Oomics-based molecular analyses of adhesion by aquatic invertebrates

Peter A. Davey, Anne Marie Power, Romana Santos, Philip Bertemes, Peter Ladurner, Pawel Palmowski, Jessica Clarke, Patrick Flammang, Birgit Lengerer, Elise Hennebert, Ute Rothbächer, Robert Pjeta, Julia Wunderer, Michal Zurovec and Nick Aldred

School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, U.K.
Ryan Institute, School of Natural Sciences, National University of Ireland Galway, Room 226, Galway, H91 TK33, Ireland
Departamento de Biologia Animal, Faculdade de Ciências, Centro de Ciências do Mar e do Ambiente (MARE), Universidade de Lisboa, Lisbon, 1749-016, Portugal
Institute of Zoology and Center of Molecular Biosciences Innsbruck, University of Innsbruck, Technikerstrasse 25, Innsbruck, 6020, Austria
Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, Mons, 7000, Belgium
Laboratory of Cell Biology, Research Institute for Biosciences, University of Mons, Place du Parc 23, Mons, 7000, Belgium
Biology Centre of the Czech Academy of Sciences and Faculty of Sciences, University of South Bohemia, České Budějovice, 370 05, Czech Republic
School of Life Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, U.K.

ABSTRACT

Many aquatic invertebrates are associated with surfaces, using adhesives to attach to the substratum for locomotion, prey capture, reproduction, building or defence. Their intriguing and sophisticated biological glues have been the focus of study for decades. In all but a couple of specific taxa, however, the precise mechanisms by which the bioadhesives stick to surfaces underwater and (in many cases) harden have proved to be elusive. Since the bulk components are known to be based on proteins in most organisms, the opportunities provided by advancing ‘omics technologies have revolutionised bioadhesion research. Time-consuming isolation and analysis of single molecules has been either replaced or augmented by the generation of massive data sets that describe the organism’s translated genes and proteins. While these new approaches have provided resources and opportunities that have enabled physiological insights and taxonomic comparisons that were not previously possible, they do not provide the complete picture and continued multi-disciplinarity is essential. This review covers the various ways in which ‘omics have contributed to our understanding of adhesion by aquatic invertebrates, with new data to illustrate key points. The associated challenges are highlighted and priorities are suggested for future research.

Key words: proteomics, genomics, transcriptomics, adhesion, bioadhesion, invertebrate, marine, aquatic

CONTENTS

I. Introduction ................................................................. 1052
II. Historic trends in bioadhesion research .......................... 1053
III. Strengths: improved understanding of established bioadhesion systems ........................................ 1054
   (1) Barnacles ............................................................ 1054
   (2) Mussels ............................................................. 1056
   (3) Echinoderms ....................................................... 1057
   (4) Caddisflies and other arthropods .......................... 1058

* Address for correspondence (E-mail: nick.aldred@essex.ac.uk)

Biological Reviews 96 (2021) 1051–1075 © 2021 The Authors. Biological Reviews published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
I. INTRODUCTION

Twenty years ago, the first version of the human genome was sequenced and published (Venter et al., 2001). Since then, sequencing technologies and the ‘omics data sets they produce have become indispensable to biological research. The progression, over the past two decades, from Sanger sequencing to next-generation sequencing (NGS) and, more recently, long-read sequencing of DNA and RNA has driven cost reductions, improved accessibility, technological progress and availability of supporting tools and resources. Many open-source software packages and annotation pipelines have been developed, not only for genomics but also for the wider family of ‘omics disciplines including proteomics, metabolomics (e.g. glycomics and lipidomics) and others. For the purposes of this review, the term ‘omics refers strictly to the genomics, transcriptomics and proteomics approaches that have become increasingly popular in the bioadhesion literature over the past decade. Characterisation of single genes or proteins in isolation does not constitute genomics or proteomics. ‘Omnis studies produce data for the entire system under investigation, which are then refined by various means to reveal the genes or proteins of interest, and their functions. Transcriptomics via RNA sequencing (RNA-seq) provides a ‘bottom-up’ method for identifying putative proteins based on the principle that molecules of messenger RNA (mRNA) are used to translate genes into proteins somewhat quantitatively. The proteins are then secreted in an unmodified state, or as post-translationally modified variants that can be identified through proteomics. Genes of interest can be targeted by various means, but often by combining differential tissue sampling, analysis and prediction methods.

At the time of writing, over 55000 genome records (https://www.ncbi.nlm.nih.gov/genome/) were publicly available on the National Centre for Biotechnology Information (NCBI) servers. The number of transcriptome and proteome data sets publicly available is less clear due to deposition of data in a variety of archives. However, total numbers for these types of data sets are also in the thousands. For example, a search of the NCBI SRA archive with the key word, ‘transcriptomic’ provides links to over 5000 BioProjects and over 11000 proteomic data sets are available on the EMBL-EBI PRIDE server. These numbers exemplify the ‘omics revolution throughout the biosciences, from the study of human diseases to crop plant production and novel compound discovery (Fukushima et al., 2009; Tanaka, 2010). It is unsurprising, therefore, that the ‘omics approach has also been adopted within bioadhesion research and it is timely to ask what the impact has been on our basic understanding of bioadhesion mechanisms.

Adhesion via a secreted chemical bioadhesive may either be reversible or irreversible, facilitating temporary or permanent attachment (Hennebert et al., 2013b; Lengerer & Ladurner, 2018). Many species use adhesion for essential processes on which their survival depends. Bioadhesion likely evolved multiple times independently and, in each case, it was an adaptation of another pre-existing physiological process. So, while bioadhesion mechanisms are diverse and complex, they are also rooted in core physiological processes that can be interrogated using ‘omics-based approaches. Indeed, it is possible that ‘omics-based studies of adhesion could identify either ancient physiological processes, such as salivary secretion (Sehnal & Akai, 1990; Yan et al., 2020), from which adhesion evolved in different lineages, or similarities among lineages through the presence of functional gene and protein domains. The desirable characteristics of bioadhesive interfaces in nature are not provided by chemistry alone, of course. So-called ‘wet’ (Wilhelm et al., 2017) and ‘dry’ (Labouze et al., 2016) adhesion systems rely on mechanics to enhance their performance: modulating adhesive contact area (Crawford et al., 2016), dissipating energy under stress at the micro- (Cohen et al., 2019) and nanoscales (Phang et al., 2010) and enabling controlled detachment (Federle &
Adhesion by aquatic invertebrates

Scale is important, since extrapolations from experiments at the nanoscale will be unlikely to reflect the true properties at the micro- or macroscales (Desmond et al., 2015). These mechanical phenomena are beyond the scope of this review, the focus of which is on identification and characterisation of secreted materials.

Adhesives secreted by aquatic invertebrates contain proteins, glycans (polysaccharides) and lipids in varying proportions. Often metals are involved and are instrumental to crosslinking (Richter, Grunwald & von Byern, 2018). Interest in bioadhesives has been driven to a significant degree by the demand for novel biomimetic adhesives with capabilities beyond the synthetic glues currently available to consumers. Understanding the mechanisms that control adhesion in aquatic systems is considered to be central to the development of bio-inspired adhesives for the construction, biomaterial and manufacturing industries, as well as for clinical therapies (Palacio & Bhushan, 2012). Many current synthetic adhesives are damaging to the surfaces they are applied to and are contaminating, toxic or hazardous to the environment. In addition, most of those currently available have low efficacy on hydrated surfaces. Substitution with bio-inspired adhesives could therefore provide more suitable and sustainable alternatives (Richter et al., 2018).

Most papers containing the key word ‘bioadhesion’ (or similar) did not focus on the fundamental biology of the natural material. Rather, the majority focused on applications, including cell adhesion technologies, biomimetic materials, and engineering applications (e.g. anti-biofouling). All search results are illustrated together by the dashed lines in Fig. 1. The small subset of those papers focusing on fundamental biological understanding of adhesion to non-self-surfaces is indicated by the solid lines. Those fundamental studies were further sub-divided into bars based upon their primary focus: biochemistry, proteins or genes. These categories are intentionally broad and not intended to be over-interpreted. ‘Biochemistry’ included studies using e.g. histological approaches and enzymatic assays as well as analytical methods to characterise the bulk adhesive secretion. Studies of proteins and genes focused on identification, purification and characterisation of single proteins/gens or proteomes/transcriptomes.

II. HISTORIC TRENDS IN BIOADHESION RESEARCH

As context for the discussion to come, it is worth considering the scale of interest in bioadhesion for a representative range of aquatic organisms, and how the methods of investigation have evolved over recent decades. Historically, barnacles and mussels have been the most intensively researched aquatic invertebrates with respect to their adhesives. To produce Fig. 1, a literature search was conducted using Web of Science (details in Fig. 1 legend) to identify papers referring to adhesion in barnacles, mussels, polychaetes, echinoderms and ‘others’ (containing references to cnidarians, ascidians and platyhelminths). Although by no means exhaustive, this exercise provided an overview of activity in bioadhesion research on a per decade basis.

Most papers containing the key word ‘bioadhesion’ (or similar) did not focus on the fundamental biology of the natural material. Rather, the majority focused on applications, including cell adhesion technologies, biomimetic materials, and engineering applications (e.g. anti-biofouling). All search results are illustrated together by the dashed lines in Fig. 1. The small subset of those papers focusing on fundamental biological understanding of adhesion to non-self-surfaces is indicated by the solid lines. Those fundamental studies were further sub-divided into bars based upon their primary focus: biochemistry, proteins or genes. These categories are intentionally broad and not intended to be over-interpreted. ‘Biochemistry’ included studies using e.g. histological approaches and enzymatic assays as well as analytical methods to characterise the bulk adhesive secretion. Studies of proteins and genes focused on identification, purification and characterisation of single proteins/gens or proteomes/transcriptomes.

It was clear from the total numbers of bioadhesion-related papers that mussels are the group of organisms most commonly referred to in the literature (Fig. 1, right axes), with 1847 papers listed in the decade 2010–2020. When only the basic biology papers (bars/solid lines) were considered, however, the difference in numbers of papers between e.g. barnacles and mussels was relatively small: 37 papers for barnacles and 42 for mussels between 2010 and 2020. For mussels, there was a clear increase in the publication of papers relating broadly to bioadhesion between 2000 and 2020 but no accompanying increase in the number of papers focussing on basic biology. The increasing interest in mussel adhesion, as measured by numbers of papers, was driven by other research priorities, foremost among which was the development of mussel-inspired adhesives. The clear discrepancy between generating basic biological understanding and translating results into bio-inspired technologies was recently commented on by Waite (2019), who advised a renewed focus on the biology.

The number of bioadhesion papers with a fundamental focus on biochemistry, proteins or genes of barnacles, echinoderms and ‘others’ has recently increased dramatically. Nevertheless, for barnacles, mussels and polychaetes these fundamental biology studies still represent a reducing proportion of total publications (increasing difference between dashed and solid lines in Fig. 1, right axes). Biological discovery, overall, remains a minor contributor to the bioadhesion literature for all organisms. For mussels, around the same number of fundamental research papers were published on proteins in the three decades between 1990 and 2020 (14, 14 and 13, respectively) despite the surge in mussel adhesion research more broadly. Echinoderms were exceptional in terms of the increasing basic protein research they attracted.

For all of the groups in Fig. 1, studies of biochemistry appeared before studies focussing on proteins and genes (Fig. 1, bars, left axes). Of particular interest is the group of
others’. These are the less well-established organisms in bioadhesion research where little work was done relating to their adhesion prior to the 2010s. Studies of proteins and genes are now prevalent for those organisms and often also include the type of biochemical work that laid the foundations for other taxa in the 1980s and 1990s. In some cases, however, only ‘omics data are presented (e.g. Davey et al., 2019), and it is important to consider how in silico-derived predictions from such studies may be tested. In barnacles, echinoderms and ‘other’ organisms, studies of proteins and genes now make up more than half of the basic bioadhesion research effort. When only papers contributing substantial new data are considered, 52 relevant papers containing ‘omics data sets have been published in the past 11 years (Table 1). Thirty-three of these included transcriptomic approaches. The same number also made use of proteomics and four used short or long-read approaches to genome assembly. None conducted metabolomics, lipidomics or glycomics.

III. STRENGTHS: IMPROVED UNDERSTANDING OF ESTABLISHED BIOADHESION SYSTEMS

The ‘omics revolution changed the culture of life-sciences research so rapidly that it can be difficult to remember a time when projects did not begin with the generation of high-throughput ‘omics data sets. Species adopted early in bioadhesion research, such as barnacles, mussels and echinoderms, were initially interrogated using the biochemical assays and histological techniques available in the 1970s and 1980s. Below we consider examples of organisms whose adhesion was initially understood using ‘non-omics’ methods, but where later adoption of ‘omics has proved advantageous.

(1) Barnacles

Barnacles, as one of the most intensively studied ‘bioadhesion models’ with literature dating back over 50 years, provide an example of how contemporary high-throughput technologies can build on pre-existing knowledge and provide new directions for research. Barnacle adhesion was recently subject to a comprehensive review (Liang et al., 2019). Our discussion here relates specifically to historic uncertainty surrounding the curing mechanism of barnacle cement, and the contribution of ‘omics to understanding that process.

It seems intuitive that the adhesive secretions of adult barnacles are released as a ‘cement’ that sets into a hardened material. In fact, this is not entirely clear. The ability of barnacles with membranous bases to ‘slide’ on smooth surfaces under their own volition is widely recognised. This is probably not due to any unique adhesive characteristic of membranous-based species, but more likely a result of the action of body movements on surfaces that are transmitted through a membrane, but not through a calcified basis.
Although the primary cement does form a hardened plaque, Kavanaugh, Quinn & Swain (2005) presented evidence of a viscous sub-layer beneath barnacles with calcified bases on a polydimethylsiloxane substrate, suggesting that the material was not completely cured. Nevertheless, the majority of effort in barnacle bioadhesion research has historically focused on the identification of major proteins and the means by which these proteins can interact to form a solid. Seminal work in this area was conducted in a series of papers between 1996 and 2015 (Kamino, Odo & Maruyama, 1996; Kamino et al., 2000; Kamino, 2001, 2010; Nakano, Shen & Kamino, 2007; Urushida et al., 2007; Kamino, Nakano & Kanai, 2012; Nakano & Kamino, 2015). Briefly, they identified strongly reducing conditions [0.5 M dithiothreitol (DTT) in 7 M guanidine hydrochloride at 60°C] necessary to solubilise up to 94% of barnacle adhesive by weight, and resolved the proteins by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Sequence information about individual proteins was then discovered using Edman degradation (N-terminal sequencing), a technique that has since largely fallen from favour. Three major proteins of 52, 68 and 100 kDa were identified in Megabalanus rosa, along with minor constituents of 20, 40 and 180 kDa (Kamino et al., 2000). A 19 kDa cement protein was identified later (Urushida et al., 2007). Although the remaining 6% was also found to be predominantly protein, it could not be solubilised and therefore remained unstudied.

From the solubilising effect of concentrated DTT it was concluded that cysteine residues probably contributed to the insolubility of the barnacle cement via inter- or intramolecular disulphide bonds, and that the alternating polar and non-polar residues discovered in the 100 kDa protein may further stabilise it through the formation of amyloids. This possibility was supported by Fourier transform infrared (FT-IR) analysis of the interface between barnacles and surfaces by Barlow et al. (2010), who identified signatures characteristic of amyloid. There was no evidence for, or speculation regarding, enzymatic crosslinking of proteins in these early studies. In fact, the self-assembly of chemically synthesised peptides based upon the 20 kDa cement protein (Nakano et al., 2007) implied that enzymatic polymerisation was not required. While a logical conclusion based on the

| Author(s) | Species | Type | Author(s) | Species | Type |
|-----------|---------|------|-----------|---------|------|
| Endrizzi & Stewart (2009) | Phragmatopoma californica | ◇ | Santos et al. (2013) | Paracentrotus lividus | ◇ |
| Becker et al. (2012) | Sabellaria alveolata | ▲ | Henmbeert et al. (2014) | Asterias rubens | ◇ |
| Buffet et al. (2018) | Sabellaria alveolata, Phragmatopoma caudata | ◇ | Lebesgue et al. (2016) | Paracentrotus lividus | ◇ |
| Chen et al. (2011) | Amphibalanus amphitrite | ◇ | Lengerer et al. (2019) | Asteroidea spp. | ◇ |
| Jonker et al. (2014) | Lepas anatifera, Pollicipes pollicipes | ▲ | Xu & Faisal (2010) | Dressenella polymorpha | ◇ |
| Liu et al. (2014) | Tetractyla japonica formosana | ◇ | Gantayet et al. (2013) | Dressenella polymorpha | ◇ |
| Wang et al. (2014) | Hesperophylax occidentalis | ◇ | Guerette et al. (2013) | Doublasia gigas, Perna viridis, P. ciliolatum | ◇ |
| Zhechen et al. (2014) | Dorsa fascicularis | ◇ | Li et al. (2017) | Pinctada fascata | ◇ |
| Wang et al. (2015) | Amphibalanus amphitrite | ◇ | Zhang et al. (2017) | Perna viridis | ◇ |
| G. Zhang et al. (2015a) | Amphibalanus amphitrite | ◇ | DeMartini et al. (2017) | Mytilus californianus | ◇ |
| Nomura et al. (2016) | Stenopside marmorata | ◇ | Li et al. (2017) | Cellana toreuma | ◇ |
| So et al. (2017) | Amphibalanus amphitrite | ◇ | Li et al. (2017) | Chlamys farri | ◇ |
| Luo et al. (2018) | Stenopside tienmushanensis | ◇ | Foulon et al. (2018) | Gracillaria gigas | ◇ |
| Abramova et al. (2019) | Balanus improvisus | ◇ | Wang et al. (2018) | Cellana toreuma | ◇ |
| Machado et al. (2019) | Leptosphaera, Pollicipes pollicipes | ◇ | Dou et al. (2018) | Ishnochiton hokkaidensis | ◇ |
| Correa-Garhwal et al. (2019) | Dolomedes triton | ◇ | Li et al. (2016) | Lumnoperla fortunei | ◇ |
| Frandsen et al. (2019) | Parapagurus doris | X | Li et al. (2018) | Dressenella bagensis | ◇ |
| Rocha et al. (2019) | Pollicipes pollicipes | ◇ | Rees et al. (2019) | Dressenella bagensis | ◇ |
| Domínguez-Perez et al. (2020) | Pollicipes pollicipes | ◇ | Kang et al. (2020) | Potamilla sulphata | ◇ |
| Engel et al. (2021) | Leptosphaera | ◇ | | | |
| Yan et al. (2020) | Megabalanus volcano | ◇ | | | |
| Pennati & Rothbacher (2015) | Ascidian spp. | ◇ | Fincher et al. (2018) | Schmidtiella mediterranea | ◇ |
| Li et al. (2019) | Ciona robusta | ▲ | Lengerer et al. (2018) | Macrostomum lignano | ◇ |
| Moya et al. (2012) | Acropora millepora | ▲ | Wunderer et al. (2019) | Macrostomum lignano | ◇ |
| Rodrigues et al. (2016) | Hydra magnipapillata | ▲ | Pjeta et al. (2019) | Minima illeanea | ◇ |
| Davey et al. (2019) | Exaiptasia pallida | ◇ | | | |
| Wang et al. (2020) | Halplanaula luciae | ▲ | | | |

Table 1. A catalogue of relevant ‘omics and functional bioadhesion studies identified at the time of writing.

- ◇, short-read genomics; X, long-read genomics; ◇, transcriptomics; ▲, quantitative polymerase chain reaction/microarray; ■, proteomics; ◇, review; ◇, in situ hybridisation; ◇, functional gene knockdown.
First, self-assembly does not prove the absence of polymerisation. Further, synthetic peptide fragments of the 20 kDa (Nakano et al., 2007) and 52 kDa cement proteins (Nakano & Kamino, 2013) self-assembled into nanofibres only under specific conditions of pH and salt. Recent evidence (Mohanram et al., 2019) from nuclear magnetic resonance (NMR) studies of the 20 kDa cement protein tertiary structure have confirmed that 12 out of 32 cysteines in the sequence engage in disulphide bonds. Molecular dynamics simulations highlighted conserved β-motifs (β7-β8), which may function as nuclei for amyloid-like nanofibres. Fears et al. (2019) argued that this 20 kDa protein was in fact a shell protein misclassified as a cement protein, and its absence in the adhesives of barnacles with membranous bases seems to support this view. In this case, the behaviour of synthetic fragments may have little bearing on our understanding of barnacle adhesion.

Dickinson et al. (2009) proposed an alternative model for the curing of barnacle cement based upon the glutamyllysine crosslinking action of transglutaminase, related to wound healing. This work proved to be controversial and one line of evidence presented by Dickinson et al. (2009), namely the presence of epsilon (γ-glutamyl) lysine crosslinks in the cured cement, was questioned by Kamino (2010). Kamino et al. (2012) maintained that enzymatic processing was not required for barnacle cement curing, at least in the case of the dominant 52 kDa protein, and that self-assembly was driven by inter- and/or intramolecular disulphide bonds, hydrogen bonds or interactions of aromatic amino acids (Nakano & Kamino, 2015). Although the matter remains unsettled, several intriguing leads have emerged from more recent ‘omics-based analyses. The early work by Kamino and others focused on isolation and characterisation of single, major proteins extracted in abundance from secreted cement. This approach had the advantage that those proteins were unambiguously present at the adhesive interface, however the focus on specific dominant proteins potentially missed a large number of minor but nevertheless important cement components. In the first large-scale ‘omics analysis of barnacle adhesion, So et al. (2016) used a modified digestion protocol involving the solvent hexafluoroisopropanol to liberate up to 90 putative proteins, identified by their alignment to a basal tissue transcriptome. Numerous new details emerged from the data set, including abundant proteins that were previously undescribed, as well as those noted by Kamino. Several of the previously undescribed proteins were enzymes, including seven oxidoreductases of which three were lysyl oxidase (LOX) homologues. In a study of oxidative activity beneath the base of an attached barnacle, So et al. (2017) identified that LOX was indeed present and active at the adhesive interface. Several of the putative adhesive proteins present in the data of So et al. (2016), including Amphibalanus amphitrite cement proteins (AaCPs) AaCP19 and AaCP52 (following their updated naming convention), contained substantial numbers of lysine residues (>10%) and could therefore represent plausible substrates for LOX activity. Given the common role of LOX in crosslinking elastin and collagen in vivo, a role in cement polymerisation cannot be ruled out. Recently, LOX was shown to be over-expressed in cyprid cement glands, thus indicating that allysine-mediated cross-links might be involved in cyprid adhesive curing as well (Yan et al., 2020). If this is proved to be the case, high-throughput ‘omics approaches will have made a substantial contribution to our understanding of one of the most high-profile bioadhesion models.

It should be noted that the sequence similarity between adhesive proteins from different barnacle species can be relatively low, and that alignment-based studies therefore need to be executed with caution. Homologous proteins from barnacle adhesive can have amino acid sequence similarity below 30%, possibly due to the different lifestyles, habitats and substrate affinities of individual species. In a comparison of the 100 kDa cement proteins (Kamino et al., 2000) of Megabalanus rosa and Amphibalanus amphitrite, it was impossible to match any peptides of over five amino acids in length (Kamino, 2010). By contrast, the settlement inducing protein complex (SIPC), used for conspecific recognition by barnacles and also present in the temporary adhesive ‘footprints’ of exploring barnacle cyprids, showed 63–76% sequence similarity among species (Yorisue et al., 2012). The conserved sequence of SIPC could be required for its proposed multifunctionality – acting both as a gregarious settlement ‘pheromone’ and adhesive constituent (particularly in larvae). Petrone et al. (2015) provided evidence for the unusual non-specific affinity of the SIPC for a wide range of surfaces and speculated on its possible role in temporary adhesion of barnacle larvae. If proved, SIPC would present an example of pre-existing physiology being diverted to a role in adhesion as this molecule has its origins in the ancient alpha-2-macroglobulin family of blood complement proteins. However, for putative barnacle adhesive proteins that diverged to a greater degree based on functional requirements, it will be necessary to look beyond simple sequence alignments and to examine other properties of the proteins. Strong biases in amino acid composition and elevated isoelectric points (pIs > 9) have been found in some barnacle adhesive proteins, such as in the 100 and 19 kDa proteins (Rocha et al., 2019, Yan et al., 2020). Amino acid sequence per se may be less important in these cases than the propensity for the protein to form particular secondary structures, such as amyloids, or perform specific interactions. Glycine enrichment (up to 20% of amino acids in 19 kDa homologues) may aid folding into amyloid cross-β sheets and the consistently high isoelectric point of 19 kDa homologues across species may suggest control of folding by the ambient pH (Tilbury et al., 2019).

(2) Mussels

The classic example in which functional understanding of an adhesion system was established prior to the ‘omics revolution is the attachment of mytilid bivalves and particularly
the blue mussel, *Mytilus edulis*. These organisms secure themselves to the substratum by producing an extra-organismic holdfast, the so-called byssus (Waite, 1985, 2017). The byssus consists of a bundle of threads connected proximally to the base of the animal’s foot, within the shell, and terminating distally with a flattened plaque which mediates adhesion to the substratum (Lee et al., 2011; Waite, 2017). The composition of the plaque was originally determined using biochemical and ultrastructural studies, revealing a collagogenous substance, a mucous material and a polyphenolic substance (Brown, 1952; Tamarin, Lewis & Askey, 1976). The phenolic substance was identified as a DOPA (3,4-dihydroxyphenyl-L-alanine)-containing protein (Waite & Tanzer, 1981). Thirty years of biochemical and molecular biology research led to the identification of nine proteins constituting the byssus attachment plaque, two collagens (preCol D and preCol NG), one thread matrix protein (TMP) and six mussel foot proteins (mfp-1–6) [see Lee et al. (2011) for review]. All mfps contain DOPA, a catecholic amino acid originating from the post-translational modification (PTM) of tyrosine residues (Waite, 2017). At the plaque adhesive interface, DOPA residues in the proteins mfp-3 and mfp-5 mediate adhesion via formation of a number of transient and covalent interactions with surfaces. In the bulk of the plaque and in the thin protective outer coating, known as the cuticle, intermolecular DOPA–metal complexation by mfp-2 and mfp-1 provides mechanical reinforcement. A fraction of the DOPA residues, in particular in the preCols, may also covalently cross-link via reaction with DOPA residues or other amino acids. These reactions are driven by the oxidation of DOPA to its quinone form, which occurs spontaneously at the pH of sea water (Waite, 2017; Priemel et al., 2020). Because of the central involvement of DOPA, most mussel-inspired adhesives developed so far are dominated by DOPA polymer constructs (Lee et al., 2011). However, as stressed by Waite (2017), mussel adhesion does not depend only on this single molecular entity. For instance, some mfps contain other post-translationally modified amino acids such as O-phosphoserine (mfp-5 and -6), 4-hydroxyproline (mfp-1), 3,4-dihydroxyproline (mfp-1), and 4-hydroxyarginine (mfp-3), which could all mediate non-covalent interactions with surfaces or with other proteins (Waite et al., 2005; Silverman & Roberto, 2007).

Guerette et al. (2013) proposed integration of transcriptomics and proteomics to accelerate the characterisation of biological materials in general, and of biological adhesives in particular. This approach allowed retrieval of the full-length sequences of five mfp orthologs from the green mussel *Perna viridis* (Pvfp-1, -2, -3, -5 and -6), which could not be identified by simple homology searches. Since then, this combined ‘omics approach has been used in several mussel species (both from marine and freshwater environments) as well as in other bivalve species such as oysters and scallops (Table 1), identifying homologues of the canonical mfps but also a whole range of novel byssal proteins. DeMartini et al. (2017) took this a step further and conducted transcriptomic analyses of the different foot glands in the mussel *Mytilus californianus*. They found around 15 highly expressed proteins that had not been characterised previously, but bore key similarities to the previously defined mfps, suggesting additional contribution to byssal function. Mass spectrometry (MS) analyses of proteins extracted from freshly secreted byssal threads and plaques confirmed their presence in the byssus. Recently, Jehle et al. (2020) proposed that some of these new cysteine-rich mfps (mfp-16–19) would function in cross-linking the byssus cuticle as well as in counteracting the spontaneous oxidation of DOPA. These recent results demonstrate quite clearly the additional interest and contribution that ‘omics technologies have made to studies of mussel adhesion, despite its long history in bioadhesion research.

(3) Echinoderms

Although studies of echinoderm adhesion used histochemical, mechanical and morphological methods for decades before the large-scale adoption of ‘omics (Fig. 1), several species have now been studied in detail using combinations of transcriptomics and proteomics. In sea stars and sea urchins, adhesion takes place via the tube feet (or podia), which consist of a proximal stem (non-adhesive part) and a distal disc (adhesive part). Tube feet can detach voluntarily, leaving the adhesive material on the substrate as a ‘footprint’.

MS-based proteome analysis of footprints of the sea star *Asterias rubens* combined with RNA-seq data (a transcriptome) identified 34 proteins in the secreted adhesive. Sequence similarity searches against the NCBI non-redundant database resulted in the functional annotation of 20 of these proteins, while 14 remained unidentified (Hennebert et al., 2015a). Whole-mount in situ hybridisation (WISH) confirmed that these 34 footprint proteins were spatially expressed in the tube foot epidermis: 22 were exclusively expressed in the disc epidermis and 12 exhibited additional expression in the stem epidermis (Lengerer et al., 2019). One abundant protein originally annotated as an immunoglobulin isotype GFC (IgGFC)-binding protein was identified as a major structural protein involved in foot-print cohesion and renamed sea star footprint protein 1 (Sfp1) (Hennebert et al., 2014). Sfp1 is 3853 amino acid residues long, contains various functional domains involved in protein–protein and protein–carbohydrate interactions [calcium-binding epidermal growth factor (EGF)-like domains, galactose-binding lectin domains, disoidin domains (also known as F5/8 type C domains), von Willebrand Factor type D (vWD) domains, trypsin inhibitor-like cysteine rich (TIL) domains and C8 domains] and is auto-catalytically cleaved into four subunits before secretion (Hennebert et al., 2014). In recent bioinformatic analyses, sequences from the *A. rubens* data set were used for Basic Local Alignment Search Tool (BLAST) searches against seven transcriptomes from Asteroidea species, highlighting substantial conservation of the large proteins that make up the structural core of the adhesive footprint (e.g. Sfp-1). Smaller, putative surface-binding proteins appeared to be more variable among sea
star species (Lengerer et al., 2019). Such comparisons were made possible by the availability of large ‘omics data sets.

For the sea urchin Paracentrotus lividus, quantitative proteomics enabled comparison of protein expression levels in the tube foot disc versus the stem, in combination with the footprint protein profile (Lebesgue et al., 2016; Toubarro et al., 2016). This resulted in the identification of 163 proteins over-expressed in the disc, yet the analysis only allowed for confident identification of highly conserved proteins. This limitation occurred because mapping of the MS-derived peptides relied on multiple incomplete publicly available sea urchin protein databases. Within these proteins, only one, nectin, had a reported adhesive function (P. lividus egg nectin significantly increases the binding of dissociated embryonic cells to the substratum; Matranga et al., 1992) and a nectin variant was also shown to be present in the tube foot adhesive secretion of adult P. lividus (Lebesgue et al., 2016). This 108 kDa protein presents phosphorylated and glycosylated isoforms and contains six discoidin domains (similar to Sfp1) that can bind molecules bearing galactose and N-acetylglucosamine residues (Santos et al., 2013; Lebesgue et al., 2016; Toubarro et al., 2016). Recently, these proteomic data were re-mapped to a new P. lividus tube foot transcriptome (Pjeta et al., 2020). This resulted in a 60% increase in the mapped disc and stem peptides, accompanied by a 71% increase in the number of identified proteins. A total of 121 transcripts were overexpressed in the tube feet discs and simultaneously present in previous disc and/or adhesive secretion proteome data sets, but not in the stem proteome. Wholomount in situ hybridisation (WISH) was performed for 59 selected transcripts, pinpointing 16 transcripts potentially involved in sea urchin adhesion. Of these, six transcripts were identified as nectin, alpha-tectorin, uncharacterised protein, myeloperoxidase, neurogenic locus notch homolog protein and alpha-macroglobulin, but simultaneously shared orthology with putative adhesion-related genes from sea stars (Lengerer et al., 2019; Pjeta et al., 2020). The advantages of ‘omics and, more specifically, the advantages of bespoke, comparative ‘omics analyses (Fig. 2) are therefore clearly evident in studies of echinoderm biodhesion.

(4) Caddisflies and other arthropods

‘Omic techniques provide convenient means to introduce taxonomic breadth into analyses, highlighting common biomolecular features in adhesives. One example is the silks of arthropods. By considering both conservation and variation within silks that are adapted to dry and hydrated surfaces, it may be possible to identify molecular tricks particular to wet-functioning bioadhesives, which may not be obvious otherwise. Insect silk is a secretory product of both caddisfly (Trichoptera) and moth (Lepidoptera) larvae, the comparison being interesting since caddisflies are aquatic ‘relatives’ of more intensively researched terrestrial silk producers, e.g. Bombyx mori. Caddisflies are freshwater specialists in the larval form, using silk to make protective cases and nets for food capture. Caddisfly species exhibit compositional differences in their cases, including the choice of particle type and grain size (Fransden et al., 2019), which may require differences in silk composition. In an example of pre-existing physiology being diverted to a role in adhesion, and in common with Lepidoptera (Yonemura et al., 2009), caddisfly silk is produced from paired labial silk glands that develop during embryogenesis as ectodermal invaginations (Sehnal & Akai, 1990).

The silks secreted by both Trichoptera and Lepidoptera have the same origin and are similar in the sense that both are composed of a fibrous core and a sticky coating (Sehnal & Zurovec, 2004). The core contains heavy- and light-chain fibroins (H- and L-fibroin) encoded by a pair of fibroin genes. Together they form hydrophobic fibres responsible for the silk’s tensile strength. The peripheral layer, on the other hand, is composed of hydrophilic molecules responsible for adhesion, including sericin proteins. Full-length sequences of some silk genes from the larvae of caddisflies were unobtainable until recently, with only partial sequences recovered from short-read RNA-seq studies (Yonemura et al., 2009; Wang et al., 2010; Ashton et al., 2013; Luo et al., 2018). The combination of transcriptomic and proteomic analyses in larvae of the common European caddisfly, Hydrophyse angustipennis, confirmed the relatively uniform structure of the fibre core, consisting of two fibroin subunits (Yonemura et al., 2009). ‘Omic methods have shown that lepidopteran and trichopteran silk is a much more complex mixture of proteins than previously thought, and that its composition needs to be revised. More than 280 proteins have been found in B. mori silk (Y. Zhang et al., 2015b) and more than 200 proteins can also be detected in the silk of the caddisfly Hydrophyse angustipennis (L. Rouhova & M. Zurovec, in preparation).

These include highly abundant structural and adhesive silk components, antimicrobial peptides and protease inhibitors, as well as numerous less-abundant cellular proteins (ribosomal proteins, metabolic enzymes), which enter the silk through apocrine-like silk gland secretion, similar to the salivary glands of Diptera (Farkaš et al., 2014). A typical example of caddisfly adhesives is nest-forming protein 1, from Hydrophyse sp. (Eum et al., 2005). It is highly abundant and contains many repetitive sequences. Nest-forming protein 1 differs from terrestrial (silkworm) serine-rich adhesives with an amino acid composition characterised by a high proportion of tyrosine, cysteine, tryptophan and histidine residues. Using ‘omic methods it may therefore be possible to compare the coatings from a number of caddisfly species, detect the major proteins and clarify the evolutionary and structural relationships among them.

A final, and slightly unusual, representative within aquatic arthropods is the European freshwater spider Argyroneta aquatica in which a ‘diving bell’ is created by a web sheet that is submerged underwater, allowing air to be transported from the surface onto the plastron, and stored, ultimately enabling this spider to breathe underwater using its tracheal system (De Bakker et al., 2006). Using existing transcriptomes of selected terrestrial species, and comparing them with a new transcriptome from A. aquatica, Strickland et al. (2018) showed...
high similarity in C-terminal amino acid sequences in spider spidroins over the range of habitats. This suggested a highly conserved mechanism of silk assembly in spiders despite their functional diversity and adaptation for life underwater in the case of *A. aquatica*. More recently, one hydrophobic amino acid motif (GV) was found to be restricted to spidroins of aquatic spiders (Correa-Garhwal et al., 2019). In fact, ‘omics provide opportunity for even broader analysis, by comparison of these aquatic silks to those of, for example, marine amphipods that manufacture silken cases in algal holdfasts (Kronenberger, Dicko & Vollrath, 2012).

### IV. OPPORTUNITIES: ORGANISMS WITH PRE-EXISTING ‘OMICS RESOURCES

In the cases of barnacles and mussels, adhesion mechanisms were proposed before the widespread adoption of high-throughput ‘omics. For barnacles, ‘omics techniques provided additional data to develop the picture more completely, highlighting previously undescribed adhesive components and new avenues of research. The same holds true for mussels, although to a lesser extent: perhaps due to the maturity of that field prior to the ‘omics revolution, or because the research community studying fundamental adhesion processes in mussels had remained largely unchanged since the 1990s (Fig. 1). It may also be that the data required to further our understanding of mussel adhesion cannot be derived from ‘omics studies (see e.g. Valois, Mirshafian & Waite, 2020). Recently, however, novel organisms for the study of bioadhesion have been selected for their existing ‘omics resources and not for more practical reasons as was the case for barnacles and mussels (both of which are problematic biofouling organisms).

The anemone *Exaiptasia pallida* was established as a model for coral/zooxanthellae endosymbiosis, which led to the sequencing of a genome in 2015 (Baumgarten et al., 2015).
Macrostomum lignano, a flatworm, was selected for developmental biology research due to its compact genome and remarkable regenerative ability (Wudarski et al., 2020). The sea squirt Ciona intestinalis represents a group of invertebrate organisms most closely related to vertebrates, and its genome was sequenced very early in the ‘omics revolution (Delsuc et al., 2006, 2018). Contrary to the cases presented in Section III, pre-existing resources were the basis for ‘omics-based studies of adhesion in these species and the typical pipeline is outlined in Fig. 2.

(1) Cnidaria

The Cnidaria are an ancestral metazoan taxon that evolved the ability to stick to surfaces at least 505 million years ago (Clarke, Davey & Aldred, 2020). The adhesion capabilities of two distinctly related species have been studied using genomic/transcriptomic pipelines: the freshwater hydrozoan Hydra vulgaris (previously H. magnipapillata; Rodrigues et al., 2016) and the marine anthozoan Eupaxiopsis pallida (Davey et al., 2019). These species are superficially similar morphologically, existing in the reproductive form as solitary polyps, although they differ significantly at the ultrastructural level (Clarke et al., 2020) and can be separated by considerable genetic distance (Kayal et al., 2018). They adhere to surfaces using secretions released from their pedal discs (Rodrigues et al., 2016; Davey et al., 2019; Clarke et al., 2020), and genome and transcriptome assemblies exist for both species (Chapman et al., 2010; Baumgarten et al., 2015; Petersen et al., 2015).

Differential expression analysis and read count abundances (pedal-disc tissue subtracted from the remaining body) were used to identify candidate genes in H. vulgaris. Expression and localisation of mRNA transcripts were then verified using WISH and high-throughput proteomics (Rodrigues et al., 2016). A similar transcriptomic subtraction approach was taken for E. pallida (Davey et al., 2019). However, in the latter, in silico analyses were used to predict the putative pedal disc secretome. Identification of follow-up candidates in both studies was facilitated by prior functional annotation of adhesion-related genes in the reference genomes, as well as similarity of functional domains to adhesion-related genes in other aquatic invertebrates. Despite their relatedness and morphological similarity, no formal comparative analysis had been completed for these two species. To demonstrate the potential of ‘omics in facilitating powerful interspecies comparison between models with well-developed resources, and to reinforce the value of these publicly available data sets, we present here the results of a conditional reciprocal best BLAST (CRBB) analysis (Aubry et al., 2014; see online Supporting Information, Appendix S1) between H. vulgaris and E. pallida. CRBB analysis finds orthologs between two sets of sequences by conducting a reciprocal BLAST and fitting a function to the distribution of alignment e-values over sequence lengths, to predict an appropriate e-value threshold cut-off (Aubry et al., 2014). In total, only 20.33%, or 11285 of 55496 H. vulgaris transcripts, had a CRBB match to the E. pallida models. Despite this low orthology, 14 candidate genes were identified in both species (Table 2). Moreover, two transcripts previously verified in H. vulgaris using WISH were identified in E. pallida. Before the time-consuming process of protocol development and probe synthesis for WISH was begun for E. pallida, therefore, this straightforward bioinformatic comparison enabled by the established genomic resources provided confidence that those efforts would likely be rewarded.

(2) Platyhelminthes

The first organism to receive comprehensive ‘omics-based investigation of its adhesion was one that had accumulated significant resources as a model for developmental biology. The flatworm, Macrostomum lignano is a meiofaunal invertebrate that uses adhesion for temporary attachment and motility. M. lignano has well-established mRNA transcript localisation and functional RNA interference (RNAi) knock-down techniques (Wudarski et al., 2020), both of which were deployed to study its adhesion. The duo-gland adhesive organs of M. lignano are located at the tail of the animals and their morphology was described by Lengerer et al. (2014, 2016). Body-region specific RNA-seq during regeneration revealed approximately 300 up-regulated transcripts in the tail (Arbore et al., 2015; Lengerer et al., 2018) and WISH confirmed that those transcripts were exclusively expressed within the cells of the duo-gland adhesive system (Lengerer et al., 2018; Weber et al., 2018). RNAi of an intermediate filament and a formin-like protein demonstrated the critical role of the support cells (so called ‘anchor cells’) in adhesion (Lengerer et al., 2014, 2018). Recently, it was confirmed that attachment of M. lignano relies on two large proteins expressed in the adhesive glands (Mlig-ap1 and Mlig-ap2; Wunderer et al., 2019). Similar transcriptomics, proteomics and expression analyses were also applied to the proseriate flatworm Mimona ileanae, where nine transcripts specific to the adhesive organs were identified (Pjeta et al., 2019). Six transcripts had similar conserved domain architecture and were rich in repetitive motifs previously observed in adhesive proteins Mlig-ap1 and Mlig-ap2 of M. lignano. Interestingly, a region comprising tandem repeats of a glycine, arginine and lysine low-complexity motif, which encompassed about two-thirds of M. lignano Mlig-ap1, was identified in a M. ileanae adhesive protein (Mile-ap3). Long-read genomic DNA sequencing using Oxford Nanopore Technology identified that four of these nine transcripts were part of two larger genes (Pjeta et al., 2019).

The morphology and composition of the adhesive organs of a wide range of Platyhelminthes are currently being investigated using ‘omics-based screening. Such a large-scale investigation may highlight parallels and contrasts among lineages and habitats. One significant habitat difference is of course salinity; some flatworm species are marine while others inhabit fresh water, unlike the aforementioned Cnidaria where the freshwater Hydra spp. are something of an

Biological Reviews 96 (2021) 1051–1075 © 2021 The Authors. Biological Reviews published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society.
Table 2. Fifteen putative adhesion orthologs up-regulated in the pedal disc of both *Eaupitaspallida* and *Hydra vulgaris*. Note AIPGENE17511 has two conditional reciprocal best BLAST (CRBB) matches in *H. vulgaris* that are likely alternative transcripts. ECM, extracellular matrix; GPCR, G-protein coupled receptor.

| Query          | Target          | ID %  | E-value | Biscore | *E. pallida* annotation                        | Role            |
|----------------|-----------------|-------|---------|---------|-----------------------------------------------|-----------------|
| HYRNA1402_11381** | AIPGENE26763    | 65.82 | 2.00E-175 | 499     | Glutamine synthetase                          | Detoxification  |
| HYRNA1402_16327 | AIPGENE1297     | 53.85 | 6.00E-13 | 69.3    | Matrix metalloproteinase-9                    | ECM degradation |
| HYRNA1402_1949  | AIPGENE1297     | 53.85 | 2.00E-12 | 71.6    | Matrix metalloproteinase-9                    | ECM degradation |
| HYRNA1402_13839 | AIPGENE14299    | 52.54 | 1.00E-13 | 72.4    | Collagen alpha-1[XII] chain                   | Structural integrity |
| HYRNA1402_19546 | AIPGENE17511    | 51.13 | 2.00E-39 | 141     | Golgi-associated plant pathogenesis-related protein 1 | Fibril formation |
| HYRNA1402_12480 | AIPGENE17511    | 44.85 | 2.00E-47 | 166     | Golgi-associated plant pathogenesis-related protein 1 | Fibril formation |
| HYRNA1402_4149  | AIPGENE1297     | 43.3  | 7.00E-17 | 85.5    | Matrix metalloproteinase-9                    | ECM degradation |
| HYRNA1402_5675.1 | AIPGENE23611    | 41.67 | 5.00E-55 | 197     | Rhamnose-binding lectin                       | α-galactosyl binding |
| HYRNA1402_18715**| AIPGENE4540     | 40.5  | 6.00E-18 | 82.4    | Latrophilin-1                                 | Adhesion-GPCR receptor |
| HYRNA1402_13504 | AIPGENE25956    | 38.62 | 1.00E-36 | 134     | Heme-binding protein 2                        | Heme binding    |
| HYRNA1402_10025 | AIPGENE2202     | 35.29 | 6.00E-17 | 80.9    | Vitelline membrane outer layer protein 1 homolog | Film stabilisation |
| HYRNA1402_4124  | AIPGENE10560    | 27.86 | 2.00E-75 | 264     | Chondroitin proteoglycan 2                    | Structural integrity |
| HYRNA1402_13799.1| AIPGENE19685    | 27.53 | 2.00E-25 | 108     | Expansin-YoaJ                                 | Carbohydrate binding |
| HYRNA1402_651   | AIPGENE1176     | 27.44 | 3.00E-49 | 194     | Hemicentin-1                                  | Structural integrity |
| HYRNA1402_322   | AIPGENE27786    | 25.29 | 1.00E-33 | 141     | Agrin                                         | Neural cell adhesion |

**,** messenger RNA (mRNA) confirmed by wholemount in situ hybridisation in the pedal disc of *H. vulgaris*. Note AIPGENE17511 has two conditional reciprocal best BLAST (CRBB) matches in *H. vulgaris* that are likely alternative transcripts. ECM, extracellular matrix; GPCR, G-protein coupled receptor.

oddity. For illustrative purposes, a novel comparison is presented here between *M. lignano* (marine) and the freshwater species *Macrocrystum poznaniense*. The same adhesive organ morphology is present in *M. lignano* (Lengerer et al., 2014) and *M. poznaniense* (Fig. 3A, C). Adhesive organs are restricted to the posterior of the animal. A transcriptome for *M. poznaniense* (Appendix S2) was assembled in a similar manner to that of *M. lignano* and was found to consist of 179871 transcripts. Assembly completeness was verified by Benchmarking Universal Single-Copy Orthologs (BUSCO), where 89% of the 954 metazoan universal single-copy orthologues were identified (Appendix S2, Fig. S1). A BLAST search against Mlig-ap1 and Mlig-ap2 identified two transcripts encoding two proteins (Mpoz-ap1 and Mpoz-ap2) which appear to have a highly similar protein domain architecture compared to *M. lignano* adhesion proteins (Fig. 4). Furthermore, highly similar conserved domain architecture was identified in those *M. poznaniense* transcripts. WISH revealed a distinct expression pattern of Mpoz-ap1 and Mpoz-ap2 in the tail plate of *M. poznaniense* at the location of the secretory cell bodies (Fig. 3D, E). In addition, an RNAi-mediated knock-down of Mpoz-ap2 led to non-adhesive animals (N = 7; Supplementary Movie S1). From these analyses we can conclude that similar adhesive proteins are used by these marine and freshwater macrostomid flatworm species. Notably, these findings highlight the conservation of the cohesive proteins (Mlig-ap1, Mpoz-ap1; Mile-ap1) among flatworm lineages. By contrast, the glue proteins, especially the repeat regions, appear to be less conserved between flatworm clades (e.g. Mlig-ap2; Mile-ap2). BLAST searches yield homologs only between closely related species (as seen above with the glue proteins, Mlig-ap2 and Mpoz-ap2, from two species of the same genus) but not between species from different clades. Again, such rapid analyses are possible only when supported by ‘omics datasets.

(3) Ascidians

The ascidian (sea squirt) *Ciona intestinalis* is a ‘true’ model organism, having been used for decades in developmental biology studies. It adheres to surfaces at the larval stage and is a problematic marine fouling organism (Aldred & Clare, 2014). The adhesion of the tadpole larva is initially reversible, with final settlement/permanent adhesion triggering metamorphosis to the adult. Adhesion is mediated by rostral papillae that secrete adhesive material by epithelial secretion (Ueki et al., 2014) at the location of the secretory collocytes containing two types of vesicles with fibrous polysaccharides and glycoproteinaceous contents, respectively (Zeng et al., 2019a). Adhesion continues throughout growth of the animal via the ampullae, which are holdfast extensions of the tunic that lack glandular organs but that seem to produce adhesive material by epithelial secretion (Ueki et al., 2018).

Both solitary and colonial ascidians have received historic interest as the closest living invertebrate relatives of vertebrates. Studies of the regulatory genome during development began with the sequencing of the *Ciona intestinalis* Type A genome 20 years ago (Delsuc et al., 2006). Besides sequence deposits for *C. intestinalis* in general public databases, more sophisticated, anatomical, expression and regulatory data sets can now be interrogated and cross-queried for 15 different species (14 sessile and one pelagic) in the ascidian network for *in situ* expression
and embryological data (ANISEED) (Dardaillon et al., 2019). Molecular analysis of C. intestinalis adhesion therefore benefits from an ever-expanding body of genomics data. Despite this, C. intestinalis has been a relative latecomer to the bioadhesion literature. Adhesive extensions (stolons) of adult Ciona robusta (formerly C. intestinalis Type A) were recently subjected to proteomic analysis. 26 proteins were identified, of which six were previously uncharacterised (Li et al., 2019). In common with the echinoderm, cnidarian and platyhelminth species discussed above, these proteins contained adhesion-relevant protein domains such as thrombospondin type 1 (TSP-1) or EGF domains, and von Willebrand Factor A (vWA) domain in the case of ascidian stolon protein 1 (ASP-1). Adhesive properties of recombinant ASP-1 were demonstrated to increase upon artificial DOPA-modification of tyrosines, but whether ascidians rely on ASP-1 for their adult adhesion remains unconfirmed, and the specific role of DOPA is circumstantial. Indeed, DOPA and TOPA (L-3,4,5-trihydroxyphenylalanine) modifications are involved in ascidian tunic wound healing and were previously utilised to generate catechol-chemistry mimetic glues (Oh et al., 2015; Zhan et al., 2017). Overall, advances in understanding commonalities with other bioadhesive systems has been rapid in ascidians due to their ‘model organism’ status and available ‘omics resources.

V. LIMITATIONS: IMPORTANT CONSIDERATIONS FOR ‘OMICS-BASED BIOADHESION STUDIES

Current ‘omics approaches are not without limitations. Sometimes failure to provide enough methodological information results in studies that lack reproducibility, clarity, interoperability and adaptability. Good programming practice and documentation in open-source software is essential for future users to comprehend fully the intended function of code and associated parameters. These and other methodological issues will be resolved as the techniques mature but, nevertheless, it seems inconceivable that all necessary information will ultimately be derived from ‘omics approaches alone. Multidisciplinarity will remain essential and specific limitations are discussed below.

1) Sampling for transcriptomics and proteomics studies

For MS-based proteomics analyses it is essential first to have an adequate transcriptome/genome assembly. Existing databases for little-studied species are rarely adequate. For example, mapping the tube foot proteome of the sea urchin Paracentrotus lividius to a corresponding transcriptome yielded almost 3.5 times more identified proteins than searching public databases using “sea urchin” as the search parameter (Pjeta et al., 2020). Even so, there remains significant variation in sampling methods for preparation of ‘omics data sets. In some cases, single animals have been sampled to represent one true biological replicate while, in others, several animals have been pooled for this purpose. Sequencing replicates of single animals is the most biologically and statistically correct way to quantify natural variation within a population, although in some cases (small or otherwise refractory organisms) researchers may struggle to produce sufficient quantities of high-quality RNA from single individuals. Some single-cell techniques do allow low-yield RNA libraries to be prepared and sequenced effectively, however, and such...
approaches should at least be considered before compromising statistical rigour.

Another issue that is common in micro-scale sampling of natural populations is the concern surrounding sampling only the target species. Extracting RNA/proteins from several cryptic or indistinguishable species is not an unlikely scenario in some environments and can lead to assemblies with inflated transcript/protein numbers, rendering the data useless. Where multiple species are sampled, either knowingly or unknowingly, the increased diversity within samples

![Conserved protein domain comparison between Macrostomum lignano ap1 and Macrostomum pozianiense ap1](image1)

![Conserved protein domain comparison between Macrostomum lignano ap2 and Macrostomum pozianiense ap2](image2)

Fig 4. Basic Local Alignment Search Tool search of (A) Macrostomum lignano adhesion protein 1 (Mlig-ap1) and (B) M. lignano adhesion protein 2 (Mlig-ap2) against the newly assembled and conserved Macrostomum pozianiense transcripts, namely Macrostomum pozianiense_DN1831_c0_g1_i26 (Mpoz-ap1) and Macrostomum pozianiense_DN2768_c0_g1_i4 (Mpoz-ap2). A conserved protein domain search of the translated transcripts showed a highly similar domain architecture between M. lignano and M. pozianiense adhesion proteins. Note that only the multi-domain sequence of Mlig-ap1 is represented. The length of the Mpoz-ap2 repetitive region remains unclear (marked as dashed line with question mark above). In addition, note that the repeat regions 1 and 2 (R1 and R2) in Mlig-ap2 are truncated (dashed lines) to show correct alignment of the conserved domain regions. C8, domain of eight conserved cysteines; CTL, C-type lectin domain; EGF, epidermal growth factor-like domain; LDL, low-density lipoprotein receptor domain; R1, repeat region 1; R2, repeat region 2; R/K, lysine/arginine-rich repeat region; S, signal peptide; ShK, Stichodactyla helianthus toxin domain; TIL, trypsin inhibitor like domain; TSP-1, thrombospondin type 1 domain; vWD, von Willebrand factor type D domain; asterisk marks stop codon.

Biological Reviews 96 (2021) 1051–1075 © 2021 The Authors. Biological Reviews published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society.
makes it much harder to define differences between samples with confidence. Life stage and life cycle are also important biological factors that need to be considered (e.g. Bechtold et al. (2016) for plants; Breckman et al. (2015) for jellyfish). It is impossible to overstate the importance of temporal as well as spatial sampling variability in RNA-seq-based experiments, particularly for species with indirect development, or where adhesion only occurs occasionally and specific transcription is likely discontinuous.

Finally, the differences in expression profiles between organisms in their natural environment and those in the laboratory can be stark. For example, the expression of nectin, an adhesive protein candidate in the sea urchin, *P. lividus*, was 1.5 times higher in the tube feet of wild individuals compared to those in aquaria (Toubarro et al., 2016). Adhesion tenacity of *P. lividus* also decreased fourfold after aquarium acclimation (Santos & Flammang, 2007). Laboratory experiments may therefore alter the natural expression of adhesive candidates, perhaps due to the lack of hydrodynamic forces and additional stressors.

### (2) Sequencing depth, read counts and short-read assembly issues in adhesion research

In RNA-seq/transcriptomic analyses, one must be mindful of sequencing depth and read count abundance. There are no strictly defined criteria for identifying transcripts relating to adhesion, but it is common to select differentially expressed transcripts with high read counts based on the presumption that adhesive proteins are usually secreted in relatively large quantities. By contrast, and in particular where deep-sequencing has been conducted, many genes may be considered to be significantly differentially expressed even if expression levels are several orders of magnitude lower than the most abundant transcripts in the organism. These genes should be treated with caution as their scarcity may imply a less important role (if any) in adhesion compared to highly abundant transcripts, despite their significant differentiation. Determining their importance may be challenging due to lack of amplification in typical 35–40 cycle (quantitative) polymerase chain reaction (PCR) procedures. Likewise, WISH probe synthesis may be difficult and positive staining may fail due to a lack of sensitivity.

Another prevalent issue in RNA-seq-based studies is failure to achieve complete sequence length with common short-read sequencing platforms and assemblers. Many sequences will be fragmented, lacking complete 3’ and 3’ ends. Short-read sequencing and assemblers have been used in all adhesion-related transcriptome studies to date. However, with increasing availability of long-read sequencing and assemblers, researchers are now able to visualise the degree of repetition within adhesion-related conserved domains that has thwarted some short-read studies to date. Proteins involved in temporary adhesion are often large and highly repetitive (Hennebert et al., 2015a; Pjeta et al., 2019; Wunderer et al., 2019). In the flatworm *M. lignano* the protein Mlig-ap2 includes two repeat regions of 255 and 221 amino acids, which occur 21 times and 25 times, respectively (Wunderer et al., 2019). These repeats could not be assembled from short reads (Lengerer et al., 2018) and the size of the Mlig-ap2 protein (14784 amino acids) was only identified by studying the genome of *M. lignano* (Wasik et al., 2015).

In caddisfly larvae, some of the sericins and fibroins are large molecules containing repeats, which represent a challenge for sequence analysis. To investigate these proteins in their entirety, long-read PACBIO® sequencing data were combined with Illumina data to form a hybrid assembly for *Stenopsyche tinn furnace*, elucidating the first complete 21 kb H-fibroin gene sequence (Luo et al., 2018). This approach facilitated the production of ultra-long scaffolds that were polished using shorter Illumina reads to provide an improved, less-fragmented assembly. A full-length H-fibroin sequence was also assembled for *Parapsyche elsis* (estimated protein size: 658 KDa) using the MinION sequencing platform from Oxford Nanopore Technology (Frandsen et al., 2019), demonstrating the important role that long-read sequencing will have in the future of bioadhesion research.

### (3) Functional gene annotation – garbage in, garbage out

Considering the difficulties faced in short-read transcriptome assembly, and the opportunities provided by inter-phylum comparison of functional domains, it is sensible also to consider knock-on effects and systematic errors that may persist from assembly through to the annotation stage of analysis. How do assembly errors influence correct functional annotation? Is correct assignment of genes compromised, or biased by short-read assembly? Importantly, to begin with, most organisms used for bioadhesion research have only low to mediocre levels of functional annotation. Second, bioadhesion studies are almost by definition searching for proteins of previously unknown function. Third, functional annotation is typically reported with query sequence percentage similarity, e-value and bit-score from BLAST searches against sequence collections on the NCBI, Uniprot or other organism-specific databases. The same values are also reported for phylogenetic comparisons among transcripts or proteins. In both cases, the length of the sequence alignment is often omitted from reporting and this can be misleading. Two sequences may appear to share 100% similarity when, on closer inspection, this may only be true for a fragment of negligible length. The gene of interest may well be differentially expressed and involved in adhesion, but before extrapolations are made based upon its functional annotation, the quality and broader implications of that annotation should be considered.

The exponential increase in data availability through web portals and databases can further complicate matters, particularly in the absence of sufficient quality control. Peer reviews rarely analyse the underlying data in forensic detail and the majority of NCBI and Uniprot (e.g. TREMBL) databases remain to be reviewed or curated. With the publication of new assemblies and annotations daily, the growth of these databases can be deceptive, with large quantities of repeated
or redundant information. For example, in *E. pallida* (Davey et al., 2019), the most differentially expressed candidate sequence, AIPGENE2358, had the functional annotation Deleted in Malignant Brain Tumours 1 (DMBT1); an annotation the gene had acquired based on its similarity to a human sequence. Cnidarians do not possess brains, however they do produce abundant glycoproteins and, with further analysis of the literature, DMBT1 was found to encode a glycoprotein of unknown function in humans (Madsen, Mollenhauer & Holmsovk, 2010). Interestingly, a protein identified in limpet mucus, *P. vulgata*_10, was also annotated as DMBT1 (Kang et al., 2020). Yet, the three proteins are only superficially similar (Fig. 5). It is therefore advisable for researchers to treat annotations with caution, to be sceptical when interpreting biological meaning from their data and to determine if/how similarity was experimentally tested.

(4) Considerations when comparing transcriptome and proteome data sets

While RNA-seq is a powerful technique for quantitative determination of gene expression in adhesive-secreting tissues, the results remain hypothetical until proteins can be identified directly in the adhesive. It is the gene products, proteins, that are of central interest. This is where high-throughput proteomics can further support contemporary studies of bioadhesion. These methods require a suitable transcriptome reference for peptide prediction and identification. So, for species without a reference genome or appropriate transcriptome, *de novo* transcriptome assembly is first required. The NGS and MS technologies used to analyse mRNA and proteins are entirely different with respect to sample collection, scientific methodology and analyses. Most practical challenges fall on the side of MS. For a variety of reasons, not all proteins are detected by MS, although reverse and oppositely charged states may be used to increase the number of peptides identified. Some peptides may be heavily modified compared to the sequences predicted from mRNA. In addition, large portions of proteome can be inaccessible following digestion with a single protease (traditionally trypsin) and therefore may also require consecutive or parallel cleavages with multiple proteases (e.g. LysC, ArgC, AspN, GluC) to increase the number of identified proteins and peptides per protein, and consequently increase proteome sequence coverage (Swaney, Wenger & Coon, 2010). The central dogma of gene transcription and translation holds that up-regulation of a protein is directly proportional to up-regulation of the encoding gene, although this is widely accepted to deviate. Systematic delays, mRNA half-lives, rates of secretion, epigenetics, PTM and vesicular trafficking can all skew the relationship between transcription and protein secretion (Haider & Pal, 2013). If the same differential tissue samples are being used to study mRNA and proteins, there may be significant differences in turnover or storage times of proteins that prevent close alignment of RNA-seq and tandem mass spectrometry (MS/MS) data. Proteins may be accumulated or lost in the tissues of interest, relative to the mRNA copy number present at the same time. It is therefore no surprise that correlations between RNA-seq and quantitative proteomics data sets can be poor, as previously demonstrated in *E. pallida* (Cziesielski et al., 2018).

Again, an example can be illustrative. Davey et al. (2019) identified a candidate list of genes up-regulated in the pedal disc tissue of *E. pallida* and conducted an *in silico* analysis of potential secretion pathways for proteins of interest. However, MS/MS was not conducted and, therefore, none of the predicted proteins were quantified *in situ*. For further down-selection of candidate adhesion-related genes it would be useful to identify those whose protein products are present at the adhesive interface. MS/MS-based analysis of the secreted adhesive was therefore conducted (see Appendix S3; ftp://massive.ucsd.edu/MSV000086094) and the results are presented here for the first time. While the RNA-seq experiment was subtractive (whole animal *versus* animal with pedal disc dissected) to identify up-regulated genes in the tissue of interest (427 in total; Davey et al., 2019), MS/MS allowed direct analysis of proteins in the adhesive footprint (174 in total; Appendix S3). The first point to note from these new data is that the number of genes up-regulated in the pedal disc was ~2.5-fold greater than the number of proteins discovered in the footprint. Of the 174 proteins discovered in the footprint, only 13 matched up-regulated genes (Davey et al., 2019). Of those 13, around half were enzymes (Table 3). Whether these 13 proteins prove to be of outstanding interest, having been highlighted in both data sets (Fig. 6; Table 3), or instead represent coincidental overlap between proteins detected efficiently by MS/MS and those present in the differential transcriptome, remains to be seen. The comparison does, however, highlight the power of combined multi-omics for down-selecting candidates for further investigation, potentially reducing the pool from 427 to 13.

MS/MS techniques also face challenges of resolution that are not problematic for RNA-seq. While mRNA transcripts are composed of, and sequenced as, four bases, proteins are far more variable in their amino acid composition, size, charge, pH and side chains. Protein extraction, separation and preparation can be challenging and time-consuming. Hard-setting adhesives, for example, are often difficult to solubilise and may not provide truly representative spectra without complex and well-optimised methods (Schultzhaus et al., 2019). In addition, these analyses should consider stability of solubilised derivatives, reproducibility, precision of MS analyses and analytically valid recovery rates after digestion (Engel et al., 2021). RNA-seq protocols are much more uniform. Nevertheless, MS/MS maintains the advantage of providing direct evidence for the presence of a protein and, thus, plays an essential confirmatory role in the bioadhesives analysis pipeline (Fig. 2).

(5) Neglected and novel areas requiring greater insight

‘Omics currently occupy a central role in the identification of candidate adhesive proteins. While proteins may constitute
large (as in barnacles) or smaller proportions of the secreted adhesive (e.g. those that have highly glycosylated ‘mucoadhesives’), they are never the entire story. Beyond the biochemical methods that paved the way for understanding mussel adhesion, the identification of polysaccharide components, lipid components as well as cross-linking chemistry within secreted adhesives all presently require ‘traditional methods’, or access to ‘omics approaches that are not yet mature. Histological methods such as Alcian blue staining have demonstrated the presence of polysaccharides within the adhesives of several aquatic species (Jonker et al., 2012; Hennebert, Gregorowicz & Flammang, 2018; Clarke et al., 2020), however these polysaccharides have not been explored in structural detail. Similarly, lipids are often present in bioadhesive samples and have demonstrated roles in the adhesives of barnacles (cyprid and adult) and mussels (Gohad et al., 2014; He et al., 2018). Lipid-binding proteins have recently been identified in the adhesive glands of marine tube-building polychaetes (Buffet et al., 2018) and in cyprid and adult barnacles (Yan et al., 2020), reinforcing the likely importance of lipids in marine adhesives. Techniques suitable for lipidomic profiling include MS, high-performance liquid chromatography (HPLC) and NMR imaging (Cajka & Fiehn, 2016).

MS and HPLC techniques are also the primary tools for investigating the glycome of organisms (Ruedl et al., 2017). Online databases, such as Glycosciences.DB are designed to collect glycan structure data, models, and glycan moieties that can form a bridge between proteomic and glycomic resources (Bohm et al., 2019). In the bioadhesion field, adhesion-related glycans have been traditionally studied using specific stains (e.g. Alcian blue) or lectins (histochemistry, blotting and enzyme-linked assays). Glycoproteins have been implicated in attachment processes of numerous fouling organisms (Jonker et al., 2012; Hennebert et al., 2015b) and have received specific mention in earlier sections of this review. In general, however, glycans are particularly abundant in non-permanent adhesives, being often conjugated with proteins (N and O-glycosylations), but the nature of the attached glycan residues seems to vary considerably both intra- and inter-phylum (Simão et al., 2020).

Adhesion proteins from most taxa investigated to date contain PTMs, such as glycosylations, that are not directly apparent from basic transcriptomic and proteomic analyses. For example, those of C. intestinalis were found via lectin affinity to be shared between three evolutionarily distant ascidian species (C. intestinalis, Phallusia mammillata and Botryllus schlosseri), while another post-translationally modified amino acid, DOPA, may play an indirect role in the adhesion of C. intestinalis larvae (Zeng et al., 2019b).

Single-cell RNA sequencing (scRNA-seq) and epigenomics/epigenetics are two additional areas that have not yet appeared in the bioadhesion literature. scRNA-seq allows for more specific examination of the transcriptional profile of specific cell lines. Single-cell transcriptomics is often achieved through combinatorial barcoding during reverse transcription.

**Fig 5.** Conserved protein domains in proteins annotated as DMBT1 (Deleted in Malignant Brain Tumours 1). CUB, complement C1r/C1s, Uegf, Bmp1 domain; Gal-bind, galactose binding domain superfamily; SRCR, scavenger receptor cysteine-rich domain; vWD, Von Willebrand factor type D domain; ZP, zona pellucida-like domain.
PCR, with these added to single cells in oil droplets (Macosko et al., 2015) or small pools of cells (Cao et al., 2017). Sequencing data are then compared and clustered into cell types according to the similarity of transcriptomes (Kiselev, Andrews & Hemberg, 2019). Recent scRNA-seq data for C. robusta (Cao et al., 2019; Sharma, Wang & Stolfi, 2019) form a rich resource with which to compare more targeted differential transcriptomes and proteomes for adhesive organs or their adhesive secretions (U. Rothbächer, unpublished data). The scRNA-seq approach is compatible with fixed cells, thus minimising detrimental effects on the cell state. Such single-cell techniques have been successfully applied to the anemone Nematostella vectensis (Sebe-Pedros et al., 2018) which, although of little direct use to the study of bioadhesion (it is a sediment-dweller), provides a toolbox of techniques that could perhaps be applied to species that are bioadhesion-relevant, such as Exaiptasia pallida. While offering a lot of potential, single-cell transcriptomes will require sufficient spatial resolution and reliable markers if they are to discriminate specific cell types and reconstruct these into scRNA-based tissue clusters.

VI. CHALLENGES: HOW CAN BIOADHESION RESEARCH CONTINUE TO BENEFIT FROM ‘OMICs’?

In the previous sections we discussed contributions by ‘omics to our understanding of bioadhesion in longstanding species of interest, as well as the ‘head start’ provided by existing resources for more established model organisms and the potential pitfalls of the ‘omics approach. In this final section we look to the future and identify a short list of challenges in bioadhesion research that could be addressed using ‘omics approaches and that should be the focus of future research efforts.

(1) Investigating the evolutionary origin of adhesive proteins

The evolutionary origin of most adhesive proteins remains elusive. In the 2000s, because of the low number of adhesive protein sequences available for a very limited range of organisms, it was considered that most adhesive systems had evolved independently and that there was no evolutionary relatedness among adhesive proteins (Kamino, 2010). The few shared features, such as the occurrence of DOPA and phosphoserines in adhesive proteins of mussels and tube-worms, were assumed to be the result of convergent evolution. The increasing number of ‘omics data sets now allows comparison of common patterns in distantly related phyla, and may eventually help to identify ancient physiological processes from which adhesion derived. In some cases, these recurring themes are believed to have evolved independently (convergence); for example, the repetitive sequence encoded by exons 9a and 9b of Bombyx mori silk sericin exhibits

---

Table 3. Thirteen genes/proteins that were both significantly up-regulated in the mRNA sequencing data set and present in the proteomic dataset of Exaiptasia pallida

| Identifier       | Annotation                                                                 | Significance (−log10) | q-value | Difference | Test statistic |
|------------------|---------------------------------------------------------------------------|-----------------------|---------|------------|----------------|
| AIPGENE18684     | Predicted protein                                                          | 3.58                  | 0.01    | 3.07       | 12.14          |
| AIPGENE28816     | Blastoidea protease 10                                                    | 3.08                  | 0.02    | 2.66       | 9.01           |
| AIPGENE25980     | MAM and LDL-receptor class A domain-containing protein 1                  | 2.91                  | 0.02    | 2.79       | 8.18           |
| AIPGENE17639     | Semaphorin-5B                                                             | 2.77                  | 0.02    | 4.17       | 7.47           |
| AIPGENE20759     | Amiloride-sensitive amine oxidase [copper-containing]                     | 2.54                  | 0.03    | 1.33       | 6.49           |
| AIPGENE27733     | Tyrosinase                                                                | 2.53                  | 0.02    | 2.86       | 6.45           |
| AIPGENE17591     | Hemicentin-1                                                              | 2.49                  | 0.02    | 2.97       | 6.31           |
| AIPGENE24832     | PREDICTED: uncharacterized protein LOC593009                               | 2.02                  | 0.04    | 1.13       | 4.68           |
| AIPGENE2512      | Lysostaphin                                                               | 1.97                  | 0.04    | 2.69       | 4.53           |
| AIPGENE1623      | Golgi-associated plant pathogenesis-related protein 1                     | 1.95                  | 0.04    | 1.15       | 4.46           |
| AIPGENE22290     | Glutathione S-transferase                                                  | 1.92                  | 0.03    | 1.34       | 4.37           |
| AIPGENE3355      | A disintegrin and metalloproteinase with thrombospondin motifs 6          | 1.83                  | 0.04    | 1.48       | 4.11           |
| AIPGENE3458      | Tauropine dehydrogenase                                                   | 1.52                  | 0.05    | 1.40       | 3.30           |
remarkable similarity (35% identity over 600 amino acids) with the protein mlp-1 from the byssus of the blue mussel (Kludkiewicz et al., 2009). In others, clear evolutionary relationships are highlighted as is the case for the glycine/serine-rich barnacle cement proteins (AaCP19 and AaCP43) that share homologies with insect and spider silks (So et al., 2016). In the latter case, sequence information derived by ‘omics approaches logically reunited barnacles with other arthropods. In many cases, however, similarities among adhesive proteins are not obvious. Sequence conservation appears to be rare and limited to species of the same phylum. However, there are recurring characteristics of putative adhesive and cohesive proteins, like biased amino acid distribution, repetitive regions and frequently identified protein domains (Fig. 7). For example, the putative cohesion proteins Msig-ap1 (flatworm), Sfp1 (sea star) and P-vulgata_1 (limpet) all contain vWD, EGF, and lectin-binding domains, which are known to mediate protein–protein and protein–carbohydrate interactions (Hennebert et al., 2014; Wunderer et al., 2019; Kang et al., 2020). Conserved blocks of different domains might be indicative of common evolutionary origin. For example, evolution of an adhesive protein from a mucus ancestor with a similar conserved domain architecture was proposed by Pjeta et al. (2019). The association vWD–C8–TIL, sometimes repeated several times, is characteristic of a super-family of gel-forming secreted proteins that includes vertebrate mucins and von Willebrand factors, but also other proteins such as tectorin, zonadhesin, IgGFc-binding protein and SCO-spondin (Lang, Hansson & Samuelsson, 2007). Several adhesive proteins (e.g. Msig-ap1, Sfp1 and P-vulgata_3) from various groups of aquatic invertebrates have been annotated within this group and share the vWD–C8–TIL architecture (Fig. 7). The surface-binding adhesive protein Msig-ap2 shares many features of the glycoprotein SCO-spondin, including the vWD–C8–TIL motif, low-density lipoprotein receptor (LDL) domains and TSP-1 repeats (Wunderer et al., 2019). Two separate proteins (P-vulgata_3 and 6), resembling respectively the N- and C-terminal parts of Msig-ap2, have also been identified in limpets (Kang et al., 2020). A protein with comparable TSP-1 repeats was detected in tunicates (Li et al., 2019). Although unlikely to be coincidental, the function and relatedness of these conserved regions remains to be confirmed.

Another interesting example is SIPC, implicated in the temporary adhesive of barnacle cyprids, which shares the functional protein domains of the alpha-2-macroglobulin family (Dreanno et al., 2006). It was proposed that SIPC derived from a duplication of an ancestral alpha-2-macroglobulin and was functionally adapted for its role as a settlement cue and potential adhesive (Dreanno et al., 2006; Petrone et al., 2015). In recent years, proteins with similar domain structures have been identified in the adhesive secretions of diverse taxa, including echiurids Asterias rubens (Hennebert et al., 2015a; Lengerer et al., 2019) and Paracentrotus lividus (Pjeta et al., 2020), the limpet Patella vulgata (Kang et al., 2020) and the tunicate Ciona robusta (Li et al., 2019) (Fig. 7). Again, if not coincidental, these findings suggest that the functional adaptation of an alpha-2-macroglobulin-like protein to an adhesive happened either evolutionarily early, or multiple times.

(2) The importance of post-translational modifications

Current ‘omics techniques have a strong focus on proteins. However they do not, alone, provide any useful indication about mechanisms beyond the often-spurious annotation data for genes of interest. Specific PTMs may transform the function of proteins, in some cases reducing the protein to a support for functional glycosylations or other PTMs. The conversion of tyrosine to DOPA and phosphorylation of serine residues have well-documented importance in the adhesion of mussels [DOPA and phosphoserine (Waite, 2017)], tube-dwelling polychaetes [DOPA and phosphoserine [Jensen & Morse, 1988; Stewart et al., 2004]] and sea cucumber Cuvierian tubules [phosphoserine (Flammang et al., 2009)]. Apart from a few examples, however, identification and characterisation methods for PTMs are not well developed and pre-omics methods still contribute substantially.

Prediction of PTMs from ‘omics data sets is a significant challenge. For successful analysis, the purity, abundance and intactness of PTMs is crucial. Often extraction, fractionation and ionisation steps of a MS protocol can lead to varying degrees of PTM degradation (Breitwieser & Colinge, 2013). It has, however, been recently touted that novel proteome shotgun sequencing methods, combining ultrafiltration with limited tryptic proteolysis (FLiP; Xiong et al., 2020), could facilitate high-throughput identification of modification sites on proteins.

Another way in which ‘omics methods can help in the understanding of PTMs is by providing access to the suite of enzymes involved in the synthesis of these modifications (Waite, 2017). For example, by performing a homology search against the mussel foot transcriptome using tyrosinase sequences from a variety of species, Guerette et al. (2013) identified several tyrosinase candidates that could be involved in the conversion of tyrosine to DOPA in the mussel P. viridis. Tyrosinases have since been detected in transcriptomes and proteomes of other mussels (Qin et al., 2016), tube-worms (Buffet et al., 2018), and sea anemones (Wang et al., 2020; see also Table 3). Using a similar approach, Wang, Suhre & Scheibl (2019) retrieved the sequence of a polyphenol oxidase in MytiBase (Venier et al., 2009). This strategy could also work for kinases (phosphorylation) or glycosyl transferases (glycosylation), and help to reconstruct the biosynthetic pathway of adhesive proteins.

PTMs have been found to be of particular importance in the formation of aquatic silks (Sinohara, 1979), where a conserved characteristic is O-glycosylation (in which a mono- or oligosaccharide is attached to the hydroxyl group of a serine or threonine residue). O-glycosylation is also characteristic of the aqueous sticky proteins that cover the silk fibres of the web produced by orb spiders (Sinohara, 1977; Tillinghast &
Several PTMs were detected in the H-fibrin of the caddisfly Hydrophilus, such as methylations and phosphorylations (L. Rouhova, unpublished data). In another caddisfly, Brachycentrus echo, more than half of the serine residues in the H-fibrin are predicted to be phosphorylated. These phosphorylated serines could contribute to silk fibre periodic substructure through Ca$^{2+}$ cross-bridging and would be an adaptation of caddisfly larval silks to aquatic habitats (Stewart & Wang, 2010). The peripheral adhesive coating of the caddisfly larva Hesperophylax occidentalis contains negatively charged glycoproteins that likely contribute to underwater adhesion (Engster, 1976; Stewart, Ransom & Hlady, 2011) and a peroxidase enzyme (peroxinectin) that catalyses covalent dityrosine cross-linking to exterior polyphenolic compounds (e.g. humic acid), known to coat the silk fibre and surfaces under natural conditions. Thus, peroxidase-mediated cross-linking may be responsible for linking the coating to the fibre core, stabilising both against the solubilising power of liquid water and to the surface (Wang et al., 2014; Wang et al., 2015). It is evident that if we are to understand aquatic bioadhesion, we must increase the range of protein modifications that we are able to identify and analyse.

**Fig 7.** Conserved protein domains of potential adhesive and cohesive proteins in different taxa. A2M, alpha-2-macroglobulin; A2M BRD, alpha-2-macroglobulin bait region domain; A2M rctp-bd, alpha-2-macroglobulin receptor-binding domain; A2M TED, alpha-macroglobulin-like TED domain; C8, complement component C8; CKCT, CUB, complement C1r/C1s, Uegf, Bmp1 domain; EGF, epidermal growth factor domain; FA58C, coagulation factor 5/8 C-terminal domain; FAMeT, farnesoic acid O-methyl transferase domain; LDL, Low-density lipoprotein domain; lectin, lectin-binding domain; NTR, netrin domain; ShKT, Stichodactyla toxin domain; TIL, trypsin inhibitor-like cysteine-rich domain; TSP-1, thrombospondin type-1 domain; vWD, Von Willebrand factor type D domain.
(3) Adhesive gene/protein validation

It is likely that the increasing availability of short- and long-read sequencing will accelerate the discovery of putative adhesive proteins. However, knowing the sequence of a gene/protein is only part of the puzzle. To confirm the adhesive role, functional genomic studies such as morpholinolino- and RNAi-mediated gene knockdown, or TALEN- and CRISPR/Cas9-mediated gene knockout must be performed.

In bioadhesion research, non-adhesive phenotypes have only been achieved in the small flatworms, M. lignano (Lengerer et al., 2014, 2018; Wunderer et al., 2019); M. poznanense (Section IV.2) and M. ileanae (Pjeta et al., 2019). Although effective for flatworms, the following potential challenges must be considered when using functional genomic techniques to study bioadhesion: (i) well-established protocols are rarely available for non-model organisms. (ii) RNAi is transient and although it can be applied to different life stages, overcoming problems with potential lethal effects during development (non-adhesive larvae will most likely not metamorphose and/or reach adulthood), it must be stressed that it is never 100% efficient. Delivering a sufficient amount of double stranded RNA, small interfering RNA or small hairpin RNA to the target tissue requires extensive optimisation. (iii) Gene editing is time consuming, requires a well annotated genome (only available for a few adhesive model organisms) and animals that can be cultured or at least grown to the targeted stage within laboratory conditions. (iv) Bioadhesives usually result from a mixture of proteins, thus knocking down or knocking out a single protein might not affect adhesion. Multiple-gene silencing/editing approaches are often inefficient and increase the risk of off-target effects.

Transgenic platforms can build upon ‘omics data to generate hypotheses for experimental testing. They can then be used for functional analyses (e.g. to identify the function of metal ions) within bioadhesives, or for biotechnology applications such as experimentally tuning material properties. In transgenic silkworms, for example, over-expression of an ion-transporting protein increased Ca\(^{2+}\) transport out of the anterior silk gland, which both increased \(\alpha\)-helix and \(\beta\)-sheet conformations and reduced \(\text{Ca}^{2+}\) content of silks, all of which enhanced some material properties including tenacity and extension of fibres (Wang et al., 2015). Alternatively, the sequences obtained via ‘omics approaches, can be used to express the proteins of interest recombinantly and test their presumed adhesive role using complementary analytical techniques. This approach has been successfully applied to adhesive proteins of mussels (Hwang et al., 2004; Hwang, Gim & Cha, 2005; Lee et al., 2008; Choi et al., 2011, 2012), barnacles (Liang et al., 2015, 2018; Tilbury et al., 2019) and echinoderms (Lefèvre et al., 2020).

Finally, there are innumerable avenues for biochemistry-based confirmation of hypotheses generated from ‘omics data. These vary depending on the intended targets, but when selected appropriately they can provide valuable support to chosen lines of investigation. For example, having identified a suite of oxidases in a basal tissue proteome from the barnacle Balanus amphirhoe, So et al. (2017) were able to demonstrate oxidase activity in vitro and in vivo using colorimetric assays, confirming the activity of ketone- and aldehyde-forming oxidases at the barnacle adhesive interface. Targeted removal of metals and disruption of imine bonds in the glue of a terrestrial slug, Arion subfuscus, identified their pivotal roles in adhesion (Braun et al., 2013), and enzymatic hydrolysis of carbohydrates confirmed the double-network nature of the material (Wilks et al., 2015). Other analytical methods such as MS/MS (Stewart & Wang, 2010), \(^{31}\)P NMR spectroscopy (Addison et al., 2013) and attenuated total reflection-FT-IR (Ashton & Stewart, 2015) have been employed to identify and quantify the phosphorylation of serine residues within caddisfly silk proteins, for example. While ‘omics approaches are often considered agnostic to any pre-existing knowledge of the system under investigation, or ‘blind’, these analytical methods are bespoke to the hypotheses being tested.

VII. CONCLUSIONS

(1) Developments in the ‘omics have enabled researchers to interrogate living systems with a scope and resolution that were not possible previously.

(2) In bioadhesion research, the most ambitious studies of the 1990s focused on small numbers of proteins or genes, whereas those of the 2010s often began with data sets capturing large, if not complete, populations of genes or proteins.

(3) This has been particularly advantageous in the study of biological adhesion where (i) the majority of materials of interest are protein based, and (ii) where tissues that produce adhesive proteins can often be separated from the rest of the organism for differential analysis.

(4) Such analyses have accelerated understanding in organisms that have traditionally been the subject of bioadhesion studies, such as mussels and barnacles, but also facilitated the introduction of new study organisms with substantial ‘omics resources.

(5) The body of data now available for this extended suite of study organisms has highlighted consistencies between unrelated taxa that point either to considerable convergent evolution, or retention of specific adhesive traits through protracted periods of evolutionary history.

(6) Knowledge of these features will better target future studies to understand natural adhesion mechanisms and incorporate these concepts into new technologies, once the remaining technical barriers have been overcome.

VIII. ACKNOWLEDGEMENTS

The authors gratefully acknowledge European COST Action CA15216: The European Network of Bioadhesion
Adhesion by aquatic invertebrates

Expertise for facilitating our productive collaborations. N.A. acknowledges Leverhulme Trust Research Project Grant RPG-2017-017 to the Universities of Newcastle and Essex for supporting generation of the novel Exaiptasia data set presented herein. B.L. is funded by a Schrödinger Fellowship of the Austrian Science Fund (FWF): [J-4071]. M.Z. thanks Czech Republic MSMT Funding Programme INTER-COST no. LTC17073. P.F. is Research Director for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex for supporting generation of the novel data set. A. Aubry, S. Addison, J. B. Essex.
Doe, H., Mao, Y., Li, Y., Li, Y., Dai, X., Zhang, X., Liang, P., Liu, W., Wang, S. & Bao, Z. (2018). Characterization of chitin Euchionothitoides haldemensis based on transcriptional sequencing. Journal of Ocean University of China 17, 632–640.

Deiwick, C., Matsukura, K., Oshima, N., Takio, K., Hirota, H., Kibe, R. R. & Clare, A. S. (2006). An a-maltoglobin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite. Proceedings of the National Academy of Sciences of the United States of America 103, 14396–14401.

El-Gebeily, S., Mistri, J., Bayram, A., Edney, S. N., Luciani, A., Potter, S. C., Quayle, A. J., Salazar, V. J., Salazar, L. F., Sonnhammer, E. L. L., Hirsch, L., Paladini, L., Piovesan, D., Tosatto, S. C. E., et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Research 47, D127–D132.

Endrizzi, B. J. & Stewart, R. J. (2009). Gue, H.Y., He, Y., Eum, J.-H., Grabherr, M.G. (1976). Studies on silk secretion in the trichoptera (Frandsen, P.B., Federle, W., Gantayet, A., Farkas,R., Biological Reviews 1072 Peter A. Davey et al. Orihuela, B. (2011). Full-length transcriptome assembly from RNA-Seq data without Rasˇka, I., Mauceli, E., Kirby, R. R. Current Genomics 2013). Experimental strategies for the identification and characterization of adhesive proteins in animals: a review. Interface Focus 5, 20140064.

Eum, J.-H., Grabherr, M. G., He, Y., Yi, H., Liu, W., Wang, S. & Bao, Z. (2018). Characterization of chitin Euchionothitoides haldemensis based on transcriptional sequencing. Journal of Ocean University of China 17, 632–640.

Engel, B., Suppan, J., Nürnbergker, S., Power, A. M. & Marchetti-Deschmann, M. (2021). Revisiting amino acid analyses for bioadhesives including a direct comparison of tick attachment cement (Demodera marginata) and barnacle cement (Lopas anserum). International Journal of Adhesion and Adhesives 105, 102798.

Engelhardt, M. S. (1976). Studies on silk secretion in the trichoptera (Lepismatidodea). I. histology, histochemistry, and ultrastructure of the silk glands. Journal of Morphology 150, 183–211.

Eum, J.-H., Yoe, S.-M., Seo, Y.-R., Kang, S.-W. & Han, S.-S. (2005). Characterization of a novel repetitive secretory protein specifically expressed in the muscular salivary gland of Euphytopus sp. (Tichopthryphidae, Hepaticoptera). Insect Biochemistry and Molecular Biology 35, 435–441.

Farkas, R., Datková, Z., Mentelová, L., Lów, P., Beňová-Linžerová, D., Benš, M., Sass, M., Hruška, P., Huřková, H., Raška, O., Smigová, J., Raška, I. & Mechler, B. M. (2014). Apatite secretion in Drosophilidae salivary glands: a subcellular origin, dynamics, and identification of secretory proteins. PLoS One 9, e94381.

Fears, K., Barnick, A., Wassick, A., Ryov, H., Schultzzhau, J. N., Orihuela, B., Scancella, J. M., So, C. R., Hunsucker, K. Z., Leary, D. H., Swain, G., Ritschroth, D., Spillmann, C. M. & Wath, E. J. (2019). Adhesion of acorn barnacles on surface active borate ceramics. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20190203.

Federle, W. & Labonte, D. (2019). Dynamic biological adhesion: mechanisms for controlling attachment during locomotion. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20190209.

Fischer, C. T., Wurzel, O., De Hoog, T., Kravarek, K. M. & Reddwen, P. W. (2018). Cell type transcriptome atlas for the planarian Schmidtea mediterranea. Science 360, eaax1736.

Flamming, P., Lambert, A., Baille, P. & Hennébert, E. (2009). Polyphosphate-containing marine adhesives. The Journal of Adhesion 85, 447–464.

Foulon, V., Artigaud, S., Bescaglia, M., Bernay, B., Fabroux, C., Petton, B., Eilies, P., Bourkerka, M., Kelloi, G., Guerard, F. & Heck, A. J. (2016). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. FEBS Journal 283, 1377–1391.

Friedman, N., Aberle, D., Pires, E., Varela Cheillo, A., Grunwald, I. & Power, A. M. (2014). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e98892.

Fukushima, A., Kimura, M., Riedtke, H., Aruta, M. & Satô, K. (2009). Integrative omics approaches in plant systems biology. Current Opinion in Plant Biology 13(5–6), 532–538.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.

Gantayet, A., Maruyama, T., Arita, M. & Piovesan, D. (2009). Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. PLoS One 9, e94381.
Adhesion by aquatic invertebrates

Lee, S. J., Han, Y. H., Nam, B. H., Kim, Y. O. & Reves, P. (2008). A novel expression system for recombinant marine mussel adhesive protein Melp1 using a truncated Ompl4 signal peptide. *Molecules and Cells* 26, 34–40.

Lengerer, B., Flammang, P., Beisel, C., Minder, M. B., Scherler, T., Humenik, M., Leclercq, M., Surin, M., Tafforeau, L., Watthiez, R., Leclère, P. & Hennebert, E. (2020). Sea-star-inspired recombinant adhesive proteins self-assemble and adsorb on surfaces in aqueous environments to form cytoocompatible coatings. *Acta Biomaterialia* S1742-7061, S0306–3038.

Lengerer, B., Flammang, P., Lengerer, B., Hennebert, E. & Flammang, P. (2019). Interspecies comparison of sea star footprint proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374, 20190195.

Mohanram, H., Kumar, A., Verma, C. S., Persuhvish, K. & Miserez, A. (2019). Three-dimensional structure of *Algalabina rus* cement protein 20 revealed by multidimensional NMR and molecular dynamics simulations. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374, 20190198.

Nakano, M. & Kamin, K. (2015). Aminyl-like conformation and interaction for the self-assembly in barnacle underwater cement. *Biochemistry* 54, 826–835.

Nakano, M., Shen, J.-R. & Kamin, K. (2007). Self-assembling peptide inspired by a barnacle underwater *Algalabina rus* molecule. *Biochemistry* 46, 8390–8395.

Ohashi, K., Ito, M., Kamamori, M., Shigeno, Y., Uchiumi, T., Arai, R. & Tsukada, M. (2015). Characterization of silk gland rhodosomes from a bivalve caddisfly, *Styponympha mirabilis*; transmembrane suppressed production of a silk protein in cold conditions. *Biochemical and Biophysical Research Communications* 469, 210–216.

Ohashi, D., Kim, S., Lee, D. & Hwang, D. S. (2015). Tunican-mimetic nanofibrillar hydrogel adhesive with improved wet adhesion. *Acta Biomaterialia* 20, 104–112.

Paladino, M. L. B. & Bhunia, B. (2012). Bioadhesion: a review of concepts and applications. *Philosophical Transactions of the Royal Society A* 370, 2321–2347.

Palmowski, P., Watson, R., Europe-Finnen, G. N., Karoleczak-Bavati, M., Porter, A., Treumann, A. & Taggart, M. J. (2019). The generation of a comprehensive spectral library for the analysis of the Guinea pig proteome by SWATH-MS. *Proteomics* 19, 1900156.

Paredes, R. & Rothbard, J. (2014). Protein aggregation formed by recombinant H. sapiens truncated OmpA signal peptide. *Acta Biomaterialia* 13, 169–181.

Peeters, H. O., Höger, S. K., Looso, M., Lengfeld, T., Kuhn, A., Warnken, U., Nishishima-Fujisawa, C., Schönholzer, M., Krüger, M., Özbek, S. & Simakov, O. (2015). A comprehensive transcriptomic and proteomic analysis of hydra head regeneration. *Molecular Biology and Evolution* 32, 1928–1947.

Peeters, T. N., Brunak, S., Von Heijne, G. & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8, 785–786.

Petrie, T., Aldred, N., Emami, K., Enander, K., Edetth, T. & Clare, A. S. (2015). Chemistry-specific surface adsorption of the barnacle settlement-inducing protein complex. *Interface Focus* 5, 20140047.

Petit, D., de Mulder, K., Philip, I., Kuales, G., Hroda, M., Eichenberger, P., Borgonie, G., Hartenstein, V. & Laderuner, P. (2007). The exceptional stem cell system of *Macrostomum lignano*; sorting for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. *Frontiers in Zoology* 4, 9.

Phan, I. Y., Aldred, N., Ling, X. Y., Huskens, J., Clare, A. S. & Vancso, G. J. (2010). Atomic force microscopy of the morphology and mechanical behaviour of barnacle foot protein footprint at the nanoscale. *Journal of the Royal Society Interface* 7, 285–296.

Pjetja, R., Wunderer, J., Bertemes, P., Hofer, T., Salvenmoser, W., Lengerer, B., Coassin, S., Erhart, G., Beisel, C., Sobral, D., Kremer, L., Linder, H. & Curry, I. (2012). Characterization of barnacle underwater adhesive protein. *BMC Bioinformatics* 13, 20110047.

Pjetja, R., Foerster, F., Hess, M. W. & Laderuner, P. (2019). Temporary adhesion of the proseriate barnacle underwater adhesive protein. *BMC Bioinformatics* 20, 20110194.

Rittschof, D. (2021) 1051. *Molecular Biology and Evolution* 38, 448, 418.

Rittschof, D. (2021) 1051. *Molecular Biology and Evolution* 38, 448, 418.

Rittschof, D. (2021) 1051. *Molecular Biology and Evolution* 38, 448, 418.

Rittschof, D. (2021) 1051. *Molecular Biology and Evolution* 38, 448, 418.
Carbohydrate content of various silk proteins and proteomics. Bioinformatics 23, 247–248.

Strickland, M., Tudorica, V., Rézag, M., Thomas, N. R. & Goodacre, S. L. (2018). Conservation of a pH-sensitive structure in the C-terminal region of spider silk extends across the entire silk gene family. Heredity 120, 574–580.

Swaney, D. L., Wenger, C. D. & Coon, J. J. (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. Journal of Proteome Research 9, 1323–1329.

Tamarin, A., Lewis, P. & Askey, J. (1976). The structure and formation of the byssus attachment plaque in Mytilus. Journal of Morphology 149, 199–221.

Takai, H. (2010). Osmics-based medicine and systems pathology. Methods in Immunology 49, 173–185.

Tilburg, M. A., McCarthy, S., Domagalska, M., Enderle, T., Power, A. M. & Wall, J. G. (2019). The expression and characterization of recombinant cp191 barnacle cement protein from Pollicipes polliceps. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20180205.

Tillinghast, E. K. & Sinhohara, H. (1994). Carbohydrates associated with the orb web silk protein of Argiope aurantia. Biochemistry International 9, 315–317.

Tourado, D., Gouveia, A., Ribeiro, R. M., Simoes, N., Da Costa, G., Cordeiro, C. & Santos, R. (2016). Cloning, characterization, and expression of the Nectin gene from the tube feet of the sea urchin Paracentrotus lividus. Marine Biotechnology 18, 372–383.

Tyanova, S., Temu, T., Sintycyn, P., Carlsson, A., Hein, M. Y., Geiger, T., Mann, M. & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of proteome data. Nature Methods 13, 731–746.

Ueki, T., Kroke, K., Fukura, I. & Yamaguchi, N. (2018). Structural and mass spectrometric imaging analyses of adhered tunic and adhesive projections of solitary ascidians. Zoological Science 35, 553–547.

Ullano-Silva, M., Americo, J. A., Brindeiro, R., Dondero, F., Prondomici, F. & Rebelo, M. F. (2014). Gene discovery through transcriptome sequencing for the invasive mussel Littorina fortunei. PLoS One 9, e102973.

Urushida, Y., Nakano, M., Matsuda, S., Inoue, N., Kanai, S., Kitamura, N., Nishino, T. & Kamiño, K. (2007). Identification and functional characterization of a novel barnacle cement protein. FEBS Journal 274, 4336–4346.

Valois, E., Mirshafian, R. & Waite, J. H. (2020). Phase-dependent redox insulation in mussel adhesion. Science Advances 6, eaaz6496.

Vennier, P., de Pitti, C., Bernardi, F., Varotto, L., de Nardi, B., Bovo, G., Roč, P., Novoa, B., Figueras, A., Pallavicini, A. & Lanfranchi, G. (2015). Mytilase: a known release of mussel (M. galloprovincialis) transcript sequences. BMC Genomics 16, 70.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. & Gocayne, J. D. (2001). The sequence of the human genome. Science 291, 1304–1351.

Volbracht, F. & Tillinghast, E. K. (1991). Glycoprotein glue beneath a spider web’s aqueous coat. Naturwissenschaften 78, 557–559.

Waite, J. H. (1985). Catechol oxidase in the byssus of the common mussel. Journal of the Marine Biological Association of the United Kingdom 65, 339–371.

Waite, J. H. (2017). Mussel adhesion - essential footwear. Journal of Experimental Biology 220, 517–530.

Waite, J. H. (2019). Translational bioadhesion research: embracing biology without tokenism. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20190207.

Waite, J. H., Andersen, N. H., Jowharst, S. & Son, C. (2005). Mussel adhesion: finding the tricks worth mimicking. Journal of Adhesion 81, 297–317.

Waite, J. H. & Tanzer, M. (1981). Polyphosphonic substance of Mytilus edulis: novel adhesive containing L-DOPA and hydroxyproline. Science 212, 1030–1040.

Wang, C.-S., Ashton, N. N., Weiss, R. B. & Stewart, R. J. (2016). Periconoinctin catalysed dityrosine crosslinking in the adhesive underwater silk of a camouflading larval horse, Hypothenemus aequalis. Insect Biochemistry and Molecular Biology 54, 69–79.

Wang, J., Schreiber, M. H. & Scheibl, T. (2019). A mussel polysaccharide oxidase-like protein shows thio-meditated antioxidant activity. European Polymer Journal 113, 305–312.

Wang, L., Teng, L., Zhang, X., Liu, X., Lyu, Q., Yang, Y. & Liu, W. (2020). Discovery and characterization of tyrosinases from sea anemone pedal disc. Journal of Adhesion Science and Technology 34, 1840–1852. https://doi.org/10.1080/10290219.2020.1731271.

Wang, X., Zhao, P., Li, Y., Yu, Q., Ma, S., Xie, K., Chen, H. & Xia, Q. (2015). Modifying the mechanical properties of silk fiber by genetically disrupting the ion terminal environment for silk formation. Biomacromolecules 16, 3119–3125.

Wang, Y., Li, C., Du, J., Huang, J., Zhang, S. & Zhang, R. (2016). The modification, proteomics and crystallization of the limpet teeth. Proteins 84, 1800194.

Wang, Y., Wang, H., Zhao, T. & Nakagaki, M. (2010). Characterization of a cysteine-rich protein specifically expressed in the silk gland of a caddisfly Stenopsyche...
Adhesion by aquatic invertebrates

1075 © 2021 The Authors. Biological Reviews published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society.

Xu, W. & Wilks, A. M., Wilhelm, M. H. Wunderer, J. Wu, J. & Yonemura, N. (Biological Reviews 1075 Received 17 September 2020; revised 22 January 2021; accepted 26 January 2021; published online 16 February 2021)

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Conditional reciprocal best BLAST (CRBB) analysis of the Hydra vulgaris and Exaiptasia pallida data sets.

Appendix S2. A transcriptome-based analysis of adhesion proteins from two flatworms inhabiting freshwater and marine environments.

Appendix S3. Proteomic analysis of Exaiptasia pallida.

Movie S1. Non-adhesive flatworms.