Research Proposal

BIODIVERSITY AND BIOGEOGRAPHY OF POLYCHAETE WORMS (ANNELIDA): GLOBALLY AND IN WALLACEA

Proposed by:
Joko Pamungkas

Principal supervisor:
Dr. Mark J. Costello

Co-supervisor:
Dr. Christopher J. Glasby

INSTITUTE OF MARINE SCIENCE
FACULTY OF SCIENCE
UNIVERSITY OF AUCKLAND
2017
Table of Contents

1. Introduction ........................................................................................................... 3
   1.1 Background ..................................................................................................... 3
   1.2 Objectives ........................................................................................................ 5
2. Literature Review .................................................................................................. 6
   2.1 Polychaete taxonomy, biology and ecology ..................................................... 6
   2.2 Global species richness: how to estimate? ....................................................... 10
   2.3 Global distribution of marine species ............................................................... 11
   2.4 Indonesian polychaetes ..................................................................................... 13
   2.5 Indo-West Pacific and Wallacea ..................................................................... 15
3. Methodology .......................................................................................................... 18
   3.1 Estimating the global species richness of polychaetes ..................................... 18
   3.2 Mapping the global distribution of polychaetes ............................................... 18
   3.3 Check-list of Indonesian polychaetes ............................................................... 19
   3.4 Polychaetes of Wallacea ................................................................................... 19
      a. Study sites .................................................................................................... 20
      b. Specimen collection .................................................................................... 21
      c. Species identification ................................................................................... 21
      d. Statistical analysis ....................................................................................... 21
   3.5 Describing a new polychaete species ............................................................... 22
      a. Identifying a new polychaete species .......................................................... 22
      b. Assigning voucher specimens ..................................................................... 22
      c. Describing and naming a new polychaete species ....................................... 22
4. Research Outputs .................................................................................................. 23
5. Timetable .............................................................................................................. 24
6. Funding Sources .................................................................................................... 26
References .................................................................................................................. 27
Appendices ................................................................................................................ 34
1. How many polychaete species (Annelida) exist on Earth? (draft) .................... 34
2. Steps to download the polychaete dataset from OBIS ....................................... 58
3. Steps to download the polychaete dataset from GBIF ....................................... 58
   collected from a mangrove habitat on Ambon Island, Indonesia
   (published paper) ................................................................................................ 59
1. Introduction

1.1. Background

The question of how many species exist on Earth is always intriguing and has been leading to long debates among scientists. This is because the answer of that question reflects how much we do and do not know about life on the planet. Also, it forms a fundamental basis for all subsequent biodiversity-based research, and is a prerequisite to know how many of them are threatened with extinction from today’s major issues: the global warming and anthropogenic activities impact.

Scientists have different approaches to estimate the approximate number of global species with varying assumptions (e.g. Hammond, 1994; May, 1994; Raeka-Kudla, 1996). One of the best approaches may be to include past species discovery rates complemented by taxonomic effort (Costello & Wilson, 2010). Using this approach, Costello et al. (2012) have predicted the global species richness to be around 1.8 to 2.0 million species, of which approximately 300,000 are marine, whereas Appeltans et al. (2012) have estimated the number of global marine species to be about 0.5 ± 0.2 million, of which between one-third and two-thirds have been described by taxonomists.

Using similar methods, other scientists have also investigated the global species richness of specific taxa of interest, such as parasites (Costello, 2016; Poulin, 2014), marine fishes (Eschmeyer et al., 2010), amphipod crustaceans (Coleman, 2015) and molluscs (Bouchet et al., 2016). Despite these studies, there has not been a comprehensive review of the global species richness of polychaete worms (phylum Annelida, class Polychaeta). To date, taxonomic reviews of the animals have been regional, such as the study of polychaetes of Southern Africa (Day, 1967), of the Philippines and China Seas (Salazar-Vallejo et al., 2014), of South China Sea (Glasby et al. 2016), of the seas surrounding the Arabian Peninsula (Wehe & Fiege, 2002) and a chronological review of polychaete taxonomy in New Zealand (Glasby & Read, 1998).

Although Appeltans et al. (2012) and Costello et al. (2012) have studied the species richness of the world, the number of polychaete species could not be estimated using such a statistical model because at the time its discovery rate – i.e. the cumulative number of species described over time – did not seem to be reaching an asymptote. Both studies also did not look specifically at the productivity of polychaete taxonomists, which then raises two questions: is the increasing rate of the species discovery due to a large “pool” of polychaete species remaining to be described; or is it because an increasing number of taxonomists are maintaining the rate (if this
is the case, the discovery rate of the taxon then might have followed the global trend that most of species have been discovered)? To answer these questions, a more detailed study of global polychaete species richness, which includes both the species discovery rate and taxonomic effort, is thus required.

The distribution of global species generally shows a spatial pattern called the latitudinal diversity gradient (LDG). That is, the increase in species richness that occurs from the poles to the tropics (e.g. Brayard et al., 2005; Gaston, 2000; Rohde, 1998). The global pattern of LGD is generally considered unimodal with a peak around the Equator (e.g. Gaston, 2000; Kaufman, 1995). However, Chaudhary et al. (2016) found that the distribution pattern of marine species was significantly bimodal with a dip in species richness near the Equator. This finding was currently argued by Fernandez & Marques (2017) that the bimodality is due to a sampling bias, but later Chaudhary et al. (2017) re-proved that it was not. With regard to this debate, a study of global distribution of all polychaete species will clarify one of these two existing paradigms as the study has not been done yet.

Polychaete species richness in some parts of the world remain poorly studied. Indonesia, geographically situated in Southeast Asia between the Indian and Pacific oceans, is one of them. Due to the absence of local taxonomists, most studies conducted by local scientists are ecological, but the identification of the animals is generally not verifiable as it is a common practice that specimens are not retained after the work has been published (pers. obs.). Collaboration with foreign scientists to improve the taxonomic knowledge of the taxon, on the other hand, is now getting more complicated as the government implements rigorous rules and regulations of both the research permit for foreign scientists and biological specimen exchange (pers. obs.).

In the country, the Siboga (1899-1900) and Snellius II (1984) Expeditions have yielded the most Indonesian polychaete specimens. Surprisingly, not one polychaete specimen from these expeditions can be found at the Museum Zoologicum Bogoriense (MZB), i.e. the only accredited zoological museum of the country established in 1894. Most of the polychaete materials obtained from the expeditions are deposited at the Naturalis Biodiversity Center in Leiden, Netherlands, many of which remain unidentified (Glasby & Al Hakim, 2017; ten Hove, pers. comm.). Over the period of more than one century (1907-2016), the MZB itself only has a small polychaete collection (i.e. about 200 lots – pers. obs.); and to date, there has not been an inventory of Indonesian polychaete species which reflects that too little attention has been paid to the taxon.
This condition is in contrast to the fact that part of the “Coral Triangle” (CT), which is recognized as the marine biodiversity hotspot of the world (Asaad et al., 2017), lies in Indonesia. The CT includes Wallacea, which contains many endemic terrestrial species (Michaux, 2010). Such a taxonomic study conducted in this region is thus likely to yield a great number of polychaete species, many of which may be new to science.

Based on the above information and research gaps, several research questions can be raised as follows.

1. Can we estimate the approximate number of polychaete species living on Earth by using the current discovery rate? How many polychaete species have been described over the past 250 years? Is the trend in the polychaete species discovery rate now following the global trend?
2. How are polychaete species distributed globally? Which hemisphere and geographic regions have more species and why? Does the global distribution of the taxon show a uni- or bimodal pattern and what are the possible reasons?
3. How many Indonesian polychaete species have been formally described to date?
4. Is the species richness of Wallacean polychaetes higher than that of elsewhere and why? Is there significant variation in species richness within Wallacea? How many endemic species can be identified through the present study?
5. How many polychaete species new to science can be identified from the studied area?

1.2. Objectives

The proposed study is going to be the first study that:

1. reviews the global species richness of polychaete worms comprehensively;
2. maps the global distribution pattern of all polychaete species;
3. reviews the historical studies of Indonesian polychaetes;
4. reveals the species richness of Wallacean polychaetes and its connectivity with that of the adjoining areas;

In addition, it will provide descriptions of new Wallacean polychaete species.
2. Literature Review

2.1. Polychaete taxonomy, biology and ecology

Polychaetes are segmented marine worms belonging to the phylum Annelida. They are a disparate group of annelids encompassing several taxa that were once considered separate classes or phyla, including clitellates, siboglinids, echiurans, sipuncula and myzostomids (Struck et al., 2011; Weigert et al., 2014). The concept of polychaetes has now so expanded to be almost equivalent to that of Annelida. Yet, in the proposed study, the taxon Polychaeta is narrowly defined as a class as currently portrayed on WoRMS.

a. General morphology

Polychaetes exhibit high levels of morphological diversity. Yet, their body plan basically consists of three distinct regions: head, trunk and pygidium. The head region of the animals comprises two distinct parts, i.e. the prostomium, followed posteriorly by the peristomium. The trunk, which is situated between the head and pygidium, is composed of a number of body segments; each of which bears parapodia. Parapodia are the legs of the animals, which bear chaetae (the animals’ bristles). Both parapodia and chaetae play an important role in the animals’ body movement. The pygidium is the posterior end of the animals. It is non-segmental (usually conical or broadly rounded) and is the organ where the anus is located (Fauchald, 1977; Fitzhugh et al., 1997) (Figure 1).

Figure 1. Schematic body plan of a polychaete: entire worm (dorsal view) (A), anterior end (lateral view) (B), cross section through a segment (C). The figure is modified after (Fitzhugh et al., 1997).
b. **Habitat**

Polychaetes are mostly marine. Nevertheless, they can also be found in brackish water environment (e.g. ten Hove & Weerdenburg, 1978), including mangrove habitats (e.g. Metcalfe & Glasby, 2008; Nordhaus et al., 2009; Pamungkas, 2013). Some species also inhabit freshwater habitats (e.g. Glasby & Timm, 2008; Pamungkas, 2017). In the marine environment, the animals show a high tolerance towards temperature. As a result, they can be found from the coldest ocean temperatures of the abyssal plain (e.g. Alalykina, 2015) to the extremely high temperatures near hydrothermal vents (e.g. McHugh & Tunnicliffe, 1994).

c. **Life habits**

Adult polychaetes typically live as benthic infauna, i.e. aquatic animals that live in the water substrate. However, some species are planktonic (e.g. Marquez-Rojas et al., 2013). Based on the development of anterior appendages and life habits, the animals are separated into two main subclasses, i.e. Errantia and Sedentaria (Day, 1967). The subclass Errantia basically contains species that are free-moving (i.e. either crawlers or active swimmers), whereas members of the subclass Sedentaria are either burrowers or tube dwellers. Recently, the Echiura and other major taxa have also been assigned to the class Polychaeta (World Polychaeta Database), based on the findings of recent studies in annelid phylogeny.

As benthic organisms, polychaetes are also found to live in symbiosis with other marine creatures. For example, syllids live in symbiosis with ascidians, cnidarians, crustaceans, echinoderms and sponges (Martin & Britayev, 1998; López et al., 2001), whereas polynoids are symbiotic with anthozoans, echinoderms, molluscs and other polychaetes (Britayev, 1991; Britayev & Zamysliak, 1996; Britayev et al., 2007).

d. **Feeding modes**

Polychaetes show several different modes for acquiring food (Fauchald & Jumars, 1979; Jumars et al., 2015). For instance, they can be deposit feeders, i.e. those that ingest organic matter from the substrate around them, either actively or selectively. Active deposit feeders (characterized as having either pointed or rounded prostomium) consume the sediment directly and excrete the undigested materials, whereas selective deposit feeders (characterized as having numerous elongated tentacles) collect only the desired particles from the sea floor’s surface. Many burrowing polychaete species such as arenicolids, capitellids and lumbrinerids are active deposit feeders, whereas some
sedentary polychaetes like the “spaghetti” or “medusa” worms (family Terebellidae) are selective deposit feeders.

Some polychaete species are also suspension feeders. That is, those that feed on particles in the water column, either actively (i.e. by moving the water through a filtering apparatus) or passively (i.e. by catching suspended matter that fall on them). Active suspension feeders are also called ‘filter feeders’. The “fan” worms (families Sabellidae and Fabriciidae), as well as “feather duster” worms (family Serpulidae), are excellent examples of filter feeders as they have a highly modified head region called ‘branchial crown’ to actively filter the food (Fauchald & Jumars, 1979; Jumars et al., 2015).

Polychaetes can also be predatory. Some species feed on other polychaetes, whereas some others consume small crustaceans (Fauchald & Jumars, 1979; Jumars et al., 2015). Interestingly, the polychaetes called ‘bobbit worms’ (family Eunicidae) are also capable of consuming fishes and octopuses (a documentary video showing this is provided by the Smithsonian Channel at https://www.youtube.com/watch?v=K_7ByiYbCYM). Polychaetes, in turn, are usually eaten by other benthic organisms such as molluscs and crustaceans, as well as fish.

e. Reproduction strategies

Polychaetes display a wide range of reproductive strategies, but in principal they reproduce both asexual (Oliver, 1984) and sexually (Wilson, 1991). In asexual reproduction, a polychaete can reproduce through binary fission. Initially, the polychaete body separates into anterior and posterior halves, then the anterior half regenerates by forming posterior segments and pygidium, while the posterior half regenerates by forming anterior segments and the head region. The animals can also reproduce by budding, i.e. an asexual reproduction mode in which clones bud off the adult, detach and eventually crawl away. The same individual is often able to reproduce sexually.

The most common sexual reproduction in polychaetes is spawning (external fertilisation) since the animals are generally dioecious (yet, some species are hermaphrodites and others may be males at particular times of their life and then become females later). When spawning, most polychaete species release eggs and sperm into the water column, where fertilization takes place. The fertilized eggs (zygote) will then develop into trochophores, i.e. planktonic ciliated polychaete larvae. Depending on the species, trochophores can spend from hours to weeks in the water column before they
metamorphose into juveniles. Juveniles will then settle to the sea floor, where they continue to grow to maturity as benthic organisms. In some species, eggs, embryos and trochophores are brooded by female individuals (Wilson, 1991).

A remarkable sexual reproduction strategy of the worms is called ‘epitoky’. Through this mode, a polychaete modifies its sexually-mature adult body into a swimming body called ‘epitoke’ (e.g. Caspers, 1984; Chatelain et al., 2008; Pamungkas & Glasby, 2015). Triggered by a lunar cycle (Bentley et al., 2001), the epitoke will swarm in the water column and perform the so-called ‘mating dance’ to shed its gametes into the water column. Afterwards, it dies.

Depending on the species, two main strategies of epitoky are employed. That is, either the whole, or only the posterior part of body metamorphoses into an epitoke. The former case (also called ‘epigamy’) is common for many nereidids (e.g. Chatelain et al., 2008; Pamungkas & Glasby, 2015) and some syllids (e.g. Fischer & Fischer, 1995). The latter case (also called ‘schizogamy’) is common for some members of Eunicidae – e.g. palolo (Caspers, 1984) and wawo (Pamungkas, 2015) worms – and Syllidae.

f. Ecological and Economic Importance

Polychaetes can be the most abundant taxon in benthic samples, both in terms of the number of species and individuals. The animals, as a result, have been used extensively as indicators of the general “health” of benthic communities in response to pollution (Dean, 2008).

Particular polychaete species have also been used as indicator species of marine conditions. For example, Capitella capitata (family Capitellidae) (Figure 2A) and Malacoceros fuliginosus (family Spionidae) have been found to be dominant in high organic situations (Pearson & Rosenberg, 1978). Additionally, Cirriformia luxuriosa (family Cirratulidae) is tolerant of high Cu concentrations, whereas the nereidid Hediste diversicolor and the polynoid Halosydna johnsoni are Hg tolerant (Reish & Gerlinger, 1997).

Humans commonly utilize polychaetes as fishing bait and feed in mariculture (e Costa et al., 2000; Olive, 1999). Uniquely, in certain places of the world, wawo (Figure 2B) and palolo worms are consumed by locals (e.g. Caspers, 1984; Martens et al., 1995; Pamungkas, 2011) – the video of the process in which the worms are caught and cooked
by the Ambonese can be viewed at https://youtu.be/6fTjiTivt8U, whereas the detailed traditional recipe of wawo dish is provided by Pamungkas (2011).

Figure 2. Capitella capitata (A) and wawo worms (B). The former species (scale bar = 0.5 mm) is commonly used as a bioindicator of marine pollution. The latter worms are consumed by the Ambonese. The photographs (A and B) are modified after Blake (2009) and Pamungkas (2015 a), respectively.

2.2 Global species richness: how to estimate?
Scientists have various approaches to estimate the number of species that might exist on Earth. Based on reviews of Hammond (1994), May (1994) and Raeka-Kudla (1996), Costello & Wilson (2010) categorized the approaches into four ways. That is, estimating the species richness based on: (1) expert opinions, (2) species-habitat relationships using the island biogeographic theory, (3) body size-species richness relationships, and (4) past species discovery rates.

While expert estimates are subjective, the relationship between species and habitats might be weak, and body sizes could have no relationship with species richness, the past rate of species discovery seems to be the best method to predict the number of species. The rate has been used by Costello et al. (2012) to estimate the global species richness, including both marine and terrestrial species. The most notable finding of this study is probably that it has now become more difficult for taxonomists to discover new species because two-thirds of the world’s species – including the most conspicuous ones – have already been discovered from 1.8 to 2.0 species that might exist on the planet. Similar to this study, Appeltans et al. (2012) have predicted that 500,000 ± 200,000 marine species exist on Earth, of which between one-third and two-thirds may have been described.

Studies of marine species richness have also been conducted regionally (i.e. in European seas) by Costello et al. (1996, 2006) and Costello & Wilson (2011). Some taxonomists have also studied the discovery rate of various taxa of interest such as parasites (Costello, 2016;
Poulin, 2014), marine fishes (Eschmeyer et al., 2010), amphipod crustaceans (Coleman, 2015) and Molluscs (Bouchet et al., 2016).

2.3. Global distribution of marine species

Global species are spatially distributed in such a pattern called the latitudinal diversity gradient (LDG) (e.g. Brayard et al., 2005; Gaston, 2000; Rohde, 1998). That is, the increase in species richness that occurs from the poles (high latitude) to the tropics (low latitude). The tropical “belt” has thus been recognized as the zone with the highest species richness on Earth (Hoeksema, 2007; Stevens, 1989).

The global pattern of LGD is generally considered unimodal with a peak around the Equator (Figure 3A), and seems to be the case for almost all terrestrial plants and animals (e.g. Gaston, 2000; Kaufman, 1995). However, the pattern does not always apply for marine taxa. Although the distribution pattern of some marine taxa appears to be unimodal (e.g. Fautin et al., 2013; Hillebrand, 2004), other taxa show a bimodal pattern (e.g. Brayard et al., 2005; Saeedi et al., 2017). Recently, based on 27 published studies and the analysis of global datasets of 65,000 recent and 50,000 fossil marine species, Chaudhary et al. (2016) found that marine species distribution was significantly bimodal with a dip in species richness near the Equator (Figure 3B). They further explained that one possible reason for this is that regions around the Equator might have been too hot for some marine species due to the global warming. Consequently, those species decline at the equator. Arguing Fernandez & Marques (2017), Chaudhary et al. (2017) have also re-examined a more recent dataset and proved that the bimodality is not due to sampling bias.

There are numerous hypotheses explaining possible mechanisms underlying the LDG. One of the most prominent ones is the mid-domain effect (MDE), which asserts that if the species’ latitudinal ranges were randomly shuffled between geographic boundaries (domains), they would tend to overlap at the center of the domain, causing a peak of species richness in this particular area (Colwell et al., 2004). Another hypothesis is the Rapoport’s rule, stating that the latitudinal range of species is generally smaller at lower than higher latitudes. This possibly has facilitated more species to coexist in the tropics, leading to extremely high biodiversity in that area (Stevens, 1989).

Costello & Chaudhary (2017) proposed three driving forces that have shaped the distribution pattern of global species. That is, temperature, productivity and habitat complexity. Most of species on Earth are centered in the tropics as the geographic regions have warmer temperatures
and a more stable climate which lead to higher levels of speciation and lower levels of extinction (Buzas et al., 2002; Gaston, 2000; Rosenzweig, 1992). Areas in higher latitudes, by contrast, have high climatic variability where only species with broad environmental tolerances can survive (Stevens, 1989). The tropics also receive optimal solar irradiation used by primary producers to produce energy via photosynthesis, which results in higher net primary productivity. As a result, the regions have more species to support; and in marine areas, the number of species typically decreases from coastal to open seas, as well as from the open ocean to deep sea environment, following the gradient of primary productivity (Costello & Chaudhary, 2017).

The Island Biogeography Theory, postulated by Wilson & MacArthur (1967), also explains why some islands have more species than others. It states that the number of species found on an island is determined by immigration and extinction rates, both of which are affected by the distance of the island from its mainland. The theory also suggests that larger islands are generally more species-rich than smaller ones since they provide more heterogeneous habitats for species to colonize. Thus, an island that is bigger and closer to its mainland, based on the theory, tend to have a higher biodiversity than a relatively small and isolated island. An isolated island, besides usually having a fewer number of species, is typically inhabited by particular species being unique to that place called ‘endemic species’ (Anderson, 1994). Endemism is thus another key factor that determine the distribution of species.

Some hypotheses explaining the distribution of global species are evolutionary. Ricklefs (2004), for example, suggested that most of the land surface on Earth was either tropical or subtropical during the Tertiary, which could explain the greater biodiversity in the tropics today as an outcome of geological processes. Another hypothesis is proposed by Clarke & Cramer (2003), theorizing that the biodiversity in the temperate regions have not yet reached equilibrium because in former times the species in these regions had insufficient time to colonize or recolonize areas due to historical perturbations. The number of species in these regions, therefore, might continue to increase until saturated.
2.4. Indonesian polychaetes

The very first study of polychaete worms in Indonesia was probably conducted by Georg Eberhard Rumphius (1627-1702), i.e. a German-born naturalist who worked for the Dutch United East Indian Company, based in Ambon, Province of Maluku, eastern Indonesia. In his book entitled ‘D’Amboinsche Rariteitkamer’, he wrote in detail about the swarming wawo worms, which he named ‘Vermiculi Marini’ (Rumphius, 1705). Nevertheless, the name is not considered valid as it is assigned to several different species of wawo.

The species of wawo worms was later formally described by Horst (1905) who named it ‘Lysidice oele’ (family Eunicidae). Nearly one century after Horst, Martens et al. (1995) discovered 13 species of wawo. These include L. oele and – interestingly – the palolo worms Palola viridis, i.e. the typical polychaete species that swarms in the Samoan waters. Later on, Pamungkas (2015) identified 25 species of wawo worms (5 families), while Pamungkas & Glasby (2015) documented additional species of Nereididae.

The Dutch Siboga Expedition (1899-1900), with its primary aim to study the deep Indonesian basins, might be one of the most important expeditions that has yielded polychaete materials. The errant polychaetes were mainly studied by Horst (1903, 1912, 1913, 1917, 1918, 1919, 1921, 1924), whereas the sedentary species were mainly studied by Mesnil & Fauvel (1939) and Caullery (1944). Besides, a number of polychaete taxonomists described the Indonesian polychaete materials collected during the expedition based on their specific taxa of interest. For example, Aguado et al. (2008), Caullery (1914), Hutchings & McRae (1993) and Southward (1961) studied the syllids, siboglinids, aphroditids and pogonophores, respectively. In addition,
Bleeker & van der Spoel (1992) compiled published polychaete species collected during the Siboga Expedition in a catalogue.

Another expedition in Indonesian waters is the Snellius II. Carried out by the Dutch RV ‘Tyro and the Indonesian RV ‘Samudera’ in 1984, the voyage yielded some additional polychaete material. The syllids were studied by Aguado et al. (2008). Afterwards, the taxonomic studies of Indonesian polychaetes were very rare and tended to be patchy, such as the polychaetes of the Natuna Islands (Province of Riau Islands) by Al-Hakim & Glasby (2004); the description of two capitellid species collected from mangrove habitats in Cilacap (Province of Central Java) and Ambon (Province of Maluku) by Pamungkas (2015b, 2017).

The biodiversity of the worms has also been investigated in many ecological studies carried out by both local university students and scientists. Nevertheless, it is a common practice in the country that specimens are disposed after the work has been published as most institutions – except the Research Center for Oceanography in Jakarta and the Research Center for Deep Sea in Ambon – do not have special space to archive them (pers. obs.). As a result, it is generally difficult or even almost impossible to confirm the validity of species identification by comparison with voucher specimens. On the other hand, the only accredited Indonesian zoological museum (MZB) only has a small collection of polychaetes – i.e. about 200 lots collected from 1907 to 2016 (Figure 4) – most of which are unpublished and remain unidentified due to the absence of polychaete curators (pers. obs.).

Figure 4. Polychaete collection at MZB (A) and its submissions of specimen lots per decade (B).
2.5. Indo-West Pacific and Wallacea

On Earth, there are some biogeographic regions with exceptionally high marine biological diversity. All are situated in the tropical “belt”, i.e. the Indo-West Pacific (IWP) and some parts of the East Pacific (EP), West Atlantic (WA) and East Atlantic (EA) (Hoeksema, 2007). In the majority, the marine biodiversity in those regions is higher in coastal zones due to higher productivity (Costello & Chaudary, 2017).

The IWP, which extends from the Red Sea and East Africa to the Central Pacific, is recognized as the center of maximum marine biodiversity due to the occurrence of coral reefs which provide favourable habitats for marine life (e.g. Ekman, 1934; Briggs, 1974; Paulay, 1997). In the IWP, the area with extraordinary marine biodiversity is centered at the so-called ‘Coral Triangle’ (CT). That is, a marine area located in the western Pacific Ocean including the waters of Indonesia, Malaysia, the Philippines, Papua New Guinea, Timor Leste and Solomon Islands (Figure 5). In this area, 75% of the world’s coral species, i.e. nearly 600 species (Spalding et al., 2001), and over 2,000 reef fish species can be found (WWF, 2017). The CT is also a home of six of seven marine turtle species that exist on Earth. To date, the CT is believed as the hotspot of global marine biodiversity (Asaad et al., 2017).

In relation to this, there are four models attempting to explain the origin of high marine species richness in the tropics. The first one is known as the Center of Speciation (Briggs, 2003), which suggests that two biodiversity hotspots in the tropics – i.e. the CT and Caribbean Sea – have become the wellsprings of marine species from where the species radiate out. The converse of this model is the Center of Accumulation (Jokiel & Martinelli, 1992; Connolly et al., 2003), which states that peripheral habitats such as oceanic archipelagos are the provenance of marine species from where the species flow out and accumulate in the two biodiversity hotspots due to prevailing currents. A third model is called the Center of Overlap (Gaither et al., 2011), which suggests that isolated species of the Indian and Pacific Oceans are co-distributed in the area of overlap. A fourth model called the Biodiversity Feedback (Rocha et al., 2008) incorporates elements of the three previous models (Figure 5).
Wallacea is a biogeographical designation for an area situated between the Asian Sunda Shelf (Sundaland) and the Australian Sahul Shelf (Sahulland) proposed by Dickerson (1928). It comprises a group of almost all of Indonesian islands including nine relatively big islands, i.e. Sulawesi, Lombok, Sumbawa, Flores, Sumba, Timor, Halmahera, Buru, Seram and many smaller islands (Figure 6). Administratively, those islands belong to ten provinces, i.e. North Sulawesi, Gorontalo, Central Sulawesi, West Sulawesi, South Sulawesi, South East Sulawesi, West Nusa Tenggara, East Nusa Tenggara, North Maluku and Maluku. Additionally, Timor Leste – situated on Timor Island and formerly was part of Indonesia – is also part of Wallacea. The area, which has a total area of 347,000 km², is separated from the Asian and Australian continental shelves by deep water straits. It is also separated by two imaginary lines called Wallace’s and Lydekker’s Lines, which separate the region from Sundaland and Sahulland, respectively (Lohman et al., 2011) (Figure 6).

It was Alfred Russel Wallace, a British naturalist, who initially proposed the Line in 1858. The Line, which has become a boundary that demarcates the distribution range of both Asian and Australian terrestrial fauna, was later named as ‘Wallace’s Line’ by Huxley (1868). By this invisible Line, the spatial distribution range of Asian fauna seems to be limited to the eastern edge of the Sunda Shelf. That is, between Kalimantan (Borneo) and Sulawesi (in northern area), and between Bali and Lombok (in southern area) (Figure 6). Crossing this Line, a striking faunal transition can be seen: the species composition is more Australian despite some overlaps (New, 2002; Wallace, 1869). Wallacea is therefore such a transitional zone between Asian and Australian terrestrial fauna (New, 2002) despite the fact that the region is also such a “home” to many endemic species (Michaux, 2010).
Figure 6. Wallacea. The region is situated between Sunda and Sahul Shelves. The map is modified after Lohman et al. (2011).
3. Methodology

3.1. Estimating the global species richness of polychaetes

The study of global species richness and discovery rate of polychaete worms has been conducted during the provisional year (Appendix 1). The study was based on a dataset containing 21,104 polychaete species records extracted from WoRMS (Read & Fauchald, 2016), downloaded on October 10th, 2016. WoRMS is a taxonomically authoritative species information system that provides inventories of valid marine species names (yet, the system also includes information about synonymy). The content of WoRMS is controlled and kept up to date by a number of selected taxonomic experts (WoRMS, 2017).

To investigate the global species richness of polychaetes, records with species names that fulfill three criteria, i.e. (1) name status is ‘accepted’, (2) have been verified by taxonomic editors, and (3) not ‘fossil only’, were first counted. Following Costello et al. (2012), the number of species described per year, as well as the cumulative number of species described over time, was plotted to see the general trend of the species discovery. As an indicator of taxonomic effort, the annual number of authors, as well as the average number of species described per author over time, was plotted. A simple correlation (Costello et al., 2013) was run to determine if there is a significant break point between an increasing number of species described per author and a decreasing number of species described per author over time.

Furthermore, the likelihood if there has been a change in specialization effort by authors over time was examined by counting the proportion of the number of taxonomists describing one species only per decade. The Pearson’s skewness coefficient in the number of species described over time was also calculated. Lastly, the number of non-polychaete species described by the top 25 most prolific authors was counted to see if taxonomists are now more specialized than they were in the 19th century.

3.2. Mapping the global distribution of polychaetes

The study of global distribution of polychaete worms will be based on datasets that will be extracted from both the Global Biodiversity Information Facility (GBIF) and the Ocean Biogeographic Information System (OBIS) (Appendix 2 and 3, respectively). GBIF is a biodiversity information system that provides an open-source database of all types of life on Earth (GBIF, 2017), whereas OBIS is such a similar system that provides a marine species database only (OBIS, 2017). Both GBIF and OBIS datasets are garnered from various research institutions all around the globe by participant nodes. The datasets basically contain information
about *what*, *where* and *when* species have been recorded. In the proposed study, a new dataset on Indonesian polychaete species that will be obtained from the fieldwork will also be included.

Following the steps in Saeedi et al. (2017), the two datasets will be first merged using R version 3.4.0. Duplication, i.e. records with the same scientific name, coordinate and collection date, will then be removed. The nomenclature of the species names will be verified on WoRMS in order to reconcile synonyms and misspellings. Records that are classified as fossils, mapped on land, have location coordinates with no precision estimates, and with location precision that is more than 100 km will be removed. By following these, one dataset with both valid scientific names and coordinates will be obtained.

ArcGIS version 10.4.1 will be employed to map the spatial distribution pattern of all polychaete species from the available records. The number of species will be plotted against 5° latitudinal bands, and Kernel Density model will be used to test if the distribution pattern of the animals differs from a unimodal pattern (Wessa, 2012). A group-average clustering method will be further used to identify geographic regions of endemcity using PRIMER-E 6 (Clarke & Warwick, 2001; Linse et al., 2006; Belanger et al., 2012).

### 3.3. Check-list of Indonesian polychaetes

The historical study of Indonesian polychaetes will be based on all published literature containing polychaete species that have been collected from Indonesian waters. The literature will mainly include taxonomic and ecological publications whose: (1) species names listed are associated with voucher specimens, and/ or (2) at least a polychaete taxonomist is involved. Ecological papers that do not fulfil those two criteria will be marked as the species identification might be doubtful and not verifiable. Yet, they will also be taken into account in order to get the “big picture” of polychaete studies in the country. Following Glasby et al. (2016), the primary taxonomic unit will be *valid species* as currently indicated on WoRMS.

### 3.4. Polychaetes of Wallacea

The study on Wallacean polychaete species richness will be based on the polychaete material that is planned to be collected from Wallacea, Indonesia (Figure 7), from October 2017 to April 2018 (Table 3). The specified time is proposed due to relatively calm weather in most studied areas (pers. obs.).
a. Study sites

The study sites are selected based on, first, Asaad’s study on delineating priority areas for marine biodiversity conservation in the Coral Triangle (Asaad et al., 2017). In this study, eight sites with exceptionally high marine biodiversity situated in and outside Wallacea are proposed (Figure 7). The study sites are also selected by considering the presence of marine research stations belonging to the Indonesian Institute of Sciences (LIPI), the Ministry of Marine Affairs and Fisheries (KKP), the University of Hasanuddin and the University of Mulawarman (Table 1). The presence of research institutions nearby is of importance as laboratory work will be done at those institutions. The distance to the airport, availability of accommodation, as well as access to the sampling sites have been other considerations.

![Figure 7. Map of the study sites. Green colour indicates islands belonging to the Wallacea, whereas red dots and numbers indicate study sites. The map is modified after Hennemann et al. (2015).](image-url)

<table>
<thead>
<tr>
<th>Site</th>
<th>Location</th>
<th>Province</th>
<th>Research stations nearby</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sulawesi</td>
<td>North Sulawesi</td>
<td>Bitung Marine Life Conservation Unit</td>
</tr>
<tr>
<td>2</td>
<td>Ambon</td>
<td>Maluku</td>
<td>Research Center for Deep Sea</td>
</tr>
<tr>
<td>3</td>
<td>Kupang</td>
<td>East Nusa Tenggara</td>
<td>Research and Development Division for Marine Bio Industry</td>
</tr>
<tr>
<td>4</td>
<td>Spermonde Islands</td>
<td>South Sulawesi</td>
<td>University of Hasanuddin</td>
</tr>
<tr>
<td>5</td>
<td>Bali Island</td>
<td>Bali</td>
<td>Institut for Mariculture Research and Development</td>
</tr>
<tr>
<td>6</td>
<td>Derawan Islands</td>
<td>East Kalimantan</td>
<td>University of Mulawarman</td>
</tr>
<tr>
<td>7</td>
<td>Biak</td>
<td>Papua</td>
<td>Biak Marine Life Conservation Unit</td>
</tr>
<tr>
<td>8</td>
<td>Dobo</td>
<td>Maluku</td>
<td>Tual Marine Life Conservation Unit</td>
</tr>
</tbody>
</table>
b. Specimen collection
At each sampling site, samples will be collected from the intertidal zone of a reef habitat during a low tide. To obtain the worms, dead coral rubble will be smashed using a hammer and a chisel. The worms observed will be pulled out with forceps then moved into a jar filled with 85% ethanol. Each vial will be labelled using a small piece of tracing paper that is put inside. Smaller pieces of rock, which typically contain smaller worms, will be placed in a bucket with sea water and left for about two hours to deoxygenate the water. Thereafter, the water will be vigorously stirred then poured, passing a sieve with 0.5 mm mesh size. This sieving step will be repeated three times in order to get as many species as possible. Retained material on the screen of the sieve will be gently removed using a plastic spoon into a jar filled with 85% ethanol. Voucher specimens is planned to be deposited at MZB in Cibinong, Bogor, Indonesia, as well as at the Reference Collection LIPI Ambon (RCLA) and the Museum and Art Gallery of the Northern Territory (MAGNT) in Darwin, Australia.

c. Species identification
Polychaetes will be identified to morphological species units using available keys. To identify the polychaete specimens, morphological examinations will be conducted using both stereo and compound microscopes equipped with digital cameras. Line drawings will also be performed using camera lucida mounted to both microscopes. To produce quality photographs, Photoshop and Helicon Focus software will be used to clean the photograph backgrounds and stack photographs with different depth of field, respectively. Primary characters to identify the worms include the morphology of pro- and peristomium, parapodia, chaetae and pygidium (Figure 1). The identification process will take place at the research stations near the sampling sites (Table 1), as well as at the Research Center for Biology (i.e. the LIPI’s research station where MZB is based), and at MAGNT.

d. Statistical analysis
To study the similarity of polychaete taxonomic compositions in and outside Wallacea, a hierarchical cluster analysis will be performed using R version 3.4.0. Principally, the analysis will first compare the taxonomic compositions of all sampling sites. Thereafter, the taxonomic compositions will be grouped in such a way that those in the same group (called a ‘cluster’) are more similar to each other than to those in other groups.
A dendogram will be created to visually illustrate the arrangement of clusters produced by the analysis.

3.5. Describing a new polychaete species

A new Wallacean polychaete species, i.e. *Capitella ambonensis* Pamungkas, 2017, has been described based on the polychaete material collected from Ambon, Province of Maluku. The work has been published in *Zootaxa* (Appendix 4), and the species has been registered on WoRMS at http://www.marinespecies.org/aphia.php?p=taxdetails&id=955060. Main steps in describing the species were as follows.

a. Identifying a new polychaete species

First of all, the new polychaete species candidate was identified to family (Capitellidae) then genus (*Capitella*) level. Afterwards, it was further identified to species level using all available literature related to all existing species belonging to the same genus. The species was justified to be new to science because it shows distinct key characters from all other existing *Capitella* spp.

b. Assigning voucher specimens

Voucher specimens of *C. ambonensis* were selected and consist of one holotype and some paratypes. The holotype was deposited at MZB, whereas the paratypes were deposited at both MZB and RCLA.

c. Describing and naming a new polychaete species

*Capitella ambonensis* was first described based on its general morphological characters. The key characters that make the species distinguishable from other *Capitella* spp. were described. Macrophotographs and photomicrographs of the species, as well as scanning electron microscope (SEM) micrographs, were taken as part of the description. Line drawings were additionally performed to describe detailed features that could not be clearly observed using photographs. Methylene blue staining pattern (MBSP), modifying the protocol in Warren et al. (1994), was also applied to examine the similarity between the material of this study and other *Capitella* species. The scientific name of the species was assigned following the rules of the International Code of Zoological Nomenclature (ICZN, 1999). A table of identification keys to all *Capitella* species was provided.
4. Research Outputs

The following outputs are expected from the proposed study.

1. Reference collection of Wallacean polychaetes (to be deposited at MZB, MAGNT and RCLA).

2. Five published scientific articles that are planned to be the chapters of the thesis as follows.
   
   - Chapter 1: ‘How many polychaete species exist on Earth?’
   - Chapter 2: ‘Global distribution of polychaetes’
   - Chapter 3: ‘Check-List of Indonesian Polychaetes (Annelida)’
   - Chapter 4: ‘Polychaetes (Annelida) of Wallacea’
   - Chapter 5: ‘Capitella ambonensis: a new polychaete species (Annelida: Capitellidae) collected from a mangrove habitat on Ambon Island, Indonesia’
## 5. Timetable

Table 2. Doctoral study timetable. The study is proposed to be conducted within a period of three and a half years. Letters indicate months. Green shading indicates tasks that have been completed during the provisional year. Red shading indicates activities that will be conducted abroad. Black shading indicates activities that are supposed to be conducted in the specified month(s).

<table>
<thead>
<tr>
<th>Activity</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Administration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctoral Candidate Needs Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety risk assessment for field work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Workshops</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Integrity Module</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DELNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctoral Skills Induction Day</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Publications</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global polychaete species richness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global polychaete distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Check-list of Indonesian polychaetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polychaetes of Wallacea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A new Wallcean polychaete species</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Proposal &amp; thesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposal writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposal submission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thesis writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thesis submission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fieldwork &amp; data analyses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fieldwork</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAGNT visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R stat's learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ArcGIS learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Presentations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposal presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>International Polychaete Conference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thesis presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Fieldwork timetable. Numbers indicate weeks. Red shading indicates the arrival and departure in Indonesia and to New Zealand, respectively. Black shading indicates activities that are supposed to be conducted in the specified week(s).

<table>
<thead>
<tr>
<th>Activity</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>October</td>
<td>November</td>
</tr>
<tr>
<td>Arrival in Indonesia</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Preparation *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field trip to site 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-site **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At RCB/ RCO ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specimens packing ****</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Preparation includes sampling equipment preparation, flight, accommodation and local transport bookings.
** The initial samples processing and identification of the animals will be conducted at the research station nearby.
*** RCB, where MZB is based, will be the basis of the proposed study as the institution has the most suitable laboratory for taxonomic work.
**** Best specimens are planned to be brought to New Zealand for further identification.
6. Funding Sources

This research will be funded by the New Zealand ASEAN Scholarship (NZAS) and the University of Auckland. Primary expenditure will be associated with both field and laboratory work as both are planned to be conducted abroad, i.e. in Indonesia. Other expenditure will be related to the participation at the 13th International Polychaete Conference (IPC) which is planned to be held in California, USA, in 2019.
References


