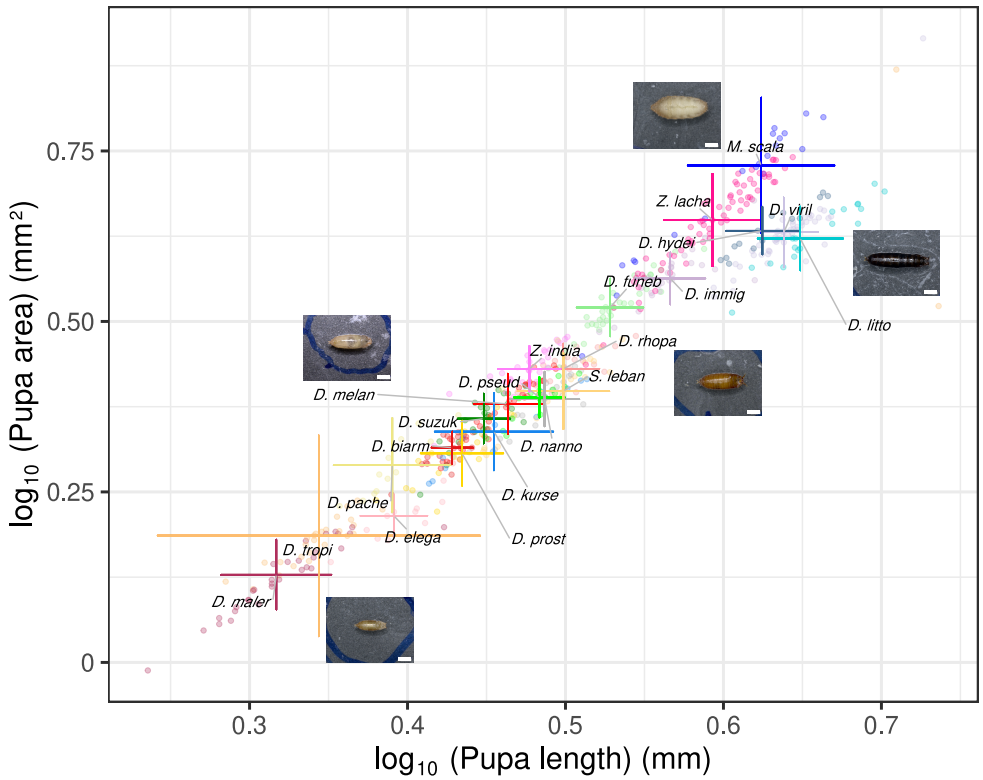


Figure 3.15: **Positive correlation between pupa length and pupa area.** Same legend as Figure 3.13. Pictures of pupae along the correlation are given. Scale bar is 1 mm.

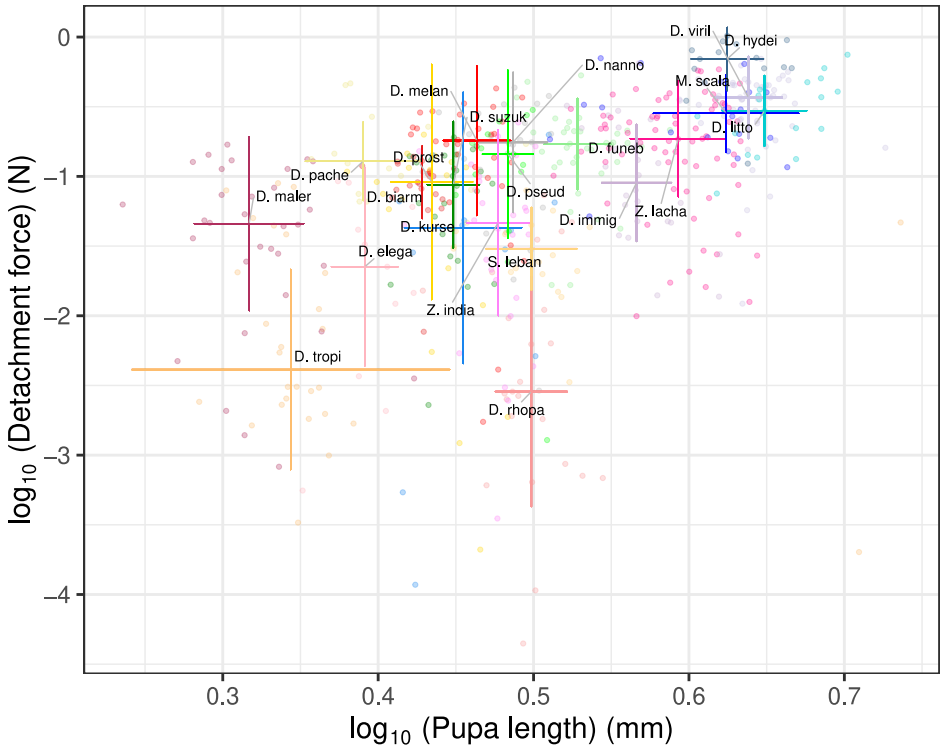


Figure 3.16: **Absence of correlation between pupa length and detachment force.** Same legend as Figure 3.13.

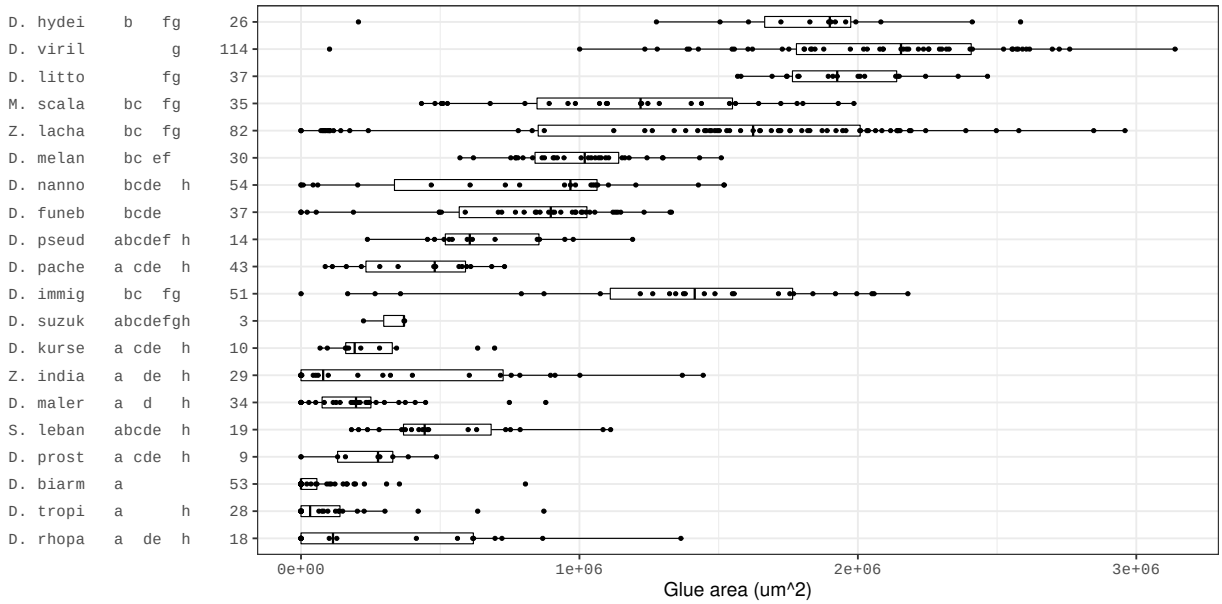


Figure 3.17: **Area of the glue surface of contact between the pupa and the glass slide.**

Same legend as Figure 3.5.

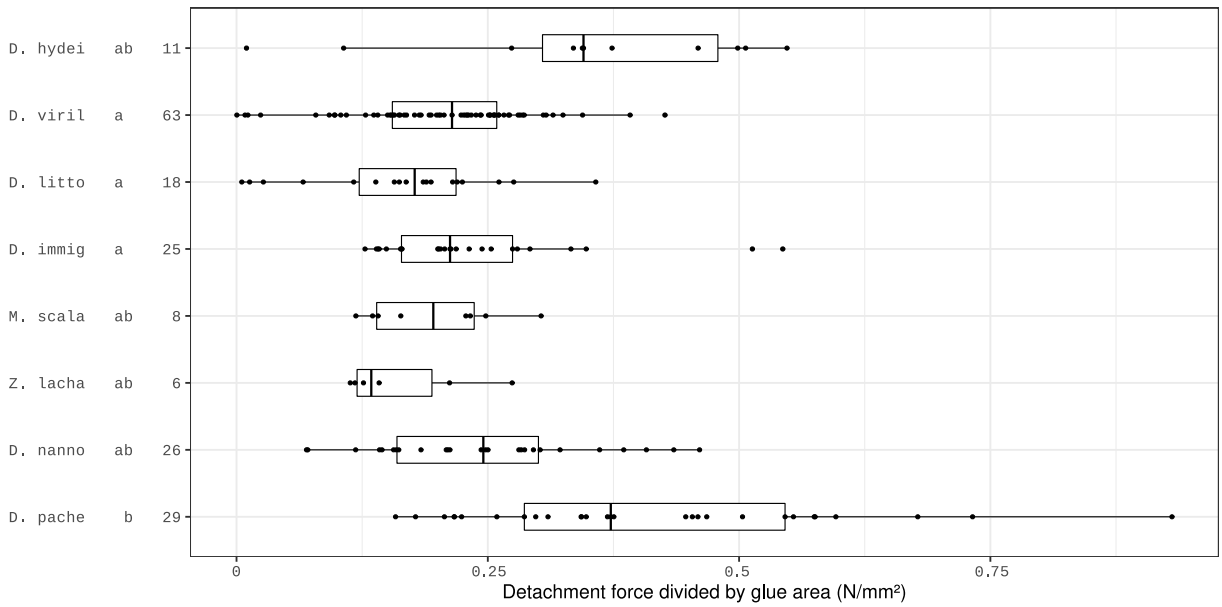


Figure 3.18: **Maximal force per unit area of glue for pupae not detached or broken after adhesion assay.**

Same legend as Figure 3.5.

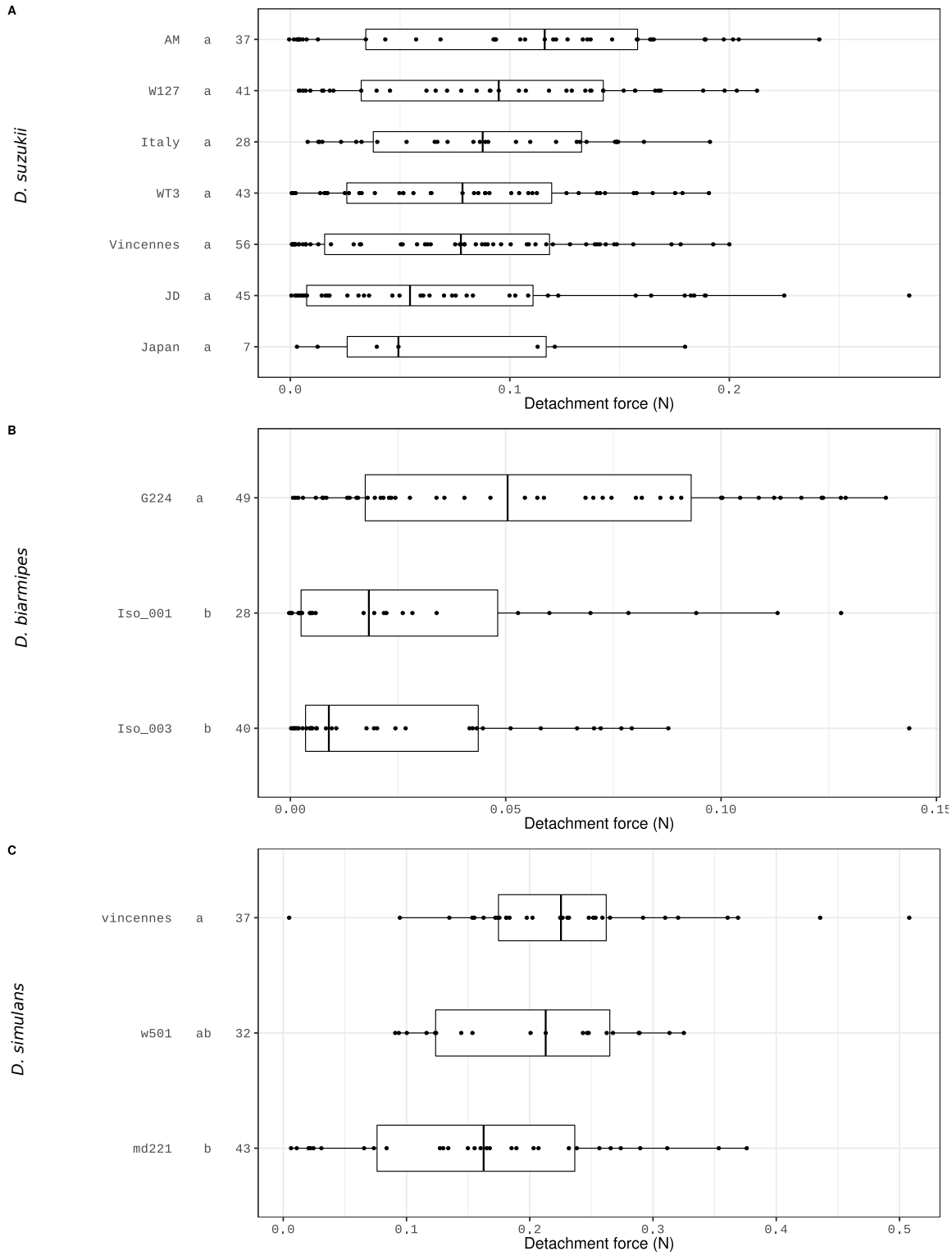


Figure 3.19: **Intraspecific variation of detachment force for *D. suzukii*, *D. biarmipes* and *D. simulans*.**

Same legend as Figure 3.5.

3.7 Supplementary tables

Table 3.1: **List of Diptera species used in this study.** Stocks, their origin and temperature of culture (°C) are given.

Species	Stock	Origin	Temperature of culture (C°)
Drosophila melanogaster	CantonS	Gift from Antoine Guichet's lab	25
Drosophila eugracilis		Collected in India, Karnataka, Bengaloru, GKVK campus in January 2013. B Prud'homme/N Gompel legacy	22
Drosophila biarmipes	G224	Collected in Cambodia,Ari Ksatr in May 1967. Genome strain 1 derived from San Diego Stock Center	22
Drosophila biarmipes	Iso_001	Given by B Prud'homme	22
Drosophila biarmipes	Iso_003	Given by B Prud'homme	22
Drosophila takahashii	14022-0311.07 Isofemale #01 and #03	Ulu Temburong National Park, Brunei (2003), Donated by Artyom Kopp (04/2004). Collection in banana or durian baits. Male morphology check by Castrezana (11/2005)	22
Drosophila elegans		San Diego Stock Center 14027-0461.03, Collected in Hong Kong and donated by A Kopp in October 2009,	22
Drosophila rhopaloa		Vietnam, Hanoi Ba Vì, near Vân Hòa, March 2005. H. Takamori leg	22
Drosophila prostipennis		San Diego Stock Center 14022-0291.00. Collected in Taiwan, Wulai, in June 1968. Lynn Throckmorton legacy.	22
Drosophila ananassae	Isofemale #01	India, Karnataka, Bengalory, GKVK campus. Collected by B Prud'homme and N Gompel	22
Drosophila kurseongensis	SaPa058	Collected in Vietnam, Lao Cai, Sa Pa in March 2009. H Takamori legacy.	22

Drosophila funebris		From approximately 10 larvae collected by Manon Monier in Pennautier, France in July 2021	22
Drosophila pachea	15090-1698.01_14.2	Arizona – Given by Michael Lang	25
Drosophila nannoptera	15090-1692.00	Oaxaca – Given by Michael Lang	25
Scaptodrosophila lebanonensis		Given by S Prigent – collected by Jean David in Prunay	25
Drosophila pseudoobscura	SKU : 14011-0121.94	Cornell Drosophila Stock Center	25
Drosophila hydei		Vincennes park (Paris, France). Collected by S. Pringent in 2020. (Borne et al., 2021b)	22
Drosophila immigrans	Vincennes	Vincennes park (Paris, France). Collected by S. Pringent in 2020. (Borne et al., 2021b)	25
Drosophila virilis	WT	Drosophila Species Stock Center #15010-1051.86, given by Ashley Farlow from Christian Schloetterer's lab in 2010. Origin: Texmelucan, Puebla, Mexico.	22
Drosophila littoralis		Given by S. Prigent	25
Drosophila malerkotliana		Given by S. Prigent	25
Drosophila mauritiana	mau12	given by Peter Andolfatto May 2020	25
Drosophila simulans		Vincennes park (Paris, France). Collected by S. Pringent in 2020. (Borne et al., 2021b)	25
Drosophila suzukii	Vincennes	Vincennes park (Paris, France). Collected by S. Pringent in 2020. (Borne et al., 2021b)	25
Drosophila suzukii	Italy	Italy (no further details). Given by Patricia Gibert.	25
Drosophila suzukii	AM	Alpes Maritimes (France). Given by Patricia Gibert.	25
Drosophila suzukii	WT3	Watsonville (USA).	25
Drosophila suzukii	W127	Watsonville (USA). Collected by V. Debat and S. Fellous in 2014.	25

Drosophila suzukii	JD	France (no further details). Collected by J. David in 2020.	25
Drosophila suzukii	Japan	Japan (no further details). Given by Patricia Gibert.	25
Drosophila tropicalis			22
Drosophila yakuba	Tai18E2	UCSD stock center 14021-0261.01 Tai forest border of Liberia and Ivory Coast, 1983 given by Peter Andolfatto May 2020	25
Megasalia scalaris		Gift from Jean David	25
Zaprionus lachaisei		Given by S. Prigent	25
Zaprionus indianus		Given by S. Prigent	25

Table 3.2: **Alternative adhesion assay protocols used.** Rows are listing the different protocols used. Columns are listing the parameters of protocols. Standard protocol is the reference protocol. For other protocols, parameters similar to standard protocol are represented by a hyphen (-). Parameters modified with respect to the standard protocol are given. Arrows pointing up and down respectively corresponding to the captor going up and down.

Protocol	**Speed**	**Pause**	**Maxima force**	**Pause force**	**Type of tape**	**Presence of glue**	**Tape on the captor**	**Pupa age**
standard	↓1 mm/min ↑0,2 mm/s	10 s	0.07 N	0.03 N	Tesa	Yes	Yes	15 to 20 hours after larva deposition
1 tape ; no glue	-	-	-	-	-	Pupa detached	-	-

2 tapes ; no glue	-	-	-	-	-	Pupa detached and placed on double sided Tesa tape	-	-
1 strong tape ; glue	-	-	-	-	Fermoflex		-	-
no tape ; glue	-	-	-	-	-	-	No	-
0,25 N	-	-	0,25N	0,21N	-	-	-	-
3 d	-	-	-	-	-	-	-	3 days after larva deposition
5 min	-	5 min	-	-	-	-	-	-
0 s	-	0 s	-	-	-	-	-	-
Speed /3	↓0,33 mm/min ↑0,06 mm/s	-	-	-	-	-	-	-
Speed x3	↓3 mm/min ↑0,6 mm/s	-	-	-	-	-	-	-
1 strong tape ; glue ; 0.25 N	-	-	0,25N	0,21N	Fermoflex		-	-

Chapter 4

Sgs genes role in *D. melanogaster* pupa adhesion

4.1 Abstract

In this section, we aim at understanding the role of *Sgs* genes in *D. melanogaster* pupa adhesion. Sgs proteins are the main components of the glue but their implication in adhesion is still unknown.

We individually inhibited each glue gene expression using the genetic tool UAS-GAL4 combined with UAS-RNAi lines. Since adhesion varies greatly between *D. melanogaster* strains (from 124 mN to 377 mN) (Borne et al., 2021a), we paid particular attention to the genomic backgrounds and always compared an RNAi targeting a given glue gene to another RNAi targeting a control gene in exactly the same genomic background. To try to avoid this issue with genomic backgrounds, we also tested a second technique, Gene-Switch, in which GAL4 is only activated in presence of RU486 in the food medium.

Overall, the goal of these experiments, is to decipher which glue genes are involved in adhesion to envisage the production of a future synthetic bioadhesive.

4.2 Introduction

Despite their name, the glue genes have not yet been shown to be involved in adhesion. They are expressed in the salivary glands and the Sgs proteins are present in the secreted glue (Borne et al., 2021a; Korge, 1975). Other small molecules, such as amino-acids and sugars, or *new glue* genes described in Chapter 2, could be involved in glue adhesion. So far, a few studies have analysed the role of glue genes in adhesion. Our work showing glue genes and pupa adhesion rapid evolution is in agreement with the hypothesis that glue genes are important for glue properties.

To study the effect of a specific gene on a given phenotype, the genetic tool UAS-GAL4 (Brand and Perrimon, 1993), combined with UAS-RNAi lines, is commonly used in *D. melanogaster*. The system is made of two parts: the GAL4 gene encoding the yeast transcription activator protein GAL4, and the UAS (Upstream Activation Sequence), an enhancer to which GAL4 specifically binds and activates transcription of a target gene (Figure 4.1). The UAS-RNAi lines are transgenic *D. melanogaster* stocks that present an insertion of two identical sequences of the target gene to be silenced : one is inserted sense and the second antisense. The product of this gene is an RNAi, for RNA interference, that will degrade the product of the gene of interest, and therefore silence it.

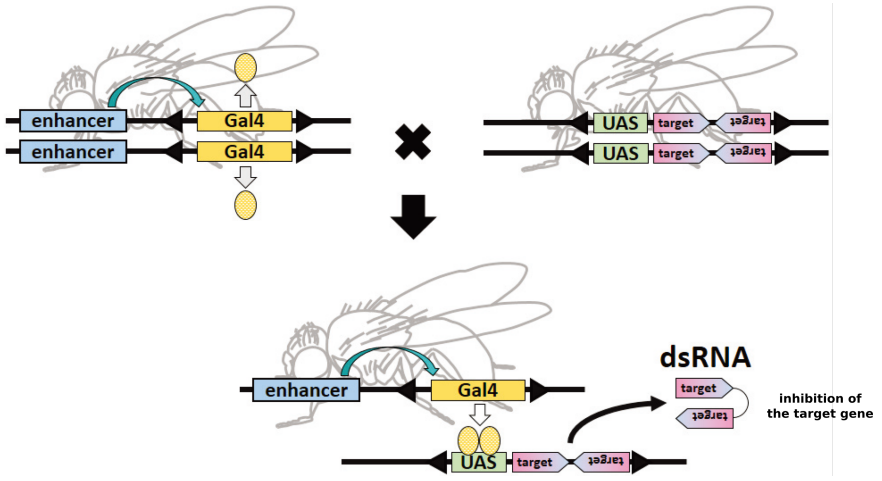


Figure 4.1: **Scheme of UAS-GAL4 system used for RNAi gene expression silencing in *D. melanogaster*.**
Two fly lines are crossed. One line presents an enhancer controlling the expression of the GAL4 gene at a precise time in a given tissue. A second line presents an Upstream Activation Sequence (UAS) activating the transcription of a double stranded RNA (dsRNA). This dsRNA comprises sequences identical to parts of a target gene (both sense and antisense). In the progeny, GAL4 transcription factor specifically binds to UAS sequences, activating the transcription of the dsRNA. The dsRNA leads to the inhibition of the expression of the target gene (Yamamoto-Hino and Goto, 2013).

Using this technique, a previous study observed a reduction in pupa adhesion in *D. melanogaster* after inhibition of the expression of a clathrin heavy chain gene (*AP-1*),

involved in glue granule formation (Burgess et al., 2011).

Our lab has been a pioneer in developing a technique to measure pupa adhesion (Borne et al., 2020), and so far no other study of glue adhesion has been published by other groups. Flora Borne, a previous PhD student in the lab, started to use UAS-RNAi lines to decipher the role of *Sgs* genes in glue adhesion (unpublished). In her experiments, the GAL4 line 1824 available at the Bloomington Drosophila Stock Center and driving expression in the salivary glands, was used. *Sgs1* and *Sgs3b* expression was inhibited independently by crossing GAL4 and UAS-RNAi lines having different genomic background between eachother. Reverse Transcriptase-quantitative PCR (q-RT-PCR) were performed in order to mesure their level of expression. *Sgs3b* expression level was high in *SG-GAL4* and reduced in both the *UAS-RNAi-Sgs3b* line and the *SG-GAL4 ; UAS-RNAi-Sgs3b* line. These results indicate that *UAS-RNAi-Sgs3b* does not express *Sgs3b* or that the UAS promoter is leaky. For *Sgs1*, the expression of *Sgs1* was high in *SG-GAL4* and reduced in *SG-GAL4 ; UAS-RNAi-Sgs1* and reduced further in *SG-GAL4 ; UAS-RNAi-Sgs1/UAS-RNAi-Sgs1* (two copies of the *UAS-RNAi-Sgs1* transgene), suggesting that the GAL4-UAS-RNAi system inhibits *Sgs1* expression as expected.

Pupa adhesion was found to be significantly reduced in *SG-GAL4 ; UAS-RNAi-Sgs1/UAS-RNAi-Sgs1* pupae (110 mN) compared to *SG-GAL4* and *SG-GAL4 ; UAS-RNAi-Sgs1* pupae (200 mN). Similarly, pupa adhesion was significantly altered in *SG-GAL4 ; UAS-RNAi-Sgs3b* (110 mN) compared to *SG-GAL4* (200 mN). The difference in adhesion that was detected between the wild-type control stock and the RNAi-mediated *Sgs* gene expression inhibition experiment is of the same order than the one previously observed between *D. melanogaster* lines (Borne et al., 2021a). Indeed, accross 12 *D. melanogaster* lines originating from diverse regions, the most adhesive ones had a median adhesion force of 377 mN and the lowest 124 mN. Therefore, the genomic background can cause large variations in adhesion. The effect previously detected with these RNAi experiments are going in the expected direction but they may be caused by the genetic background and not by alteration of the *Sgs* gene expression levels.

In this section, we aimed at testing the role of each *Sgs* and *new glue* genes in adhesion via two different approaches where the genetic background is controlled.

A first approach consisted in using the GAL4 UAS-RNAi system to inhibit *Sgs* expression. We used *UAS-RNAi* lines and crossed them to a unique GAL4 line. The resulting progeny pupae were tested for their adhesion and, as they have the same genetic background, compared together.

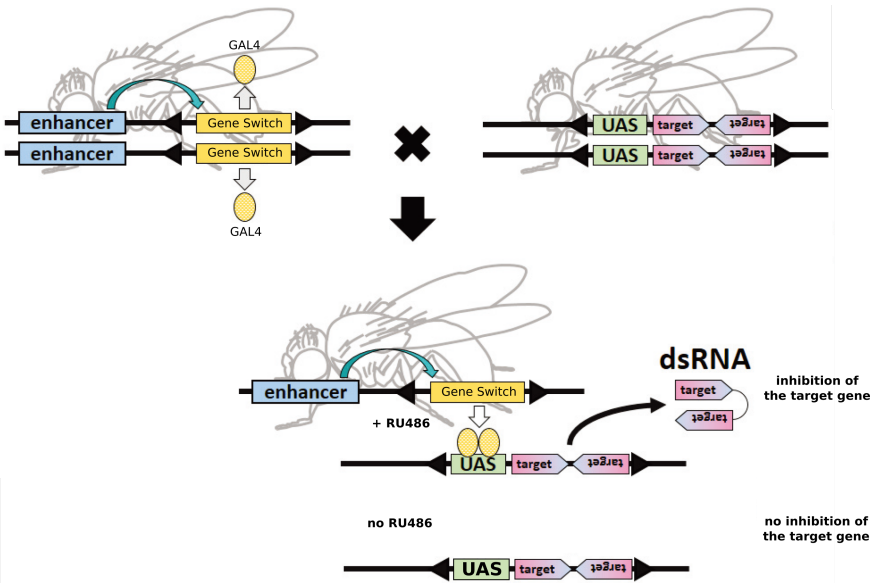


Figure 4.2: **Scheme of the Gene Switch system silencing in *D. melanogaster*.** Gene Switch system is based on the UAS-GAL4 system (Figure 4.1). The modified GAL4 protein is inactive in absence of the antiprogestin RU486. In its presence, the modified GAL4 protein can activate transcription. It will bind to UAS sequencesn activate dsRNA transcription and lead to inhibition of the expression of the target gene. Adapted from (Yamamoto-Hino and Goto, 2013)

In a second approach, we used the Gene-Switch technique. This technique enables to turn on and off the expression of a transgene at the desired time in a given tissue (Roman et al., 2001) (Figure 4.2). Gene-Switch is a chimeric gene, it is composed of portions of three coding sequences, a GAL4 DNA-binding domain, the human progesterone receptor-ligand-binding domain, and the activation domain from the human protein, p65. The antiprogestin RU486 acts as an activator. When it is present, the chimeric GAL4 protein binds to a UAS sequence, and activates the downstream target genes. When it is absent, there is no transcription of the target genes. We used this technique to analyse the adhesion of *UAS-RNAi* lines that are not viable at the homozygous state and were crossed with balancer lines. The genetic background between these lines differ and their adhesion can not be compared together. Gene-Switch technique is particularly useful to get rid of the influence of the genetic background: the same line is being compared with and without RU486, thus in the same genomic background. The difference observed between the two lines corresponds to the effect of the inhibition induced by RNAi.

To test the contribution of each *Sgs* gene to adhesion properties, we performed adhesion assays on pupae resulting from crosses between a GAL4 or GAL4-Switch drive and a UAS-RNAi reporter line. This study was done with the help of Isabelle Nuez, engineer in the lab, and in collaboration with Kelly Ten Hagen from the National Institutes of Health (Bethesda, USA) who analyses the consequence of RNAi on the glue aspect via electron microscopy.

4.3 Material and methods

4.3.1 Fly culture

Three different culture medium were used. Medium 1 is a standard medium (4 l: 83.5 g yeast, 335.0 g cornmeal, 40.0 g agar, 233.5 g saccharose, 67.0 ml Moldex, 6.0 ml propionic acid). Medium 2 is the same as medium 1 with 25 μL of 96% ethanol added in 50 mL of medium. Medium 3 is the same as medium 1 and is supplemented with RU486. A stock solution at 20 mg/mL was made with 1.56 g of RU486 in powder diluted in 96% ethanol. 25 μL of that solution was added in 50 mL of medium, resulting in a final concentration of 10 μg/mL of RU486 in culture medium.

4.3.2 Fly lines and crosses

UAS-GAL4

We used the Gal4 line *c135-Gal4 , Sgs3b-GFP*, kindly provided by Kelly Ten Hagen (Syed et al., 2022). C135 drives expression of GAL4 active in the salivary glands at the wandering stage and this line also expresses the Green Fluorescent Protein (GFP), under the control of *Sgs3b* cis-regulatory elements. For *UAS-RNAi* lines, we used medium 1.

We used *UAS-RNAi* lines available at Bloomington Drosophila Stock Center. Their targeted genes, stock number, and transgene site of insertion are given in the Table 4.1. These lines were created from the same fly stock by Perkins et al. (2015), by inserting *pVALIUM20* and *pVALIUM10 UAS-RNAi* constructs into two attP sites: either attP40 on the left arm of the chromosome 2 at 25C6 or attP2 on the left arm of the chromosome 3 at 68A4.

Table 4.1: **List of the *VALIUM UAS-RNAi* lines targetting glue genes and other candidate genes.** The targetted genes, lines and site of insertion of the transgene are given. These lines are available at the Bloomington Drosophila Stock Center.

Gene to target	Line	Insertion site
Sgs1	*pVALIUM20-UAS-dsRNA-Sgs1* (51421)	Chr 3
Sgs3b	*pVALIUM20-UAS-dsRNA-Sgs3b* (62944)	Chr 2
Sgs4	*pVALIUM20-UAS-dsRNA-Sgs4* (63675)	Chr 2
Sgs5	*pVALIUM20-UAS-dsRNA-Sgs5* (64857)	Chr 2
Sgs8	*pVALIUM20-UAS-dsRNA-Sgs8/CyO* (65165)	Chr 2
CG33272	*pVALIUM20-UAS-dsRNA-CG33272/CyO* (60411)	Chr 2
pgant9	*pVALIUM20-UAS-dsRNA-pgant9/TM3, Sb[1]* (50617)	Chr 3
sage	*pVALIUM10-UAS-dsRNA-Sage* (25980)	Chr 3

Gr59e	*pVALIUM10-UAS-dsRNA-Gr59e* (25815)	Chr 3
Gr33a	*pVALIUM20-UAS-dsRNA-Gr33a* (62940)	Chr 2

We crossed *c135-Gal4 , Sgs3b-GFP* virgin females with UAS-RNAi males and measured adhesion of the resulting progeny.

As a positive control, we used a *UAS-RNAi-sage* line. *Sage* is only expressed in the salivary glands and its RNAi silencing leads to the downregulation of genes expressed in the salivary glands at the wandering stage (Li and White, 2003). We expect its inhibition to lead to the downregulation of *Sgs* genes and reduce pupa adhesion.

As a negative control, we used *UAS-RNAi* lines targeting the genes *Gr59e* and *Gr33a*. These genes both encode for gustatory receptors and are not expressed in the salivary glands at the wandering stage. Their inhibition should not lead to any change in pupa adhesion.

Gene-Switch

For Gene-Switch experiments, we tested two GAL4 lines, *GAL4-Switch-40990* (40990) and *GAL4-Switch-40259* (40259), available at Bloomington Drosophila Stock Center. Both lines were reported to drive expression of GAL4 in the salivary glands and in other organs (Nicholson et al., 2008). To confirm their expression in salivary glands, we crossed both of them with a *UAS-mCD8-GFP* line. *mCD8-GFP* is a genetic construct that expresses *GFP* in cell membranes (Lee and Luo, 2001). Crosses were cultured on medium 2 and 3. For 20 progeny larvae tested, the line 40259 did not display any fluorescence in the salivary glands when supplemented with RU486. For 15 larvae cultured in presence of RU486, the line 40990 displays fluorescence in salivary glands and in some parts of the gut. Larvae from the line 40990 did not display fluorescence in absence of RU486 (n=20). As the *GAL4* line 40990 is the only one to express *GFP* in salivary glands, under activation by RU486, we decided to use this line for our experiments involving Gene-Switch.

We cultured RNAi lines on a medium supplemented in RU486 (medium 3) or on a medium deprived of it (medium 2).

As a control line, we used a CantonS strain, kindly given by Guichet lab.

4.3.3 Adhesion assay

Adhesion assays were performed using the standard protocol described in Chapter 3.

4.3.4 Statistical analysis

We compared adhesion force between genotypes by using Kruskal Wallis test and pairwise Wilcoxon test in R version 4.1.2 (2021-11-01) (https://www.r-project.org).

4.4 Results

4.4.1 *Sgs3b* inhibition reduces pupa adhesion

We conducted adhesion assays on pupae obtained from crosses between *c135-GAL4 ; Sgs3b-GFP* females and *UAS-RNAi* males. Among the tested lines, the ones inhibiting *Gr33a*, *CG33272*, *Sgs8*, *Sgs5*, *Sgs4* and *Sgs3b* present an insertion of the *VALIUM UAS-RNAi* construct on the chromosome 2, at the site attP40. The other lines inhibiting *Gr59e* and *sage* have the *VALIUM UAS-RNAi* construct inserted on the chromosome 3 at the site attP2. The lines having the *VALIUM UAS-RNAi* construct inserted on the same chromosome have the same genomic background and can thus be compared together for their adhesion. Silencing of *Gr33a* and *Gr59e* gave a median detachment force respectively of 223 and 227 mN (Figure 4.3). As these genes are not related to adhesion and not expressed in salivary glands, their detachment force represent the wild-type condition with normal activity of *Sgs* genes. Silencing of *sage* led to a significant decrease in adhesion with a median detachment force of 58 mN (Figure 4.3). This results confirms that *sage* acts as a regulator of *Sgs* genes expression (Li and White, 2003).

A strong and significant decrease in adhesion was also observed when *Sgs3b* was silenced with a median detachment force of 97 mN. *Sgs4* and *Sgs5* inhibition did not lead to a significant decrease in adhesion. In Chapter 2, *CG33272* was found 200 bp away from *Sgs7* in *D. melanogaster* genome and identified as one of the *new glue* genes. Inhibition of *CG33272* leads to a slight decrease in adhesion, with a median detachment force approximately of 170 mN, but not significantly different from the control lines *Gr33a* and *Gr59e*.

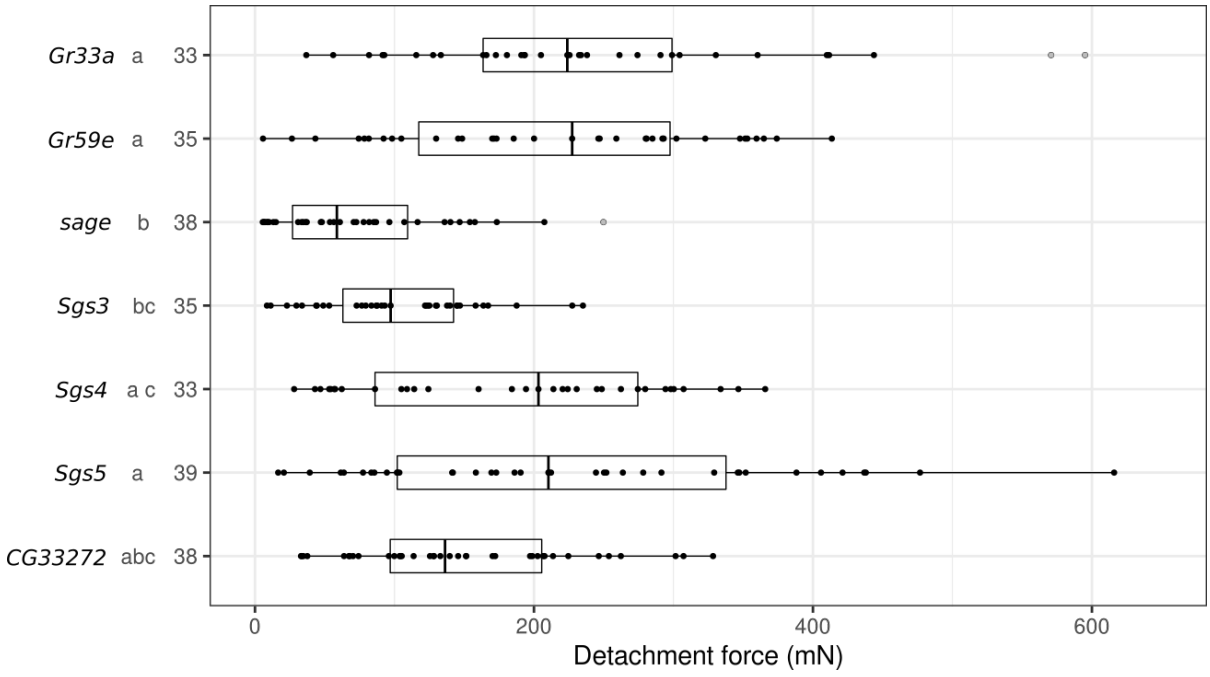


Figure 4.3: **Detachment force of *c135-GAL4 UAS-RNAi* pupae**

Each dot corresponds to a single pupa. Ends of the boxes define the first and third quartiles. The black vertical line represents the median. The horizontal line on the top of the box extends to the largest value no further than 1.5 × IQR from the upper hinge of the box. The horizontal line on the bottom of the box extends to the smallest value at most 1.5 × IQR of the hinge (IQR, inter-quartile range is the distance between the first and the third quartiles). Data beyond the end of these lines are 'outlying' points represented in grey. An ANOVA followed by all pairwise comparisons after Tukey correction (P 0.05) was performed on the set of protocol groups. Groups that are not significantly different from each other share a letter. Figures on the right side of statistical letters indicates the total number of pupae tested for each condition.

4.4.2 Gene-Switch experiments are not conclusive

For UAS-RNAi lines that are not viable at the homozygous state and have been crossed with a balancer line, we cannot use the crosses performed above to infer the role of candidate genes as these lines present a different genetic background. We decided to use the Gene-Switch technique (Roman et al., 2001) to assess the adhesion force without the UAS-RNAi activity in a given genomic background by adding or depriving the food medium in RU486.

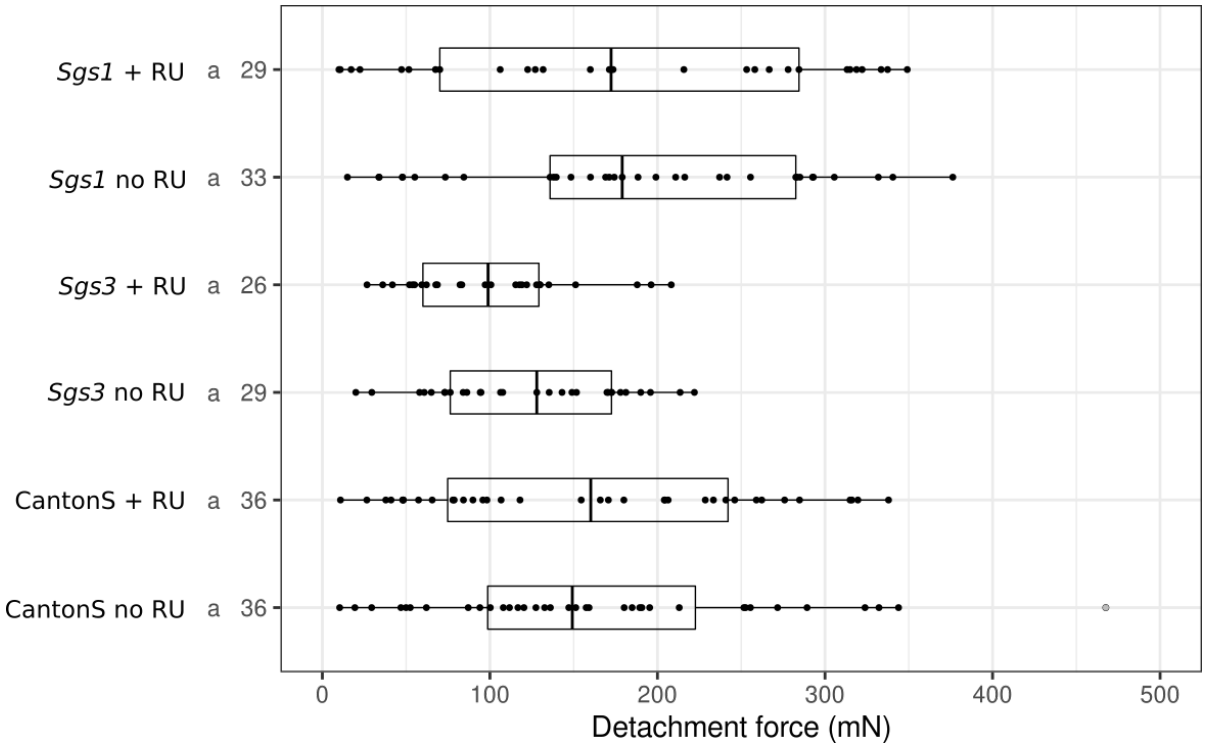


Figure 4.4: **Detachment force for Gene-Switch lines.** Same legend as Figure 4.3

.

We cultured the Gene-Switch UAS-RNAi lines and the CantonS strain on a medium supplemented in RU486 (medium 3) and a medium deprived of it (medium 2).

Adhesion assays were performed on pupae resulting from crosses between the GAL4-Switch and UAS-RNAi line (Figure 4.4). No significant difference of adhesion was observed between lines cultured on RU486 supplemented medium and lines cultured on a medium deprived of it. These results suggest that the RNAi expression was not activated.

4.5 Discussion

Our preliminary results reveal a strong decrease in adhesion when *Sgs3b* expression is silenced using RNAi lines, and a slight but not significant decrease when of one the *new glue* genes, *CG33272*, is silenced. We plan to measure the surface of glue between the pupa and the glass slide for the lines tested and to compare it with adhesion. We also plan to perform qRT-PCR on the wandering third instar larvae progeny from the crosses, to measure the level of expression of the target *Sgs* genes.

We also intend to analyse the adhesion properties of these pupae with the same R code used in Chapter 3. Indeed, it is possible that the rigidity and reversibility of the deformation might change according to the glue composition.

We would like to inhibit other *Sgs* and *new glue* genes expression using the Gene-Switch technique, as it may be a good way to test the genetic background of our lines. However, in our study the expression of the RNAi does not seem to be sufficiently high in presence of RU486. This can be due to a low *GAL4* activity. We plan to test other GAL4 Gene-Switch lines that would express *GAL4* in larger quantity. As observed in a previous study conducted by Flora Borne (unpublished), two copies of the *UAS-RNAi* construct might be needed to see an effect on adhesion. We plan to design lines that are homozygous for the *UAS-RNAi* construct. Finally, increasing RU486 concentration in the culture medium could increase its activity.

Chapter 5

Conclusion

The aim of my thesis was to better understand the evolution of Drosophila glue genes and adhesive properties. I first studied the evolution of glue genes between several Drosophila species and compared their evolution rate. In a second part, I studied adhesive properties for several Diptera species and examined pupa morphology and glue quantity. Finally, I inhibited individually the expression of glue genes to test their role in *D. melanogaster* pupa adhesion. Overall, my thesis aims at understanding better the evolution of glue genes and glue adhesive properties to develop Drosophila glue as a promising model for bioadhesion.

5.1 Evolution of the glue genes and their genomic regions

In this section, I annotated a total of 102 *Sgs* genes and identified *new glue* genes copies not yet annotated by analysing the genome of 24 Drosophila species. This study benefited from recently published genomes that are more accurate than the ones used in a previous study from the laboratory (Da Lage et al., 2019). We used *Sgs* sequences from closely related species in order to find orthologs and annotated more copies than the previous study. Finally, using *Sgs* genes synteny enabled us to find a new locus with *Sgs3* copies and the presence of "*new glue*" genes.

We observed two evolutionary dynamics among the glue genes. On one hand, the genes *Sgs3b*, *Sgs7* and *Sgs8* present a high rate of inversion, duplication, gene loss and conversion, and are in a region containing multiple *new glue* genes. On the other hand, *Sgs1*, *Sgs3x* and *Sgs3e* show a few gene losses but no duplication, no local inversion and no gene conversion. We suggest that the *new glue* genes could favor the rapid evolution of *Sgs3b*, *Sgs7* and *Sgs8*. Indeed, repeated sequences can promote duplications and deletions of genes (Hastings et al., 2009). Our work proposes that the rapid evolution of some *Sgs* genes can be due to the mutation process inducing structural changes in the genome.

We can suppose that some of the *Sgs* duplicates, if they are expressed, have gained new properties that could be involved in glue adhesion. As the *new glue* genes appear to be linked to *Sgs* genes evolution, they could also be involved in glue production and its adhesion. In *D. melanogaster*, *Sgs4* is surrounded by three *new glue* genes (*ng-1*, *ng-2* and *ng-3*) that are expressed exclusively in salivary glands (Furia et al., 1993) between the beginning of the third larval instar until the early wandering stage (D'Avino et al., 1995). In this study, I did not analyze the evolutionary dynamics of *Sgs4*. It would be interesting to know more about the rate of *Sgs4* evolutionary dynamics and whether the presence of these *new glue* genes is associated with a high gene turn over in *Sgs4* as well as in *Sgs3*, *Sgs7* and *Sgs8*.

In our study, all the 24 Drosophila species genomes possess at least one *Sgs3* copy, and most of them have several. This could suggest an important role played by Sgs3 protein in pupal survival. In the previous study conducted in the laboratory (Da Lage et al., 2019), *Sgs5bis* was also found in many different species and was supposedly present at the ancestral state with *Sgs3*. It thus seems that there are two important glue genes that are present in all species, *Sgs3* and *Sgs5bis*.

5.2 Adhesive properties of several Diptera species

5.2.1 The adhesive properties vary rapidly between species

In this second part, we studied adhesive properties of 27 Diptera species. We used an automated adhesion assay program developed in the lab to measure the adhesion force and we created a new R script analysis to assess additional properties relative to the pupa and its glue. By comparing these properties between species, we would like to understand better glue properties evolution and find promising species that could be useful for future applications.

Calculation of glue properties revealed that the detached pupa is elastic whereas the attached pupa is plastic. We found that the rigidity of the animal and its glue is not correlated with adhesion force. We calculated the energy of detachment, which gives a proxy of the energy necessary for a predator to detach pupae from their substrate. This energy increases with the adhesion force, suggesting that a predator needs to provide more energy to detach strongly attached pupae.

We found that the adhesion force varies greatly between species. Indeed, closely related species can differ in adhesive properties, with many gain and losses of adhesion observed along the phylogenetic tree. The rapid evolution of glue adhesion can be due to an adaptation to different pupation substrates. Species pupating in exposed areas can require a stronger glue than the ones being burrowed during metamorphosis. In further analysis, we plan to compare the adhesion forces obtained with the available information

concerning ecological pupation sites.

We also observed adhesion force variations between pupae from the same species. This could be due to the position of the animal during adhesion assay and its shape affecting our measure.

5.2.2 Adhesion force is correlated with glue quantity

To understand parameters that can affect the adhesion properties, we assessed pupa morphology by measuring its area and length, and the surface of glue secreted. We found that adhesiveness increases with the quantity of glue and is not correlated with pupa morphology.

By dividing the adhesion force by the area of glue secreted, the adhesion strength is comparable between most of the species. However, a few species stand out, such as *D. hydei*, which presents a higher adhesion force per unit of area. Testing more pupae for these species will help for further conclusions as we have tested less than fifteen per species. These results suggest that the glue components responsible for glue adhesion are present in most of the species studied and that variation in the amount of glue secreted explains adhesion diversity between pupae. The common glue components among species could be Sgs3 and Sgs5bis proteins, as their genes are found in most of the species studied. The number of *Sgs* copies in the genome and their resulting proteins, if expressed, could explain the differences observed in the amount of glue between species. However, among the 27 species tested for their adhesion, only fourteen have been annotated for their *Sgs* genes in a previous study (Da Lage et al., 2019) and in Chapter 2. Annotations of the glue genes in the remaining thirteen species should be performed to draw any conclusions on a possible correlation between the presence/absence of *Sgs* genes and the glue surface area.

In this study, we considered as the glue surface, the area of glue in contact between the pupa and the substrate, which is the area of glue contributing to adhesion. As we are not yet able to trigger larva expectoration, we can not collect the glue directly from the animal. We could assess the total volume of glue produced by dissecting the salivary glands and measuring the amount of glue secreted in salivary gland lumen before expectoration.

Another parameter that could be responsible for adhesion interspecific variation would be salivary glands size. A previous study found great variation in the salivary gland cell number among several Diptera species (Babišová et al., 2023). By comparing this information with our results, we did not find any obvious relationship between the adhesive properties and the salivary gland cell number obtained for eleven of the common studied species.

5.3 Role of glue genes in pupa adhesion

In this thesis, we gained evidence that *Sgs* genes might be involved in adhesion, as it is significantly reduced when *Sgs3b* is inhibited, and slightly decreased but not significantly when a *new glue* gene, *CG33272* is silenced. These genes were inhibited using the *GAL4 UAS-RNAi* technique and the resulting lines have the same genetic background. For other lines displaying a different genetic background, we used the Gene-Switch technique which enables to turn on and off RNAi expression. However, the expression of the RNAi does not seem sufficient in these experiments. We plan to improve our protocol by using a different *GAL4* line with a higher activity and homozygous *UAS-RNAi* lines.

Pictures taken of pupae before the adhesion assays will be measured to compare the quantity of glue between RNAi induced lines. In Chapter 3, we have observed a large quantity of glue produced for some species presenting a few *Sgs* genes annotated in their genomes (*D. virilis*). Whether a few genes are producing most of the glue proteins is still unknown. In addition, we would like to use the R script analysis developed in Chapter 3 to compare adhesive properties between RNAi induced lines. So far, we have measured the adhesion force but we would like to compare the elasticity, plasticity and rigidity between these lines, as the glue composition could affect them. In the situation where no significant decrease in adhesion per unit of area is observed in *Sgs* inhibited lines, this could mean that a single *Sgs* inhibition is not sufficient to observed an effect. The adhesive function could be redundant between proteins, and inhibition of several glue genes in the same line could be necessary to observe a significant decrease.

Our collaborator Kelly Ten Hagen is studying the effect of *Sgs* genes inhibition on the glue structure within the salivary gland granules. This work is providing an interesting insight about the glue aspect before its expectoration. Together with our adhesion tests, we aim at understanding better how *Sgs* genes affect the glue structure and properties from its production in salivary glands until its expectoration on the animal body.

So far, nothing is known about the level of expression of the glue genes in other species than *D. melanogaster*. To our knowledge, transcriptomics of salivary glands has only been performed in one Drosophila species, *D. melanogaster*. Genetic tools such as UAS-GAL4 and Gene-Switch systems are only developed in *D. melanogaster* and only the eight *Sgs* genes annotated in this species can be silenced for now. However, it would be interesting to know the role of other *Sgs* copies annotated in different species in Chapter 2. In particular, we want to gain information about the *Sgs* genes expressed by species tested in the lab for their adhesion. We plan to conduct transcriptomics on the salivary glands of these species. Two approaches are possible: RNA sequencing or single-cell RNA sequencing. We are considering the later option, as it would be cost efficient to bulk salivary glands from different species together and to map the reads to a concatenated genome sequence of all the species. Single-cell RNA sequencing would provide information about the *Sgs*

and other genes expressed and the quantity of RNA in the salivary glands at the wandering stage. We plan to correlate this information with the adhesion properties and the *Sgs* genes annotations obtained from Chapter 2. Overall, future projects in the lab will focus on understanding the glue composition for the species of interest.

5.4 Drosophila glue as a promising future bioadhesive

In this thesis, I found that glue genes and glue amount are varying greatly between species, suggesting that the glue is under the action of selective forces. However, adhesion per unit of area is comparable between most species. The selection operated on the glue could be due to other functions than adhesion. Indeed, we know that the glue is involved in protecting pupae from ants predation (Borne et al., 2021b). In the glue, Sgs1, Sgs3, Sgs4 and Eig71Ee proteins are characteristic of mucins, which present antimicrobial properties in other animals (Bakshani et al., 2018). Pupation substrates such as decaying fruits, corpses or soil are rich in bacteria and fungi. We can assume that the glue presents antimicrobial properties to protect the immobile pupa, vulnerable to these pathogens during metamorphosis. We could test for Drosophila glue antimicrobial properties by placing pupae in contaminated Petri dishes. We would observe how the animal and its glue are affected and how bacteria proliferation is affected by the presence of the pupa. Drosophila glue could also be involved in facilitating adult emergence from the pupal case. We could compare the emergence of adults between pupae naturally attached with their own glue and pupae manually detached. In the same way we have tested the influence of *Sgs* genes inhibition on glue adhesion, we could also test for its effect on antimicrobial properties and adult emergence.

Pupa adhesion had never been tested on natural substrates that are found in the wild and on which pupae have been observed. We could test their adhesion properties on leaves, wood or organic substrates such as fruits. These experiments would provide information about the pupa interaction with a natural substrate, but not about the glue itself as its properties and area would be difficult to assess. Moreover, some larvae were found to chew the food medium or paper and mixing it with their glue before pupation (Babišová et al., 2023). It would be interesting to compare pupa adhesion according to the material mixed with the glue. So far, this behavior has never been tested for its resulting pupa adhesion. Assessing glue adhesion on these substrates would represent an interest for future applications in industry.

This study of Diptera glue properties aims at defining a new model of bioadhesion. *D. hydei* and *D. virilis* appear to be good candidates for future applications, as their glue

is the strongest and produced in high amounts among the 27 fly species tested. These species are easy to culture in laboratory conditions and are found in most part of the world. Moreover, *D. virilis* has only three *Sgs* genes copies identified (Da Lage et al., 2019), restricting the number of candidate glue proteins that would be responsible for adhesion.

So far, most of bioadhesives applications concern biomedical innovations, mainly as wound sealants (Bianco-Peled and Davidovich-Pinhas, 2015). However, innovative solutions are needed to replace adhesives used in other domains. As adhesives can prevent from recycling the materials attached together, developing biodegradable adhesives is of high relevance (Onusseit, 2006). Terrestrial wet bioadhesives are poorly studied but represent a great opportunity as they are mostly made of proteins and adapted to our environment. In particular, Drosophila glue is present in terrestrial and aerial environments in the wild, and sticks to a wide range of substrates. A bioadhesive inspired from Drosophila glue could be suited to food packaging adhesives that need to be biocompatible and biodegradable. Cosmetics and pharmaceutics products could also be concerned by our study for the same reasons.

Overall, this thesis takes as a model a living organism, Drosophila, to conceptualise a future invention, a bioadhesive. Through this work, we highlight that evolution is a source of inspiration and that diversity among living organisms is a wealth to conceive future solutions.

Appendix A

Public outreach

During my PhD, I had the opportunity to present my research in two videos.

The first one is directed by the Université Paris Cité and is available with the link: https://www.youtube.com/watch?v=E-WqoU2YILU

The second one is directed by the chanel Nature = Futur ! which puts light on biomimicry applications and is available with the link:

https://youtu.be/kDZfrjrv_j0

Bibliography

Abou, B., Gay, C., Laurent, B., Cardoso, O., Voigt, D., Peisker, H., and Gorb, S. (2010). Extensive collection of femtolitre pad secretion droplets in the beetle leptinotarsa decemlineata allows nanolitre microrheology. *Journal of The Royal Society Interface*, 7(53):1745–1752.

Adler, P. H. and Courtney, G. W. (2019). Ecological and societal services of aquatic diptera. *Insects*, 10(3):70.

Ahmed-Braimah, Y. H. and Sweigart, A. L. (2015). A single gene causes an interspecific difference in pigmentation in drosophila. *Genetics*, 200(1):331–342.

Ahn, J.-S., Choi, H.-K., Lee, K.-H., Nahm, J.-H., and Cho, C.-S. (2001). Novel mucoadhesive polymer prepared by template polymerization of acrylic acid in the presence of silk sericin. *Journal of applied polymer science*, 80(2):274–280.

Akai, H., Nagashima, T., and Aoyagi, S. (1993). Ultrastructure of posterior silk gland cells and liquid silk in indian tasar silkworm, antheraea mylitta drury (lepidoptera: Saturniidae). *International Journal of Insect Morphology and Embryology*, 22(5):497–506.

Akam, M. E., Roberts, D. B., Richards, G. P., and Ashburner, M. (1978). Drosophila: the genetics of two major larval proteins. *Cell*, 13(2):215–225.

Andres, A. J., Fletcher, J. C., Karim, F. D., and Thummel, C. S. (1993). Molecular analysis of the initiation of insect metamorphosis: a comparative study of drosophila ecdysteroid-regulated transcription. *Developmental biology*, 160(2):388–404.

Andrew, D. J., Henderson, K. D., and Seshaiah, P. (2000). Salivary gland development in drosophila melanogaster. *Mechanisms of development*, 92(1):5–17.

Arnott, S. and Turner, B. (2008). Post-feeding larval behaviour in the blowfly, calliphora vicina: effects on post-mortem interval estimates. *Forensic science international*, 177(2-3):162–167.

Ashburner, M. et al. (1989). *Drosophila. A laboratory handbook.* Cold spring harbor laboratory press.

Babišová, K., Mentelová, L., Geisseová, T. K., Beňová-Liszeková, D., Beňo, M., Chase, B. A., and Farkaš, R. (2023). Apocrine secretion in the salivary glands of drosophilidae and other dipterans is evolutionarily conserved. *Frontiers in Cell and Developmental Biology*, 10:2458.

Bakshani, C. R., Morales-Garcia, A. L., Althaus, M., Wilcox, M. D., Pearson, J. P., Bythell, J. C., and Burgess, J. G. (2018). Evolutionary conservation of the antimicrobial function of mucus: a first defence against infection. *npj Biofilms and Microbiomes*, 4(1):14.

Barnes, H. A. (1994). Rheology of emulsions—a review. *Colloids and surfaces A: physicochemical and engineering aspects*, 91:89–95.

Beltramí, M., Medina-Muñoz, M. C., Arce, D., and Godoy-Herrera, R. (2010). Drosophila pupation behavior in the wild. *Evolutionary ecology*, 24:347–358.

Benedict, C. V. and Waite, J. H. (1986). Composition and ultrastructure of the byssus ofMytilus edulis. *Journal of Morphology*, 189(3):261–270.

Beňová-Liszeková, D., Mentelová, L., Babišová, K., Beňo, M., Pechan, T., Chase, B. A., and Farkaš, R. (2021). An apocrine mechanism delivers a fully immunocompetent exocrine secretion. *Scientific reports*, 11(1):15915.

Benyus, J. M. (1997). Biomimicry: Innovation inspired by nature.

Berendes, H. D. (1965). Salivary gland function and chromosomal puffing patterns in drosophila hydei. *Chromosoma*, 17(1):35–77.

Betz, O. (2010). Adhesive Exocrine Glands in Insects: Morphology, Ultrastructure, and Adhesive Secretion. In von Byern, J. and Grunwald, I., editors, *Biological Adhesive Systems*, pages 111–152. Springer Vienna, Vienna.

Beutel, R. and Gorb, S. (2001). Ultrastructure of attachment specializations of hexapods (arthropoda): evolutionary patterns inferred from a revised ordinal phylogeny. *Journal of Zoological Systematics and Evolutionary Research*, 39(4):177–207.

Beutel, R. G. and Gorb, S. N. (2006). A revised interpretation of the evolution of attachment structures in hexapoda with special emphasis on mantophasmatodea. *Arthropod Systematics & Phylogeny*, 64(1):3–25.

Bianco-Peled, H. and Davidovich-Pinhas, M. (2015). *Bioadhesion and biomimetics: from nature to applications*. CRC Press.

Biomimetics, I. . (2015). Biomimetics — Terminology, concepts and methodology. Standard, International Organization for Standardization.

Borne, F., Kovalev, A., Gorb, S., and Courtier-Orgogozo, V. (2020). The glue produced by drosophila melanogaster for pupa adhesion is universal. *Journal of Experimental Biology*, 223(8):jeb220608.

Borne, F., Kulathinal, R. J., and Courtier-Orgogozo, V. (2021a). Glue genes are subjected to diverse selective forces during drosophila development. *Genome Biology and Evolution*, 13(12):evab248.

Borne, F., Prigent, S. R., Molet, M., and Courtier-Orgogozo, V. (2021b). Drosophila glue protects from predation. *Proceedings of the Royal Society B*, 288(1947):20210088.

Boudouris, B. W., Frisbie, C. D., and Hillmyer, M. A. (2008). Nanoporous poly (3-alkylthiophene) thin films generated from block copolymer templates. *Macromolecules*, 41(1):67–75.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *development*, 118(2):401–415.

Brau, F., Lanterbecq, D., Zghikh, L.-N., Bels, V., and Damman, P. (2016). Dynamics of prey prehension by chameleons through viscous adhesion. *Nature Physics*, 12(10):931–935.

Burgess, J., Jauregui, M., Tan, J., Rollins, J., Lallet, S., Leventis, P. A., Boulianne, G. L., Chang, H. C., Le Borgne, R., Krämer, H., et al. (2011). Ap-1 and clathrin are essential for secretory granule biogenesis in drosophila. *Molecular biology of the cell*, 22(12):2094–2105.

Burke, S. A., Ritter-Jones, M., Lee, B. P., and Messersmith, P. B. (2007). Thermal gelation and tissue adhesion of biomimetic hydrogels. *Biomedical materials*, 2(4):203.

Büscher, T. H. and Gorb, S. N. (2021). Physical constraints lead to parallel evolution of micro-and nanostructures of animal adhesive pads: a review. *Beilstein Journal of Nanotechnology*, 12(1):725–743.

Butt, H.-J., Cappella, B., and Kappl, M. (2005). Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surface science reports*, 59(1-6):1–152.

Büscher, T. H. and Gorb, S. N. (2021). Physical constraints lead to parallel evolution of micro- and nanostructures of animal adhesive pads: a review. *Beilstein Journal of Nanotechnology*, 12:725–743.

Büscher, T. H., Lohar, R., Kaul, M.-C., and Gorb, S. N. (2020). Multifunctional Adhesives on the Eggs of the Leaf Insect Phyllium philippinicum (Phasmatodea: Phylliidae): Solvent Influence and Biomimetic Implications. *Biomimetics*, 5(4):66.

Campos, J. and Hartenstein, V. (1985). The embryonic development of drosophila melanogaster. *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormone, Springer-Verlag, Berlin, Germany*, pages 218–225.

Cao, T.-T. and Zhang, Y.-Q. (2016). Processing and characterization of silk sericin from bombyx mori and its application in biomaterials and biomedicines. *Materials Science and Engineering: C*, 61:940–952.

Carson, H. L. (1967). The association between drosophila carcinophila wheeler and its host, the land crab gecarcinus ruricola (l.). *American Midland Naturalist*, pages 324–343.

Carson, H. L., Hardy, D. E., Spieth, H. T., and Stone, W. S. (1970). *The evolutionary biology of the Hawaiian Drosophilidae*. Springer.

Carson, H. L. and Stalker, H. D. (1951). Natural breeding sites for some wild species of drosophila in the eastern united states. *Ecology*, 32(2):317–330.

Chirila, T. V., Suzuki, S., and McKirdy, N. C. (2016). Further development of silk sericin as a biomaterial: comparative investigation of the procedures for its isolation from bombyx mori silk cocoons. *Progress in biomaterials*, 5(2):135–145.

Clemente, C. J., Bullock, J. M., Beale, A., and Federle, W. (2010). Evidence for self-cleaning in fluid-based smooth and hairy adhesive systems of insects. *Journal of Experimental Biology*, 213(4):635–642.

Coleman, J. M., Benowitz, K. M., Jost, A. G., and Matzkin, L. M. (2018). Behavioral evolution accompanying host shifts in cactophilic drosophila larvae. *Ecology and Evolution*, 8(14):6921–6931.

Costa, B. E., Rohde, C., and Valente, V. L. d. S. (2003). Temperature, urbanization and body color polymorphism in south brazilian populations of drosophila kikkawai (diptera, drosophilidae). *Iheringia. Série Zoologia*, 93:381–393.

Da Lage, J.-L., Thomas, G. W., Bonneau, M., and Courtier-Orgogozo, V. (2019). Evolution of salivary glue genes in drosophila species. *BMC evolutionary biology*, 19(1):1–22.

Davey, P. A., Power, A. M., Santos, R., Bertemes, P., Ladurner, P., Palmowski, P., Clarke, J., Flammang, P., Lengerer, B., Hennebert, E., et al. (2021). Omics-based molecular analyses of adhesion by aquatic invertebrates. *Biological Reviews*, 96(3):1051–1075.

D'Avino, P. P., Crispi, S., Polito, L. C., and Furia, M. (1995). The role of the br-c locus on the expression of genes located at the ecdysone-regulated 3c puff of drosophila melanogaster. *Mechanisms of development*, 49(3):161–171.

Delaney, K. T. and Fredrickson, G. H. (2017). Theory of polyelectrolyte complexation—complex coacervates are self-coacervates. *The Journal of chemical physics*, 146(22):224902.

Delroisse, J., Kang, V., Gouveneaux, A., Santos, R., and Flammang, P. (2023). Convergent evolution of attachment mechanisms in aquatic animals. In *Convergent Evolution: Animal Form and Function*, pages 523–557. Springer.

DeMartini, D. G., Errico, J. M., Sjoestroem, S., Fenster, A., and Waite, J. H. (2017). A cohort of new adhesive proteins identified from transcriptomic analysis of mussel foot glands. *Journal of the Royal Society Interface*, 14(131):20170151.

Dirks, J.-H. (2014). Physical principles of fluid-mediated insect attachment-shouldn't insects slip? *Beilstein journal of nanotechnology*, 5(1):1160–1166.

Dirks, J.-H., Clemente, C. J., and Federle, W. (2010). Insect tricks: two-phasic foot pad secretion prevents slipping. *Journal of the Royal Society Interface*, 7(45):587–593.

Dirks, J.-H. and Federle, W. (2011). Fluid-based adhesion in insects–principles and challenges. *Soft Matter*, 7(23):11047–11053.

Ditsche, P. and Summers, A. P. (2014). Aquatic versus terrestrial attachment: Water makes a difference. *Beilstein Journal of Nanotechnology*, 5:2424–2439.

Dorow, W., Maschwitz, U., and Rapp, S. (1990). The natural history of polyrhachis (myrmhopla) muelleri forel 1893 (formicidae formicinae), a weaver ant with mimetic larvae and an unusual nesting behaviour. *Tropical Zoology*, 3(2):181–190.

dos Santos-Pinto, J. R. A., Esteves, F. G., Tormena, C. F., Perez-Riverol, A., Lasa, A. M., Bueno, O. C., and Palma, M. S. (2022). Proteomic characterization of the fibroin-based silk fibers produced by weaver ant camponotus textor. *Journal of Proteomics*, 261:104579.

Drechsler, P. and Federle, W. (2006). Biomechanics of smooth adhesive pads in insects: influence of tarsal secretion on attachment performance. *Journal of Comparative Physiology A*, 192:1213–1222.

Du, D., Chen, X., Shi, C., Zhang, Z., Shi, D., Kaneko, D., Kaneko, T., and Hua, Z. (2021). Mussel-inspired epoxy bioadhesive with enhanced interfacial interactions for wound repair. *Acta Biomaterialia*, 136:223–232.

Duan, J., Zhao, Y., Li, H., Habernig, L., Gordon, M. D., Miao, X., Engström, Y., and Büttner, S. (2020). Bab2 functions as an ecdysone-responsive transcriptional repressor during drosophila development. *Cell reports*, 32(4):107972.

Eisner, T. and Aneshansley, D. J. (2000). Defense by foot adhesion in a beetle (hemisphaerota cyanea). *Proceedings of the National Academy of Sciences*, 97(12):6568–6573.

Epstein, L. and Nicholson, R. (2016). Adhesion and Adhesives of Fungi and Oomycetes. In Smith, A. M., editor, *Biological Adhesives*, pages 25–55. Springer International Publishing, Cham.

Erezyilmaz, D. F. and Stern, D. L. (2013). Pupariation site preference within and between drosophila sibling species. *Evolution*, 67(9):2714–2727.

Farkaš, R. (2016). The complex secretions of the salivary glands of drosophila melanogaster, a model system. *Extracellular composite matrices in Arthropods*, pages 557–600.

Farkaš, R., Ďatková, Z., Mentelova, L., Löw, P., Beňová-Liszeková, D., Beňo, M., Sass, M., Řehulka, P., Řehulková, H., Raška, O., et al. (2014). Apocrine secretion in drosophila salivary glands: subcellular origin, dynamics, and identification of secretory proteins. *PLoS One*, 9(4):e94383.

Fayemi, P.-E., Maranzana, N., Aoussat, A., Bersano, G., et al. (2014). Bio-inspired design characterisation and its links with problem solving tools. In *DS 77: Proceedings of the DESIGN 2014 13th International Design Conference*, pages 173–182.

Federle, W., Riehle, M., Curtis, A. S., and Full, R. J. (2002). An integrative study of insect adhesion: mechanics and wet adhesion of pretarsal pads in ants. *Integrative and Comparative Biology*, 42(6):1100–1106.

Flammang, P. (2020). Adhesion in echinoderms. In Jangoux, M. and Lawrence, J. M., editors, *Echinoderm Studies*, pages 1–60. CRC Press, 1 edition.

Flammang, P., Demeuldre, M., Hennebert, E., and Santos, R. (2016). Adhesive Secretions in Echinoderms: A Review. In Smith, A. M., editor, *Biological Adhesives*, pages 193–222. Springer International Publishing, Cham.

Flammang, P., Demeulenaere, S., and Jangoux, M. (1994). The Role of Podial Secretions in Adhesion in Two Species of Sea Stars (Echinodermata). *The Biological Bulletin*, 187(1):35–47.

Flammang, P., Michel, A., Van Cauwenberge, A., Alexandre, H., and Jangoux, M. (1998). A study of the temporary adhesion of the podia in the sea star asterias rubens (echinodermata, asteroidea) through their footprints. *Journal of Experimental Biology*, 201(16):2383–2395.

Flammang, P. and Walker, G. (1997). Measurement of the Adhesion of the Podia in the Asteroid *Asterias Rubens* (Echinodermata). *Journal of the Marine Biological Association of the United Kingdom*, 77(4):1251–1254.

Flatt, T. and Heyland, A. (2011). *Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs*. Oxford university press.

Fraenkel, G. and Brookes, V. J. (1953). The process by which the puparia of many species of flies become fixed to a substrate. *The Biological Bulletin*, 105(3):442–449.

Furia, M., D'Avino, P. P., Crispi, S., Artiaco, D., and Polito, L. C. (1993). Dense cluster of genes is located at the ecdysone-regulated 3c puff of drosphila melanogaster. *Journal of molecular biology*, 231(2):531–538.

Garfinkel, M. D., Pruitt, R. E., and Meyerowitz, E. M. (1983). Dna sequences, gene regulation and modular protein evolution in the drosophila 68c glue gene cluster. *Journal of molecular biology*, 168(4):765–789.

Godoy-Herrera, R., Cifuentes, L., de Arcaya, M. F. D., Fernández, M., Fuentes, M., Reyes, I., and Valderrama, C. (1989). The behaviour of drosophila melanogaster larvae during pupation. *Animal behaviour*, 37:820–829.

Godoy-Herrera, R. and Silva-Cuadra, J. L. (1998). The behavior of sympatric chilean populations of drosophila larvae during pupation. *Genetics and molecular biology*, 21:31–39.

Gorb, E. V. and Gorb, S. N. (2017). Anti-adhesive effects of plant wax coverage on insect attachment. *Journal of Experimental Botany*, 68(19):5323–5337.

Gorb, E. V., Lemke, W., and Gorb, S. N. (2019). Porous substrate affects a subsequent attachment ability of the beetle harmonia axyridis (coleoptera, coccinellidae). *Journal of the Royal Society Interface*, 16(150):20180696.

Gorb, S. N. (1998). The design of the fly adhesive pad: distal tenent setae are adapted to the delivery of an adhesive secretion. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1398):747–752.

Gorb, S. N. (2007). Smooth attachment devices in insects: functional morphology and biomechanics. *Advances in insect physiology*, 34:81–115.

Gorb, S. N. (2008). Biological attachment devices: exploring nature's diversity for biomimetics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1870):1557–1574.

Gorb, S. N., Beutel, R. G., Gorb, E. V., Jiao, Y., Kastner, V., Niederegger, S., Popov, V. L., Scherge, M., Schwarz, U., and Vötsch, W. (2002). Structural design and biomechanics of friction-based releasable attachment devices in insects. *Integrative and Comparative Biology*, 42(6):1127–1139.

Greenberg, B. (1990). Behavior of postfeeding larvae of some calliphoridae and a muscid (diptera). *Annals of the Entomological Society of America*, 83(6):1210–1214.

Guerette, P. A., Hoon, S., Seow, Y., Raida, M., Masic, A., Wong, F. T., Ho, V. H., Kong, K. W., Demirel, M. C., Pena-Francesch, A., et al. (2013). Accelerating the design of biomimetic materials by integrating rna-seq with proteomics and materials science. *Nature biotechnology*, 31(10):908–915.

Haller, C., Buerzle, W., Kivelio, A., Perrini, M., Brubaker, C., Gubeli, R., Mallik, A., Weber, W., Messersmith, P., Mazza, E., et al. (2012). Mussel-mimetic tissue adhesive for fetal membrane repair: an ex vivo evaluation. *Acta biomaterialia*, 8(12):4365–4370.

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., and Ira, G. (2009). Mechanisms of change in gene copy number. *Nature Reviews Genetics*, 10(8):551–564.

Heinrich, L. A. (2019). Future opportunities for bio-based adhesives–advantages beyond renewability. *Green chemistry*, 21(8):1866–1888.

Hennebert, E., Haesaerts, D., Dubois, P., and Flammang, P. (2010). Evaluation of the different forces brought into play during tube foot activities in sea stars. *Journal of Experimental Biology*, 213(7):1162–1174.

Hennebert, E., Leroy, B., Wattiez, R., and Ladurner, P. (2015). An integrated transcriptomic and proteomic analysis of sea star epidermal secretions identifies proteins involved in defense and adhesion. *Journal of proteomics*, 128:83–91.

Hennebert, E., Santos, R., and Flammang, P. (2012). Echinoderms don't suck: evidence against the involvement of suction in tube foot attachment*. *Zoosymposia*, 7(1):25–32.

Hennebert, E., Wattiez, R., Demeuldre, M., Ladurner, P., Hwang, D. S., Waite, J. H., and Flammang, P. (2014). Sea star tenacity mediated by a protein that fragments, then aggregates. *Proceedings of the National Academy of Sciences*, 111(17):6317–6322.

Hennebert, E., Wattiez, R., and Flammang, P. (2011). Characterisation of the Carbohydrate Fraction of the Temporary Adhesive Secreted by the Tube Feet of the Sea Star Asterias rubens. *Marine Biotechnology*, 13(3):484–495.

Heredia, F., Volonté, Y., Pereirinha, J., Fernandez-Acosta, M., Casimiro, A. P., Belém, C. G., Viegas, F., Tanaka, K., Menezes, J., Arana, M., et al. (2021). The steroid-hormone ecdysone coordinates parallel pupariation neuromotor and morphogenetic subprograms via epidermis-to-neuron dilp8-lgr3 signal induction. *Nature communications*, 12(1):3328.

Hermans, C. (1983). The duo-gland adhesive system. *Oceanography and Marine Biology*, 21:283–339.

Hofmann, A., Garfinkel, M. D., and Meyerowitz, E. M. (1991). cis-acting sequences required for expression of the divergently transcribed drosophila melanogaster sgs-7 and sgs-8 glue protein genes. *Molecular and cellular biology*, 11(6):2971–2979.

Hölldobler, B. and Wilson, E. O. (1977). Weaver ants: social establishment and maintenance of territory. *Science*, 195(4281):900–902.

Ideo, S., Watada, M., Mitsui, H., and Kimura, M. T. (2008). Host range of asobara japonica (hymenoptera: Braconidae), a larval parasitoid of drosophilid flies. *Entomological Science*, 11(1):1–6.

Israelachvili, J. N. (2011). *Intermolecular and surface forces*. Academic press.

Jangoux, M. and Lawrence, J. M., editors (1996). *Echinoderm studies. Volume 5*. A.A. Balkema, Rotterdam. OCLC: 1182800126.

Ji, S., Samara, N. L., Revoredo, L., Zhang, L., Tran, D. T., Muirhead, K., Tabak, L. A., and Ten Hagen, K. G. (2018). A molecular switch orchestrates enzyme specificity and secretory granule morphology. *Nature Communications*, 9(1):3508.

Johnson, P. H. and Russell, R. C. (2019). Effects of attachment substrate, larval diet, and temperature on development and survival of immature coquillettidia linealis (skuse) and coquillettidia xanthogaster (edwards). *Journal of Vector Ecology*, 44(1):138–148.

Josso, C., Moiroux, J., Vernon, P., Van Baaren, J., and van Alphen, J. J. (2011). Temperature and parasitism by asobara tabida (hymenoptera: Braconidae) influence larval pupation behaviour in two drosophila species. *Naturwissenschaften*, 98:705–709.

Karim, F. D., Guild, G. M., and Thummel, C. S. (1993). The drosophila broad-complex plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis. *Development*, 118(3):977–988.

Khanlari, S. and Dubé, M. A. (2013). Bioadhesives: A Review: Bioadhesives: A Review. *Macromolecular Reaction Engineering*, 7(11):573–587.

Kiel, E. and Röder, T. (2002). Gelelectrophoretic studies on labial gland secretions of immature blackflies (simuliidae, diptera). *Limnologica*, 32(3):201–205.

Kirillova, A., Kelly, C., von Windheim, N., and Gall, K. (2018). Bioinspired mineral–organic bioresorbable bone adhesive. *Advanced Healthcare Materials*, 7(17):1800467.

Kivelio, A., DeKoninck, P., Perrini, M., Brubaker, C. E., Messersmith, P. B., Mazza, E., Deprest, J., Zimmermann, R., Ehrbar, M., and Ochsenbein-Koelble, N. (2013). Mussel mimetic tissue adhesive for fetal membrane repair: initial in vivo investigation in rabbits. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 171(2):240–245.

Kord Forooshani, P. and Lee, B. P. (2017). Recent approaches in designing bioadhesive materials inspired by mussel adhesive protein. *Journal of Polymer Science Part A: Polymer Chemistry*, 55(1):9–33.

Korge, G. (1975). Chromosome puff activity and protein synthesis in larval salivary glands of drosophila melanogaster. *Proceedings of the National Academy of Sciences*, 72(11):4550–4554.

Kress, H. (1982). Biochemical and ontogenetic aspects of glycoprotein synthesis in drosophila virilis salivary glands. *Developmental biology*, 93(1):231–239.

Kumar, S., Suleski, M., Craig, J. M., Kasprowicz, A. E., Sanderford, M., Li, M., Stecher, G., and Hedges, S. B. (2022). Timetree 5: An expanded resource for species divergence times. *Molecular Biology and Evolution*, 39(8):msac174.

Lang, M., Polihronakis Richmond, M., Acurio, A., Markow, T., and Orgogozo, V. (2014). Radiation of the d rosophila nannoptera species group in m exico. *Journal of evolutionary biology*, 27(3):575–584.

Lee, B. P., Messersmith, P. B., Israelachvili, J. N., and Waite, J. H. (2011). Mussel-inspired adhesives and coatings. *Annual review of materials research*, 41:99–132.

Lee, T. and Luo, L. (2001). Mosaic analysis with a repressible cell marker (marcm) for drosophila neural development. *Trends in neurosciences*, 24(5):251–254.

Lee, Y., Xu, C., Sebastin, M., Lee, A., Holwell, N., Xu, C., Miranda Nieves, D., Mu, L., Langer, R. S., Lin, C., et al. (2015). Bioinspired nanoparticulate medical glues for minimally invasive tissue repair. *Advanced healthcare materials*, 4(16):2587–2596.

Lefèvre, B. M., Mienanzambi, S., and Lang, M. (2022). Developmental timing of drosophila pachea pupae is robust to temperature changes. *Journal of Thermal Biology*, 106:103232.

Li, T.-R. and White, K. P. (2003). Tissue-specific gene expression and ecdysone-regulated genomic networks in drosophila. *Developmental cell*, 5(1):59–72.

Li, Z., Qian, W., Song, W., Zhao, T., Yang, Y., Wang, W., Wei, L., Zhao, D., Li, Y., Perrimon, N., et al. (2022). A salivary gland-secreted peptide regulates insect systemic growth. *Cell reports*, 38(8):110397.

Manning, M. and Markow, T. A. (1981). Light-dependent pupation site preferences in drosophila. ii. drosophila melanogaster and drosophila simulans. *Behavior genetics*, 11(6):557–563.

Marchiori, C. and Silva, C. (2003). First occurrence of parasitoid spalangia endius (walker)(hymenoptera: Pteromalidae) in pupae of zaprionus indianus gupta (diptera: Drosophilidae) in brazil. *Brazilian Journal of Biology*, 63:361–362.

Mayhew, P. J. (2007). Why are there so many insect species? perspectives from fossils and phylogenies. *Biological Reviews*, 82(3):425–454.

McCombie, W. R., McPherson, J. D., and Mardis, E. R. (2019). Next-generation sequencing technologies. *Cold Spring Harbor perspectives in medicine*, 9(11):a036798.

Melrose, J. (2022). High performance marine and terrestrial bioadhesives and the biomedical applications they have inspired. *Molecules*, 27(24):8982.

Merritt, R. W., Courtney, G. W., and Keiper, J. B. (2009). Diptera:(flies, mosquitoes, midges, gnats). In *Encyclopedia of insects*, pages 284–297. Elsevier.

Messersmith, P. B., Brubaker, C., and Zisch, A. H. (2013). Sealants for fetal membrane repair. US Patent 8,409,602.

Monier, M. and Courtier-Orgogozo, V. (2022). Drosophila glue: a promising model for bioadhesion. *Insects*, 13(8):734.

Mooi, R. (1986). Non-respiratory podia of clypeasteroids (Echinodermata, Echinoides): I. Functional anatomy. *Zoomorphology*, 106(1):21–30.

Mulla, M., Chaudhury, M., et al. (1968). Effects of surface tension on pupae of culex pipiens quinquefasciatus say and aedes aegypti (l.). *Mosquito News*, 28(2):187–91.

Nicholson, L., Singh, G. K., Osterwalder, T., Roman, G. W., Davis, R. L., and Keshishian, H. (2008). Spatial and temporal control of gene expression in drosophila using the inducible geneswitch gal4 system. i. screen for larval nervous system drivers. *Genetics*, 178(1):215–234.

Norton, M. R., Kay, G. W., Brown, M. C., and Cochran, D. L. (2020). Bone glue-the final frontier for fracture repair and implantable device stabilization. *International Journal of Adhesion and Adhesives*, 102:102647.

Onusseit, H. (2006). The influence of adhesives on recycling. *Resources, Conservation and Recycling*, 46(2):168–181.

Papanicolaou, A., Woo, A., Brei, B., Ma, D., Masedunskas, A., Gray, E., Xiao, G. G., Cho, S., and Brockhouse, C. (2013). Novel aquatic silk genes from simulium (psilozia) vittatum (zett) diptera: Simuliidae. *Insect biochemistry and molecular biology*, 43(12):1181–1188.

Peisker, H., Heepe, L., Kovalev, A. E., and Gorb, S. N. (2014). Comparative study of the fluid viscosity in tarsal hairy attachment systems of flies and beetles. *Journal of the Royal Society Interface*, 11(99):20140752.

Peppas, N. A. and Buri, P. A. (1985). Surface, interfacial and molecular aspects of polymer bioadhesion on soft tissues. *Journal of Controlled Release*, 2:257–275.

Perkins, L. A., Holderbaum, L., Tao, R., Hu, Y., Sopko, R., McCall, K., Yang-Zhou, D., Flockhart, I., Binari, R., Shim, H.-S., et al. (2015). The transgenic rnai project at harvard medical school: resources and validation. *Genetics*, 201(3):843–852.

Perrini, M., Barrett, D., Ochsenbein-Koelble, N., Zimmermann, R., Messersmith, P., and Ehrbar, M. (2016). A comparative investigation of mussel-mimetic sealants for fetal membrane repair. *Journal of the mechanical behavior of biomedical materials*, 58:57–64.

Putthanarat, S., Stribeck, N., Fossey, S., Eby, R., and Adams, W. (2000). Investigation of the nanofibrils of silk fibers. *Polymer*, 41(21):7735–7747.

Ramesh, S. and Kalisch, W.-E. (1988). Glue proteins in drosophila nasuta. *Biochemical genetics*, 26(7-8):527–541.

Rawlings, A. E., Bramble, J. P., and Staniland, S. S. (2012). Innovation through imitation: biomimetic, bioinspired and biokleptic research. *Soft Matter*, 8(25):6675–6679.

Reddy, N., Xu, H., and Yang, Y. (2011). Unique natural-protein hollow-nanofiber membranes produced by weaver ants for medical applications. *Biotechnology and Bioengineering*, 108(7):1726–1733.

Reibe, S. and Madea, B. (2010). Use of megaselia scalaris (diptera: Phoridae) for postmortem interval estimation indoors. *Parasitology Research*, 106:637–640.

Reigada, C., Giao, J. Z., Galindo, L. A., and Godoy, W. A. C. (2011). Survival of submerged blowfly species and their parasitoids: implications for postmortem submersion interval. *Forensic Science International*, 212(1-3):126–129.

Reynolds, H. M., Zhang, L., Tran, D. T., and Ten Hagen, K. G. (2019). Tango1 coordinates the formation of endoplasmic reticulum/golgi docking sites to mediate secretory granule formation. *Journal of Biological Chemistry*, 294(51):19498–19510.

Richards, A. G. and Richards, P. A. (1979). The cuticular protuberances of insects. *International Journal of Insect Morphology and Embryology*, 8(3-4):143–157.

Rizki, M. and Davis Jr, C. G. (1953). Light as an ecological determinant of interspecific competition between drosophila willistoni and drosophila melanogaster. *The American Naturalist*, 87(837):389–392.

Rohlfs, M. and Hoffmeister, T. S. (2004). Spatial aggregation across ephemeral resource patches in insect communities: an adaptive response to natural enemies? *Oecologia*, 140:654–661.

Roman, G., Endo, K., Zong, L., and Davis, R. L. (2001). P {Switch}, a system for spatial and temporal control of gene expression in drosophila melanogaster. *Proceedings of the national academy of sciences*, 98(22):12602–12607.

Rota-Stabelli, O., Blaxter, M., and Anfora, G. (2013). Drosophila suzukii. *Current Biology*, 23(1):R8–R9.

Rzepecki, L. M., Hansen, K. M., and Waite, J. H. (1992). Characterization of a Cystine-Rich Polyphenolic Protein Family from the Blue Mussel *Mytilus edulis* L. *The Biological Bulletin*, 183(1):123–137.

Sameoto, D. and Miller, R. S. (1968). Selection of pupation site by drosophila melanogaster and d. simulans. *Ecology*, 49(1):177–180.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). Dna sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences*, 74(12):5463–5467.

Santos, R., Gorb, S., Jamar, V., and Flammang, P. (2005). Adhesion of echinoderm tube feet to rough surfaces. *Journal of Experimental Biology*, 208(13):2555–2567.

Schnebel, E. M. and Grossfield, J. (1992). Temperature effects on pupation-height response in four drosophila species group triads. *Journal of insect physiology*, 38(10):727–732.

Sehnal, F. and Sutherland, T. (2008). Silks produced by insect labial glands. *Prion*, 2(4):145–153.

Shirk, P. D., Roberts, P. A., and Harn, C. H. (1988). Synthesis and secretion of salivary gland proteins in drosophila gibberosa during larval and prepupal development. *Roux's archives of developmental biology*, 197:66–74.

Singh, D. and Greenberg, B. (1994). Survival after submergence in the pupae of five species of blow flies (diptera: Calliphoridae). *Journal of Medical Entomology*, 31(5):757–759.

Siri, S. and Maensiri, S. (2010). Alternative biomaterials: Natural, non-woven, fibroin-based silk nanofibers of weaver ants (Oecophylla smaragdina). *International Journal of Biological Macromolecules*, 46(5):529–534.

Slatko, B. E., Gardner, A. F., and Ausubel, F. M. (2018). Overview of next-generation sequencing technologies. *Current protocols in molecular biology*, 122(1):e59.

Sokolowski, M. B. (1985). Genetics and ecology of drosophila melanogaster larval foraging and pupation behaviour. *Journal of insect physiology*, 31(11):857–864.

Sokolowski, M. B., Bauer, S. J., Wai-Ping, V., Rodriguez, L., Wong, J. L., and Kent, C. (1986). Ecological genetics and behaviour of drosophila melanogaster larvae in nature. *Animal Behaviour*, 34(2):403–408.

Sokolowski, M. B., Kent, C., and Wong, J. (1984). Drosophila larval foraging behaviour: developmental stages. *Animal behaviour*, 32(3):645–651.

Stevens, M. J., Steren, R. E., Hlady, V., and Stewart, R. J. (2007). Multiscale Structure of the Underwater Adhesive of *Phragmatopoma Californica* : a Nanostructured Latex with a Steep Microporosity Gradient. *Langmuir*, 23(9):5045–5049.

Stewart, R. J., Wang, C. S., and Shao, H. (2011). Complex coacervates as a foundation for synthetic underwater adhesives. *Advances in Colloid and Interface Science*, 167(1-2):85–93.

Stewart, R. J., Weaver, J. C., Morse, D. E., and Waite, J. H. (2004). The tube cement of *Phragmatopoma californica:* a solid foam. *Journal of Experimental Biology*, 207(26):4727–4734.

Stuart, A. and Hunter, F. (1995). A re-description of the cocoon-spinning behaviour of simulium vittatum (diptera simuliidae). *Ethology Ecology & Evolution*, 7(4):363–377.

Su, Y., Wang, F., Wu, S., Fan, Y., Bai, W., Wang, S., Sun, H., Zhu, Z., Liang, W., and Li, A. (2021). Template-assisted preparation of conjugated microporous polymers membranes for selective separation. *Separation and Purification Technology*, 259:118203.

Sun, J., Bauman, L., Yu, L., and Zhao, B. (2023). Gecko-and-inchworm-inspired untethered soft robot for climbing on walls and ceilings. *Cell Reports Physical Science*, page 101241.

Sutherland, T. D., Young, J. H., Weisman, S., Hayashi, C. Y., and Merritt, D. J. (2010). Insect Silk: One Name, Many Materials. *Annual Review of Entomology*, 55(1):171–188.

Suvorov, A., Kim, B. Y., Wang, J., Armstrong, E. E., Peede, D., D'agostino, E. R., Price, D. K., Waddell, P. J., Lang, M., Courtier-Orgogozo, V., et al. (2022). Widespread introgression across a phylogeny of 155 drosophila genomes. *Current Biology*, 32(1):111–123.

Syed, Z. A., Zhang, L., Tran, D. T., Bleck, C. K., and Ten Hagen, K. G. (2022). Regulated restructuring of mucins during secretory granule maturation in vivo. *Proceedings of the National Academy of Sciences*, 119(43):e2209750119.

Teramoto, H., Kameda, T., and Tamada, Y. (2008). Preparation of gel film from bombyx mori silk sericin and its characterization as a wound dressing. *Bioscience, biotechnology, and biochemistry*, 72(12):3189–3196.

Teramoto, H. and Miyazawa, M. (2005). Molecular orientation behavior of silk sericin film as revealed by atr infrared spectroscopy. *Biomacromolecules*, 6(4):2049–2057.

Thomas, L. A. and Hermans, C. O. (1985). ADHESIVE INTERACTIONS BETWEEN THE TUBE FEET OF A STARFISH, *LEPTASTERIAS HEXACTIS* , AND SUBSTRATA. *The Biological Bulletin*, 169(3):675–688.

Tran, M. L. T. Q., De Muijlder, T., Pittenger, B., Flammang, P., Hennebert, E., and Leclère, P. (2021). On the nanomechanical and viscoelastic properties of coatings made of recombinant sea star adhesive proteins. *Bioadhesion*.

Tyler, S. (1988). The role of function in determination of homology and convergence-examples from invertebrates adhesive organs. *Fortsch Zool*, 36:331–347.

Ude, A., Eshkoor, R., Zulkifili, R., Ariffin, A., Dzuraidah, A., and Azhari, C. (2014). Bombyx mori silk fibre and its composite: a review of contemporary developments. *Materials & Design*, 57:298–305.

van den Boogaart, L. M., Langowski, J. K., and Amador, G. J. (2022). Studying stickiness: Methods, trade-offs, and perspectives in measuring reversible biological adhesion and friction. *Biomimetics*, 7(3):134.

Vandal, N. B., Siddalingamurthy, G. S., and Shivanna, N. (2008). Larval pupation site preference on fruit in different species of drosophila. *Entomological Research*, 38(3):188–194.

Vanin, S. (2016). Advances in forensic entomology in missing persons investigations. *Handbook of missing persons*, pages 309–317.

von Byern, J. and Grunwald, I., editors (2010). *Biological Adhesive Systems*. Springer Vienna, Vienna.

von Kalm, L., Crossgrove, K., Von Seggern, D., Guild, G. M., and Beckendorf, S. K. (1994). The broad-complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of drosophila metamorphosis. *The EMBO journal*, 13(15):3505–3516.

Waite, J. H. (2017). Mussel adhesion - essential footwork. *The Journal of Experimental Biology*, 220(Pt 4):517–530.

Waite, J. H., Jensen, R. A., and Morse, D. E. (1992). Cement precursor proteins of the reef-building polychaete Phragmatopoma californica (Fewkes). *Biochemistry*, 31(25):5733–5738.

Wanasingha, N., Dutta, N. K., and Choudhury, N. R. (2021). Emerging bioadhesives: From traditional bioactive and bioinert to a new biomimetic protein-based approach. *Advances in Colloid and Interface Science*, 296:102521.

Wei, W., Tan, Y., Martinez Rodriguez, N. R., Yu, J., Israelachvili, J. N., and Waite, J. H. (2014). A mussel-derived one component adhesive coacervate. *Acta Biomaterialia*, 10(4):1663–1670.

Wright, L. G., Chen, T., Thummel, C. S., and Guild, G. M. (1996). Molecular characterization of the 71e late puff indrosophila melanogasterreveals a family of novel genes. *Journal of molecular biology*, 255(3):387–400.

Yamamoto-Hino, M. and Goto, S. (2013). In vivo rnai-based screens: studies in model organisms. *Genes*, 4(4):646–665.

Yu, J., Wei, W., Danner, E., Ashley, R. K., Israelachvili, J. N., and Waite, J. H. (2011). Mussel protein adhesion depends on interprotein thiol-mediated redox modulation. *Nature Chemical Biology*, 7(9):588–590.

Zhang, Y.-Q. (2002). Applications of natural silk protein sericin in biomaterials. *Biotechnology advances*, 20(2):91–100.

Zhao, H., Robertson, N. B., Jewhurst, S. A., and Waite, J. H. (2006). Probing the Adhesive Footprints of Mytilus californianus Byssus. *Journal of Biological Chemistry*, 281(16):11090–11096.

Zhao, H., Sun, C., Stewart, R. J., and Waite, J. H. (2005). Cement Proteins of the Tube-building Polychaete Phragmatopoma californica. *Journal of Biological Chemistry*, 280(52):42938–42944.