University of Stavanger

Faculty of Science and Technology

Master's Thesis

Study program/Specialization:	Spring semester, 2020
Environmental Engineering/Water Science and Technology	Confidential
Writer: Aleksandra Szwedo	(Writer's signature) 

Faculty supervisor:

Professor Daniela Maria Pampanin

Thesis title:

The application of digital image analysis to quantify histological pathological changes in *Mytilus edulis* from the Water Column Monitoring program

Credits (ECTS): 30

Key words:	Pages:
Digital pathology, *Mytilus edulis*, histopathology, gonads, gills, QuPath	61 pages + 10 pages Appendices Stavanger, 15 June 2020

ABSTRACT

Use of histopathology to identify changes in marine species in response to contaminants is a long-known practice. The suitability of histopathological biological markers in the assessment of marine environment quality derives from their relatively high ecological relevance, as alterations on tissue level are often irreversible. Also, affecting the physiology of an individual could impose a potential effect on the whole population. Visual assessment, traditionally used in the scoring of histological parameters, is prone to bias, which is a major disadvantage. The advances in image acquisition and digital image analysis over the last decades created a suitable, dynamically evolving environment for analysis of high-content histological images. Digital pathology, taking advantage of computer vision and the high potential for automation, could address the limitations persistent in manual scoring methodology. In this study, we propose the application of digital image analysis in performing analysis of histological parameters in genital and gill tissues of *Mytilus edulis*. Mussels were collected as a part of Water Column Monitoring (2017) program from the rig stations in various proximities to the offshore platforms in Tampen area. Using QuPath, an open-source digital analysis software, specifically designed to handle whole slide images, we developed a method for gender recognition in Hematoxylin & Eosin (H&E) colored slides with >97% success rate. We performed a comprehensive analysis of male genital tissue with emphasis on gonad area and cells in various stages of gametogenesis. The comparison of our results with the available scores from the manual assessment of gonadal status and spawning stage failed to replicate the original classification scores. The differences between groups were significant for a subset of parameters. Stemming from the small size and narrow diversity our study sample, the analysis was limited. Using slides stained with Alcian Blue – Periodic Acid-Schiff, we quantified mucus cells in gill tissue and performed classification of mucus cells according to the properties of the stain. The comparison of our results with the original scoring of the lesion "Abnormal mucus secretion" did not reach significance. The manual scoring of the lesion was performed on H&E colored slides, which might explain the lack of correlation with our results. Lastly, we investigated the potential correlations between mucus cell quantity and concentration of various contaminants determined in mussels as a part of WCM 2017 program. The study showed that digital pathology could be a viable approach for quantification of histological parameters in *Mytilus edulis*, especially when used with lesion-specific staining. In the study, results obtained by digital image analysis did not always reflect the visual scoring, which might be due to bias in visual scoring or limitations of developed digital methods. In order to fully benefit from digital pathology, broad expertise in digital image analysis and histopathology, as well as data mining should be brought together.

ACKNOWLEDGEMENTS & CONTRIBUTIONS

I would like to express my gratitude to everyone that contributed to this work and without whom it could have never been completed. First and foremost I would like to thank my supervisor, Professor Daniela Maria Pampanin, for offering me participation in this project. Through the work on this thesis, I got to know Professor Daniela as versatile researcher, whose ambition and expertise is expressed through inspiration, advice and constructive criticism, which she provided along the path of completing the project.

I would like to thank Melinda Lillesand, who was a great companion during our training in image analysis software, Visiopharm. Melinda provided valuable insights from her histopathology and image analysis experience, which I hoped to put in best use in this thesis.

My sincere gratitude goes to Jodi Maple-Grødem, my PhD supervisor, for being supportive of continuation of my master degree alongside the doctoral position. Her true dedication to research has always served as motivation and inspiration to go forward.

I would like to say thank you to Abdelnour Alhourani, my colleague, best friend and fiancé, who has been of invaluable support and motivation in all aspects of the thesis work. I greatly appreciate his suggestions in image analysis and advice on data analysis, as well as continual readiness to help.

I would like to express my gratitude to Associate Professor Hanne Røland Hagland and Associate Professor Roald Kommedal, both of whom kept me motivated to pursue master thesis when facing a dilemma due to demands of my PhD work.

Lastly, I would like to express the appreciation for support from my friends, fellow students, colleagues and family, who never stopped encouraging me to pursue the project and research in general.

Sincere thanks to all of you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AB – Alcian blue

AB-PAS - Alcian Blue – Periodic Acid-Schiff

AI – Artificial intelligence

DAB – 3,3'-Diaminobenzidine

DF - degrees of freedom

FDA - Food and Drug Administration

GPUs - Graphical Processing Units

H&E – Haematoxylin & eosin

MILKYS – Miljøgifter i norske kystområder, Norwergian coastal environmental monitoring program

OD - Optical density

OSPAR – Convention for the Protection of the Marine Environment of the North-East Atlantic

PAHs - polyaromatic hydrocarbons

PAS – Periodic Acid-Schiff

PCBs - Polychlorinated biphenyls

RGB – Red Green Blue

SD - Standard deviation

WCM – Water Column Monitoring

WSI – Whole slide imaging

LIST OF TABLES

LIST OF FIGURES

LITERATURE REVIEW

INTRODUCTION

In recent years, growing trend towards studying causal link between exposure to pollution and alterations in marine organisms can be observed. Preference to include biomarkers in environmental pollution monitoring programs is driven by their capability to identify actual biological effects exerted by contaminants (Bignell et al., 2011). In this study, particular focus is given to histopathological biological markers. Their abundancy and variety in marine species are clear in the literature. Use of histopathology in environmental monitoring is well documented and confirms their reliability and sensitivity. Traditionally, the assessment of histopathological biological markers is performed according to scoring systems. However, literature resources indicate that different scoring systems are used. This negatively impacts the possibility to integrate and compare data, especially on an international scale. Therefore, efforts are made to unify the assessment methods. Combination of digital pathology and artificial intelligence (AI) offers digitalization and high degree of automation of the process (Tizhoosh and Pantanowitz, 2018). The application of digital image analysis could address the problem of scoring bias, commonly recognized in visual assessment by human eye and hard to resolve. Being a relatively novel approach, AI in digital pathology still awaits addressing several challenges that the method faces (Tizhoosh and Pantanowitz, 2018). Nevertheless, shift from human eye to computer vision in scoring systems carries big potential regarding automation and standardization. Moreover, AI use in environmental sciences could provide essential feedback about the methodology for its further implementation in medical diagnostics.

HISTOPATHOLOGICAL BIOMARKERS

The response to environmental contaminants in an individual organism or a more complex biological system offers valuable information on the effects of a contaminant on biological life (Hylland et al., 2017). A measurable response can be used as a biomarker of environmental exposure to contaminants (Aarab et al., 2011). Biomarkers provide information on the actual effects that contaminants exerts in living organisms on a variety of biological organization levels (Figure 1): molecular, biochemical, physiological, histo- and -cytopathological, organismal, population and community (Au, 2004; López-Galindo et al., 2010). In regard to environmental pollution monitoring, the relevance of biomarkers is highly determined by their abilities to predict the effect of a contaminant on a higher level of organization (i.e. population/community/ecosystem). In general, with increasing level of biological organization, responses become more indicative of relevant ecological effects, however, they might be less specific (Au, 2004; Beyer et al., 2013). In addition, responses at higher biological organization levels occur only after long preceding exposure to a contaminant and indicate already deteriorated

quality of the ecosystem (Figure 1). In contrary, responses on the low levels of biological organization, despite their high specificity, sensitivity and generally better reproducibility, are often insufficient to provide findings relevant for environmental monitoring purposes (Au, 2004).

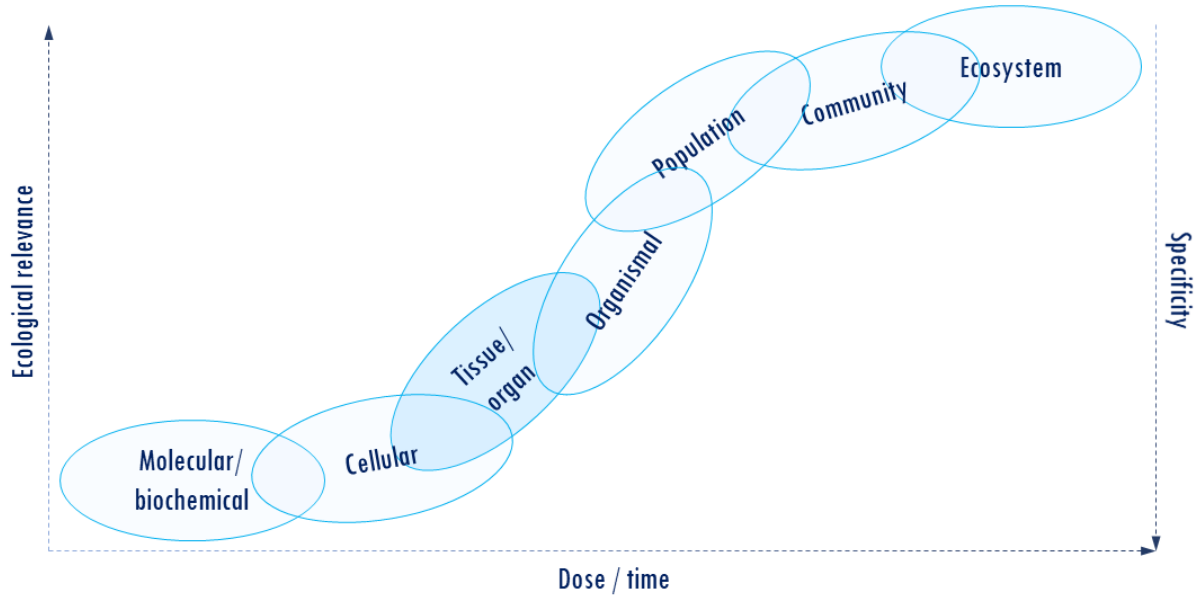

Figure 1 Levels of biological organisation concerning biomarker's specificity and ecological relevance (vertical axes), as well as dose/time relationship (horizontal axis). Level corresponding to biomarker in this study (histopathology) marked bright blue. Based on (Hansson et al., 2013).

Biomarkers of high potential in monitoring programs are characterized by: 1) high ecological relevance, 2) specificity, 3) responsiveness to the realistic concentrations of the pollutant in the environment and ideally 4) good dose-response relationship (Au, 2004). Biomarkers on a tissue level address the majority of those requirements. Representing a moderate level of biological organization, histopathological changes provide diverse and valuable information. Tissue alterations are a direct result of adverse changes on molecular, biochemical and physiological levels and therefore could occur in a relatively early stage of the exposure. An exclusive advantage of histopathology evaluations is the ability to identify the precise locations of tissue alterations and detect the presence of pathogens (Ben Ameur et al., 2012). Reaching the tissue level, alterations are often irreversible and have a potential to alter the physiology of the whole organism. The insight on physiological functions (e.g. nutrition, respiration, reproduction, growth) and presence of pathogens, reflective of the health of an individual organism, can be used to project on the higher levels of biological organization (Au, 2004; Bignell et al., 2008; Yancheva et al., 2016). Therefore, histopathological alterations can serve as a reference to compare with other types of biomarkers to provide a more informed assessment of biological effects in environmental monitoring (Aarab et al., 2008).

MYTILUS SPP. AS BIOINDICATOR ORGANISMS

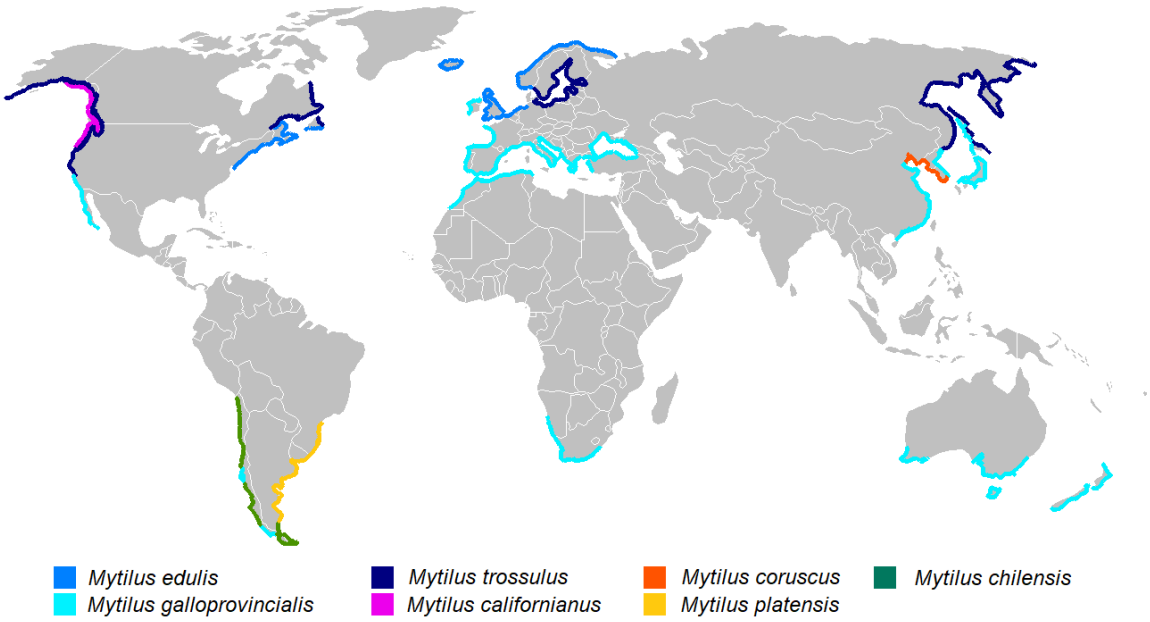

Figure 2 Global distribution of mussels of the genus *Mytilus*. Based on (Gaitán-Espitia et al., 2016; Gosling, 2015).

The suitability of mussels (*Mytilus* spp.) in environmental pollution monitoring is reflected by the application of the species in mussel watch programs, dating back to 1960s (Beyer et al., 2017). Originally dominated by the fish species, monitoring programs progressively incorporate the bivalve species as a part of their routine (Cuevas et al., 2015). In Norway, *Mytilus edulis* has been incorporated in the national coastal environmental program (MILKYS) since 1981 (Beyer et al., 2017). The utility of blue mussel (*Mytilus edulis*) as a sentinel organism derives from the biology and ecological attributes of the species (Table 1), and is in line with the criteria set by OSPAR commission for a species suitable for environmental monitoring (Beyer et al., 2017). *Mytilus* spp. are sessile, which makes them excellent sentinels to access the location-specific environmental quality. Due to their tolerance to handling, they can be employed in transplant caging studies, which allows better control over the experimental design in monitoring studies (Beyer et al., 2017). The use of caged mussels is often favored even where the native populations are abundant, as it offers a wider choice of spatial positioning, ease of sampling and, better control over the confounding factors (e.g. genetic differences between the subspecies, variation of health condition between individuals, non-uniform developmental and spawning stages, earlier exposure to contaminants), especially when combined with the use of farmed mussels (Beyer et al., 2017). Mussels are of reasonable size to provide sampling material for different types of analyses and are relatively easy to sample due to their tendency to form large mussel beds. *Mytilus edulis* can tolerate the typical environmental concentrations of contaminants. however, it demonstrates a sensitivity suitable for monitoring of the adverse biological effects (Aarab et al., 2008). Their utility as sentinel organisms

is also attributed to their filter-feeding mechanism and ability, contributing to bioconcentration and bioaccumulation of contaminants from the large volume of filtered seawater and the food particles. Lastly, the wide geographical distribution of the *Mytilus* spp. (Figure 2) makes mussel watch programs suitable for international monitoring programs, data exchange and comparison (Au, 2004; Beyer et al., 2017; Gaitán-Espitia et al., 2016).

Table 1 Characteristic of *Mytilus* spp. as suitable for environmental monitoring. Based on (Au, 2004; Beyer et al., 2017; Gaitán-Espitia et al., 2016).

Characteristic		Description
Biological	Sessile	Provide location-specific information Ease of sampling
	Medium-sized	Even small sample provides enough tissue material for the analysis
	Hardy	Ease of sampling and transportation; Can survive in laboratory conditions, allowing for culturing and validation exposure experiments in a controlled environment
	Filter-feeding	Accumulation of contaminants from the seawater by bioconcentration and bioaccumulation
	Sensitive	Respond to changes in the surrounding environment
	Tolerant	Can withstand concentrations of contaminants encountered in the marine environment without serious adverse effect on their condition
Ecological	Global distribution in temperate waters	Potential for international monitoring programs and exchange of data
	Form mussel beds	Ease of material collection

HISTOLOGICIAL PATHOLOGIES IN *MYTILUS EDULIS*

According to Bignell et al., histopathological parameters in health assessment can be linked to: 1) parasite and pathogen infections, 2) inflammation and non-specific pathologies and 3) reproductive or physiological status (Bignell et al., 2008). In pollutant exposure studies, commonly reported alterations are associated with morphological changes and inflammatory responses. The latter are also common in bacterial/viral infections. Although parasitic infections usually do not elicit inflammatory responses in mussels, they remain included in health assessment parameters (Bignell et al., 2011). According to Cuevas et al., parasitosis might lead to misinterpretation of histopathological indices, which emphasizes the importance of monitoring of parasite presence (Cuevas et al., 2015).

GILL TISSUES

In mussels, gills are a multifunctional organ, associated not only with respiration, but also filter-feeding process (Gosling, 2015). Gills are therefore a key organ in providing the entry for energetic substrates and gas exchange. Due to their physiology, gills constitute the main site of entry for the

contaminants in polluted seawater (Beyer et al., 2017), as illustrated in Figure 3. Increased exposure and elevated concentration of the contaminant in gill tissue derive from the high filtration capacity and the accumulation tendency (Au, 2004). The gill epithelium is one of the first-line defense against toxic compounds, which is reflected in many histopathological alterations identified in this tissue of the exposed individuals (Beyer et al., 2017). Some morphological changes in the gill epithelium are considered as early adaptations to toxic compounds, meant to hinder routes of their uptake or compensate for the impaired capacity of the feeding and gas exchange processes (Pagano et al., 2016). Uptake of contaminants in mussel's gills occurs mainly by passive diffusion and active transport through transmembrane ion pumps, and, to some extent, by endocytosis (Beyer et al., 2017). Uptake of hydrophobic organic contaminants, like polyaromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs), is believed to occur mainly through the external surfaces and the gut epithelium, following the equilibrium partitioning processes. According to Au et al., the gill histopathology can be indicative of general stress conditions due to exposure to metals, oil, pulp mill effluents, organic pollutants, toxic algae and suspended solids (Au, 2004). Examples of gill histopathological alterations in response to contaminants are described in **Feil! Fant ikke referansekilden.**. The most typical pathologies in the gill tissue are presented in Table 3. Despite a wide range of the histopathological alterations in gill tissue, similar pathologies seem to appear in a response to a variety of contaminants (Table 2). Therefore, gill tissue histopathology might be a good method to assess the general severity of environmental contamination, rather than contaminant-specific evaluation.

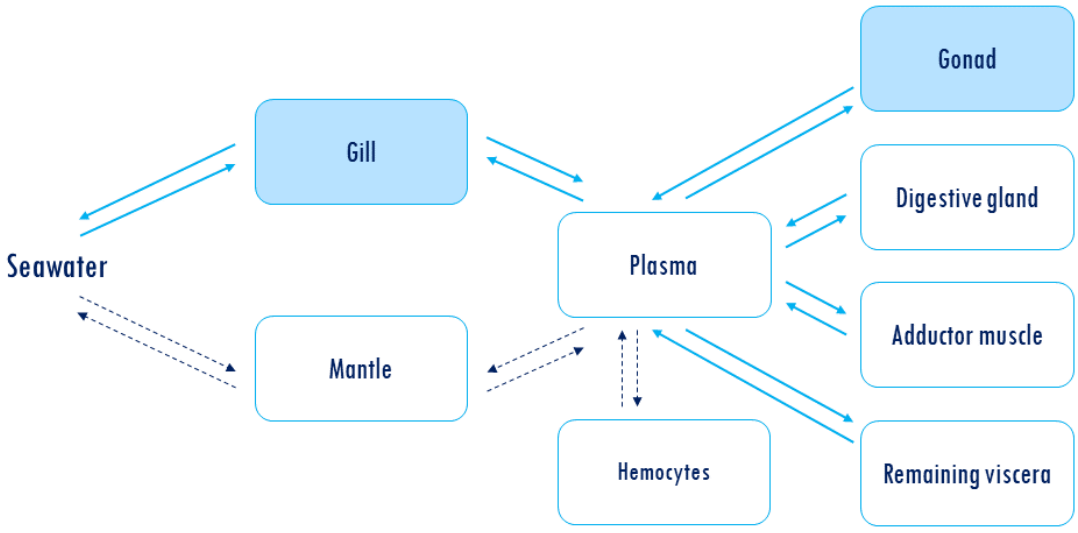

Figure 3 Schematic illustration of pathways for transport and distribution of contaminants in blue mussels. Bright blue, solid arrows indicate major pathways, dark dashed arrows – alternative pathways. Tissues of interest in this study are highlighted in blue. Reprinted from (Beyer et al., 2017)

Table 2 Examples of histopathological alterations in molluscs in response to various contaminants, reported in the literature.

Species	Contaminants/contaminants source	Sample source	Alterations	Reference
Mytilus galloprovincialis	Seawater with Commercial Diesel and dispersant agent (BIOVERSAL HC); (simulation of untreated oily wastewater)	Laboratory exposure, 15 days	Erosion of lateral cilia, heamocyte infiltration. Severe cellular alterations: disorganised microvilli layer, vaucolated cytoplasm, presence of mucous granules.	(Pirrone et al., 2018)
Mytilus galloprovincialis	Seawater with BIOVERSAL HC	Laboratory exposure, 15 days	Gill filament hyperplasia, heamocyte infiltration. Moderate cellular alterations: cytoplasm vacuolisation, epithelial disorganisation.	(Pirrone et al., 2018)
Mytilus galloprovincialis	Seawater with Commercial Diesel	Laboratory exposure, 15 days	Gill filament hyperplasia, heamocyte infiltration. Moderate cellular alterations: irregular nuclei shape, cytoplasm vacuolization, epithelial disorganisation.	(Pirrone et al., 2018)
Mytilus galloprovincialis	BF-MBR treated oily wastewater	Laboratory exposure, 15 days	Relatively normal morphology. Minor cellular aletrations: presence of mucous granules, cytoplasm vacuolisation, disorganisation of cilia and microvilli.	(Pirrone et al., 2018)
Mytilus galloprovincialis	Metal contamination (Cd, Pb, Zn)	Field experiment, 60 days exposure	Moderate heamocyte infiltration, minor gill filaments thinning, minor frontal and lateral cilia erosion.	(Rossi et al., 2016)
Mytilus galloprovincialis	Metal (As, Pb, Hg, Zn, Fe, Co, Cd, Cu) and C10-C50 hydrocarbons pollution	Field experiment, 60 days exposure	Hyperplasia, gill filaments thinning, frontal and lateral cilia erosion, heamocyte infiltration. Irregular nuclei shape, cytoplasm vacuolization.	(Rossi et al., 2016)
Mytilus galloprovincialis	Preservative agent, Quaternium-15 (hexamethylenetetramine chloroallyl chloride)	Laboratory exposure, 18 days	Massive melanin deposits in gill epithelium, loss of frontal and lateral cilia, enlarged heamolymph vessel, heamocyte infiltration. Increased mucous cells and mucous production with increasing pollutant concentration.	(Pagano et al., 2016)
Diplodon expansus	Atrazine (2-chloro-4ethylamino-6-isopropylamino-s-triazine), herbicide	Laboratory exposure, 7 days	Low concentration exposure: epithelial adherence, fusion of gill filaments, epithelial detachment and dilation of intercellular space, heamocyte presence in the lumen. High concentration exposure: odema, epitheliam detachment, loss of filament integrity, heamocyte infiltration.	(Nogarol et al., 2016)
Mytilus galloprovincialis	Sodium hypochlorite, NaClO	Laboratory exposure, 14 days	Allterated lipofuscin granules amount, presence of heamocytes in the lumen. Severe decrease in the amount and extent of cilia. Vacuolization of epithelial cells.	(López-Galindo et al., 2010)
Mytilus galloprovincialis	Mexel® 432 (alkylamino)-3-aminopropane	Laboratory exposure, 14 days	Alterated lipofuscin granules amount, presence of heamocytes in expanded lumen. Decrease in the amount and extent of cilia. Minor vacuolization of epithelial cells.	(López-Galindo et al., 2010)
Mytella falcata	Domestic sewage, steel mill plant discharge. High PAH concentration and metal content in the sediment.	Field sampling	Detachment of the gill epithelium, cell swelling in the intermediate zone. Heamocyte presence in the lumen and infiltration into epithelium. Rare hyperplasia. Increased amount of mucous cells.	(David et al., 2008)
Mytilus galloprovincialis	Silver nanoparticles	Laboratory exposure, 12 hours	Hyperplasia and hypertrophy of epithelium. Severe infiltration of heamocytes. Some detachment of epithelium. Loss of epithelial thickness.	(Bouallegui et al., 2017)

Table 3 Selected gill pathological changes in blue mussels.

Alteration category and name		Description/significance	Reference
Parasitic	Copepods	Reduced overall condition in intense infections.	(Bignell et al., 2008)
	Ancistrum mytili/unknown gill ciliates	Pear-shaped body, large macronucleus, small micronuclei. Reported in nearly 100% mussels. Generally, do not elicit host response.	(Bignell et al., 2008)
	Rickettsia /chlamydia-like organisms	Basophilic inclusion bodies present within epithelial cells. Large colonies can disrupt cell morphology and surrounding tissue architecture. Generally, do not elicit host response.	(Bignell et al., 2008)
	Gregarine	Pear-shaped parasite, eosinophilic cytoplasm, large macronucleus, small micronucleus. Generally, do not appear to elicit host response.	(Bignell et al., 2008)
Morphological	Detachment/lifting of gill epithelium; oedema	Related to presence of chemical contaminants. Caused reduction in gills surface and may impair the gas exchange process.	(Bouallegui et al., 2017; David et al., 2008)
	Increase in epithelium thickness; hyperplasia; hypertrophy	Increase in tissue volume by abnormal proliferation without cell volume change (hyperplasia) or abnormal cell volume/swelling (hypertrophy). May lead to hyperplasia of gill filaments, reduction of inter-lamellar space and fusion of lamellae. Increase in epithelium thickness may also result from cell swelling following the membrane permeability change.	(David et al., 2008; Santos et al., 2014)
	Loss of epithelium thickness	Can be associated with epithelium lifting. May indicate shortening of life cycle, apoptosis and inhibition of cell division. May be related to DNA damage and oxidative stress.	(Bouallegui et al., 2017)
	Increase in mucus cell number	Promotes higher mucus production. Mucus has important role in filter-feeding to ease capture of particulate matter. Mucus is also involved in pseudofeces formation. In some cases, increased mucus production is related to higher metabolic activity of mucus cells, rather than increase in their number. It can be detected using TEM, as presence of smaller and more abundant mitochondria. Abnormal mucus secretion can be relatively easy observed with naked eye while sampling.	(Bouallegui et al., 2017; David et al., 2008)
	Loss of microvilli on epithelial cells	Indicates either damage of the cell following absorption of undesired substances or absorption prevention by decrease in contact area.	(David et al., 2008)
	Loss/erosion of cilia	Lack of cilia indicates onset of gill exfoliation, which progressively leads to impaired food-filtering and gas exchange which adversely affects survival capacity of the organism.	(Pagano et al., 2016; Rossi et al., 2016)
	Increase in mitochondrial number	Possibly an adaptive response to compensate for lower ATP production in damaged tissues, necessary to conduct gas exchange.	(David et al., 2008)
	Skeletal rod differentiation	Thickening and differentiation of skeletal rod. Possibly an adaptation to prevent gill filaments separation and disorganization of gill structure.	(David et al., 2008)
Inflammatory	Haemocytes in haemolymph vessel/invasion into epithelium	May lead to deformation of gill filament. Considered as manifestation of defense mechanisms linked to increased cell turnover (reabsorption of damaged or old epithelial cells) in these regions. Some describe cell invasion as moderate alteration, possible to reverse when removing the stressor.	(David et al., 2008)
	Haemocytic neoplasia/ disseminated neoplasia	Proliferation of abnormal haemocytes. Increased infiltration, which may lead to the whole space between lamellae filled with haemocytes. Impaired defense mechanisms. HCN is progressive and leads to mortality. Progressive proliferation of haemocytes causes displacement, compression and gill necrosis.	(Bignell et al., 2008; Gosling, 2015)
	Melanosis	Abnormal deposits of melanin. Indicates inflammatory activity of pro-phenoloxidase linked to recognition, cytotoxicity and encapsulation of foreign bodies in the tissue.	(Pagano et al., 2016)

GONADS

Gonads in mussel are the main tissue of concern when assessing the reproductive condition of the individual (Bignell et al., 2008). Gonad developmental stages are typically a subject to seasonal variation, which might be a confounding factor when assessing of the mussel's condition in a histopathology-based manner (Cuevas et al., 2015). During a spawning event, mussel's biomass might decrease by up to 40%, representing a major physiological challenge. It was observed that the developmental stage itself can be subjected to other factors, like the relative position of an individual in a mussel bed or food availability (Beyer et al., 2017). Guidelines for assessment of the gonad developmental stages and the spawning stages are as described by Seed (Seed, 1969; Seed and Brown, 1977) and are presented in Table 4 and Table 5, respectively. Schematic illustration of gonad maturation in both female and male *Mytilus edulis* is shown in Figure 4. The assessment of developmental stages using histological methods is prone to bias, which represents a major disadvantage. Methods employing quantification of volume fractions of different gonad constituents have been used to decrease the classification bias (Gosling, 2015). The changes in genital tissue during the developmental cycle of *Mytilus edulis* concern mainly the gonad size and cell differentiation through gametogenesis (Figure 4). Such morphological changes could be quantified using digital image analysis, leading to more automated gender recognition and gonadal status classification. Image analysis approach to determine the ratio of gonad to the entire visceral area was reported as indicative of gonadal status in oysters (Gosling, 2015).

Apart from their central role in assessing the developmental stage, gonads in blue mussel also exhibit a range of histopathological alterations (Table 6).

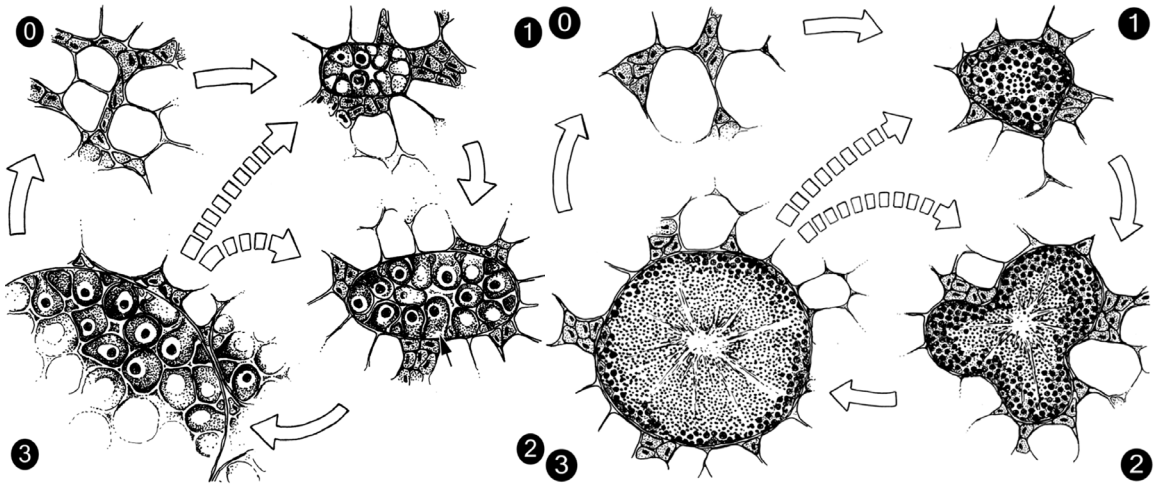

Figure 4 Schematic representation of ovarian and testicular developmental stages in *Mytilus edulis*. Stage 0 – empty gonad, Stage 1 – early filling, Stage 2 – filling, Stage 3 – mature gonad. Gonad development indicated by arrows. Dash arrows indicate changes in gonads due to partial spawning, leading to various stages of gonad maturation. Reprinted from (Duinker et al., 2008).

Table 4 Guidelines for assessment of gonad developmental stage (Seed, 1969; Seed and Brown, 1977).

Stage		Short description	Description
Resting / spent	Stage 0	The resting or spent gonad Undifferentiated; no gametes present	No traces of sexuality can be observed. It includes virgin animals where the reproductive system is rudimentary, and those which have completed spawning. During this period, stores of fat and glycogen are accumulated in the connective tissue and this frequently obscures the genital canals.
Developing (maturation) stages	Stage 1	Developing gametes appear; no ripe gametes	Characterized by the onset of gametogenesis, islands of germinal tissue appearing in the matrix of dense connective tissue. Very early stages are frequently difficult to separate into males and females. No ova or spermatozoa are present at this stage.
	Stage 2	Ripe gametes appear but majority still developing	Ripe gametes appear in the center of the follicles although these are mainly occupied by early stages of gametogenesis (small, numerous oocytes attached to the germinal epithelium in the female, and spermatogonia and spermatocytes in the male).
	Stage 3	Gonad half full of ripe gametes	There is general increase in the mass of the gonad at the expense of the stored food in the connective tissue. This is a stage of rapid gametogenesis with approximately half of each follicle occupied by ripe gametes, and half with early stages of gametogenesis. The area occupied by genital tissue is half of that of the fully ripe conditions.
	Stage 4	Majority of gametes ripe; a few developing gametes still present	Maximum proliferation of genital tissue is almost attained. There is a preponderance of ripe gametes in each follicle, with a general reduction in the earlier stages of gametogenesis. Gametogenesis is, however, in progress.
Ripe gonad	Stage 5	Gonad full of ripe gametes	Fully ripe conditions. It differs from 4 stage only in the greater reduction of early stages of gametogenesis (a few small oocytes in the germinal epithelium of the female, and a narrow band one of two cells deep, of spermatogonia and spermatocytes in the male). Ova are compacted into polygonal configuration, whilst in the males the follicles are distended with morphologically ripe spermatozoa. The time taken from morphological to physiological ripeness varies from a few weeks to a few months.

Table 5 Guidelines for assessment of spawning stage (Seed, 1969).

Stage	Short description	Description
Stage IV	General reduction in density of gametes occurs	The follicles are still relatively full of ripe gametes, but active discharge of there is now in progress. This is obvious from the general reduction in density of spermatozoa, and the rounding of the remaining ova as the pressure within the follicles is reduced following partial emission.
Stage III	Gonad half empty; unlike developing gonad very few early gametes are present	Similar to Stage 3 development in as much the follicles are approximately half full of mature gametes. Here, however, unlike developing Stage 3, relatively few early stages of gametogenesis are present. In the female, the ripe eggs are rounded rather than polygonal in appearance. There is a general reduction in the area of mantle covered by genital tissue.
Stage II	Gonad approx. ¾ spawned	At this stage the follicles are considerably less than half full of mature gametes. There is still further general reduction in the area occupied by genital tissue.
Stage I	Only residual gametes present; cytolysis may be in progress	Residual spermatozoa and ova are still present in the follicles and can often be seen undergoing cytolysis by amoeboid phagocytes (there are eosinophilic, and their nuclei stain intensely with hematoxylin). The center of the follicles is often filled with a yellow-brown matrix – the result of cytolysis of the residual gametes.

Table 6 Selected histopathological alterations in gonads in blue mussel.

Alteration type/name/description		Reference
Morphological	Degeneration of the ovarian follicles	(Aarab et al., 2004)
	Abnormally big and numerous ovarian follicles	(Aarab et al., 2004)
	Atretic/necrotic oocytes	(Cuevas et al., 2015)
Inflammatory	Heamocytic infiltration	(Cuevas et al., 2015)
	Haemocytic aggregates – accumulation of haemocytes inside the gonadal follicles or in the connective tissue	(Aarab et al., 2011; Bignell et al., 2008)
	Melanised haemocytic aggregates	(Aarab et al., 2011; Cuevas et al., 2015)
	Granulocytomas – often accompany severe inflammation	(Bignell et al., 2008; Cuevas et al., 2015)
	Brown cells - associated with lipofuscin-like pigments	(Cuevas et al., 2015)
Parasitic	*Nematopsis sp.*	(Cuevas et al., 2015)
	Tubelarria sp.	(Cuevas et al., 2015)

IMAGE ANALYSIS, WHOLE-SLIDE IMAGING, DIGITAL PATHOLOGY & ARTIFICIAL INTELLIGENCE

The aim of the image analysis is to obtain meaningful information from an image in an objective and reproducible manner. The concept of the image analysis originated early after the invention of the microscopy, when Anton van Leeuwenhoek developed a system to measure microscopic objects using a small reference object in the 17th century (Aeffner et al., 2019; Meijer et al., 1997). For nearly 300 years, image analysis progressed slowly, until digital imaging techniques and digital analysis became available in the second half of the 20th century (Aeffner et al., 2019). Particular improvements in the image analysis were brought in the last two decades, in parallel with the increasing technological advancement, especially with regard to computing power, storage capacity, image acquisition techniques and digital analysis. The wide availability of whole-slide imaging and digital image analysis techniques have led to a shift from the traditional pathology (glass-slide microscopy, light microscopy) towards a digital pathology (Niazi et al., 2019; Snead et al., 2016). Apart from the facilitated, lab-independent analysis, the adoption of digital images promotes remote consultation, collaboration, innovation and virtual education, and encourages structured and efficient work (Tizhoosh and Pantanowitz, 2018).

Data in form of an image might contain information of great complexity, and therefore the analysis of such could be challenging. Images generated by whole slide imaging (WSI) technique, used in digital pathology, represent an exceptional source of information: they are characterized by high complexity due to their large size and high resolution; they contain color information; they have no

evident anatomical orientation; they provide information at multiple scales (depending on magnification) and multiple z-stack levels (each cross-section will produce a different image). A complete, objective and reproducible analysis of such extent of visual information is beyond the capacity of a human eye (Niazi et al., 2019).

These obstacles can be overcome by the computer vision and artificial intelligence techniques. Computer vision is the science of making machines capable of extracting the information from an image and understanding its content, which is achieved by application of artificial intelligence (AI) – science that seeks to make machines produce intelligent content. The versatility of computer vision-based image analysis stems from automation, completeness and access to invisible data, features that are impossible to achieve by human perception, which relies on event selection by association (Danuser, 2011). High potential of AI technology is reflected by its progressing popularity and application in many fields - Food and Drug Administration (FDA) has approved AI-based apps for radiology and ophthalmology (Tizhoosh and Pantanowitz, 2018).

Table 7 Artificial intelligence and digital pathology – challenges and opportunities. Based on (Tizhoosh and Pantanowitz, 2018)

Challenges/disadvantages	Need for high quality labeled training images, manually annotated by experts (time consuming, challenging in low-quality images)
	High variation of histopathological pattern (need for training images addressing each variation)
	Processing large dimensions of WSI is computationally expensive using not optimized AI algorithms capacity (need for dividing image into smaller tiles or down-sampling (loss of potential image information)
	Lack of automated evaluation methods (pathologist as the ultimate reviewer of AI workflow and outcomes)
	Lack of universal algorithms to perform multitask work (each AI algorithm performs a specific task)
	Requirement for high-capacity Graphical Processing Units (GPUs) to train and apply AI algorithms to pathology images, which might be financially challenging
	Sensitivity of deep learning algorithms to a small artifacts or noise in the image; small artifacts may have large effect on the final result
	Lack of transparency and interpretability of the AI results (impossibility to trace the rationale behind the AI's decision)
Opportunities/advantages	Possibility to adapt pre-trained network to a new set of images
	Use of computer vision
	Possibility to apply generative models that can (re)generate data without making a decision
	Supervised & unsupervised learning
	Assistance in the pathologist's job (knowledge extraction)
	Automation and simplification of the mundane and complex tasks
	Extraction of previously unavailable information from the image stained with common stains (e.g. hematoxylin and eosin)
	Opportunity to include the data science into pathologist's job

DIGITAL IMAGE ANALYSIS SOFTWARE

In parallel with advancements in techniques to acquire and store high content images, a dynamic development has been observed in the field of bioimage analysis software. Dominated by ImageJ for nearly 25 years (Schneider et al., 2012), digital image analysis environment currently offers numerous options, both commercial and open source, with features adapted to specialized fields in bioimage analysis. Customized workflows, plugins and scripts, actively developed and shared by the software users worldwide have led to a great enhancement in the software capabilities. The growing popularity of whole slide imaging techniques and, consequently the large, high-content images became a challenge for most of the available open source software, incapable of handling image data of such size and complexity. For a long time, addressing the challenge of computation and visualization of large images was dominated by commercial image analysis software, often too expensive for a regular user and a financial burden to a research facility. In light of limited options, the quality of analysis of complex digital images was often constrained by image down-sampling or using only a subset of the original data (Bankhead et al., 2017).

QUPATH

The QuPath is an open source digital image analysis software, which was designed to address the WSI visualization and analysis by Pete Bankhead at the Queen's University Belfast. The presence of readily available, user-friendly tools to perform the most common analysis strategies in pathology makes the software suitable for the novice to image analysis. Single analysis steps can be combined into customized workflows, allowing batch-processing of image series. Indigenous capabilities of the software can be further extended by scripting functionality and addition of custom extensions, which is particularly attractive for programming-skilled users. A core feature of QuPath is the hierarchical approach of 'object-based' image data annotations/detections, which can be used to represent 'parent-child' relationship between image objects. Such approach can be further applied in training of object classifiers based on machine learning algorithms (Bankhead et al., 2017).

AIMS OF STUDY

OBJECTIVE

The overall aim of this study was to develop a method to analyze developmental stage, identify and quantify histological lesions in *Mytilus edulis* gills, using a digital image analysis system.

SPECIFIC AIMS

- Finding an open-source software for digital image analysis capable of handling WSI images used in the study and offering an AI analysis;
- Development of image analysis methods to identify and quantify histological lesions in gills of *Mytilus edulis*;
- Comparison of image analysis results with measurements of contamination in corresponding rig stations;
- Comparison between developed methods and visual assessment.

MATERIALS AND METHODS

MATERIALS

Mussel samples for the histopathology slide imaging and analysis were collected as a part of the Water column monitoring in 2017 (Pampanin et al., 2019), A subset of mussels was sampled in order to establish pre-exposure health parameters (time zero, T0). Mussels were caged around two offshore installations: Statfjord A and Statfjord B. Exact cage locations were chosen according to high probability of the PW plume direction in pre-simulations (Figure 5). In total, 19 mussel cages were deployed (Table 8). Mussels were collected after 6 weeks of the field exposure. Tissues for the histological study were stored in formalin solution (10% neutral buffered, ChemiTeknikk) at +4°C. Further handling and staining procedures were carried out as described in subsection Staining procedures.

Table 8 Mussel rigs distribution and location details. SFB – Statfjord B, SFA – Statfjord A, Ref – reference station. Based on (Pampanin et al., 2019).

Rig no.	Location	Station name	Distance	Direction	Description
4	Statfjord B	SFB 500 SE	500	SE	Near field
3		SFB 1000 SE	1,000	SE	Intermediate distance
1		SFB 10000 SE	10,000	SE	Far field
13	Statfjord A	SFA 500 E	500	E	Near field
12		SFA 500 SE 1	500	SE	Near field
11		SFA 500 SE 2	500	SE	Near field
10		SFA 500 SW	500	SW	Near field
14		SFA 500 NE	500	NE	Near field
7		SFA 1000 E	1,000	E	Intermediate distance
8		SFA 1000 SE	1,000	SE	Intermediate distance
9		SFA 1000 SW	1,000	SW	Intermediate distance
16		SFA 1000 NW	1,000	NW	Intermediate distance
15		SFA 1000 NE	1,000	NE	Intermediate distance
6		SFA 2000 E	2,000	E	Intermediate distance
5		SFA 2000 SE	2,000	SE	Intermediate distance
17		SFA 2000 NW	2,000	NW	Intermediate distance
2		SFA 10000 SE	10,000	SE	Far field
18	Ref.	Ref 1	-	-	Reference area
19		Ref 2	-	-	Reference area

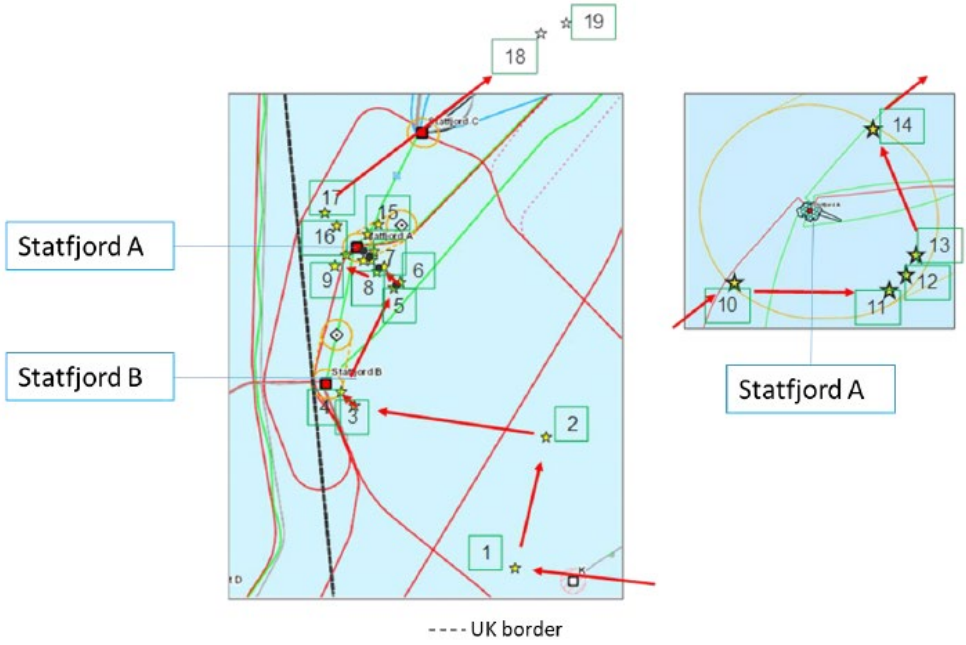

Figure 5 Location of mussel cage stations (marked as yellow stars) in relation to platforms (marked as red squares) Statfjord A and Statfjord B. Red arrows and numbers indicate the order of field deployment. Reprinted from (Pampanin et al., 2019).

METHODS

All histological sample preparation was performed at the Department of Pathology, Stavanger University Hospital, Norway, by Melinda Lillesand.

STAINING PROCEDURES

Preparation of the histology slides

Gonad and gill samples were carefully dissected, orientated and placed in histocassettes (ChemiTeknikk, Norway). Tissue processing was performed overnight using an automated Tissue Processor (Leica, Shandon Excelsior), during this process, tissue samples were dehydrated through 100% alcohol and a clearent Xylen after which there were fixed in 4% buffered formaldehyde. Samples were then embedded with melted 60°C paraffin (Q-Path, France). Paraffin embedded tissue samples were cut by semi-automated Microtome (Leica). Paraffin sections of 3 µm depth were mounted onto Superfrost Plus slides (Menzel, Braunschweig, Germany) and dried overnight at 37˚C. Before proceeding with the staining procedures, sections were deparaffinised and rehydrated.

Haematoxylin & Eosin staining procedure

Haematoxylin and eosin (H&E) staining was performed using Symphony automated stainer (Roche, France) on the deparaffinised and rehydrated sections.

Alcian Blue – Periodic Acid-Schiff staining

The special staining was performed using Bench Mark automated special stainer (Roche, France). Sections were stained by Alcian Blue – Periodic Acid-Schiff (AB-PAS) with diastase ready to use special staining kit (Roche, France). Acidic mucosubstances stained blue by incubating sections in Alcian Blue for 15 minutes, while neutral mucosubstances stained magenta red by incubating sections in PAS Period Acid for 4 minutes following by PAS Schiff's for 12 minutes. Cell nuclei were stained by PAS Hematoxylin. Sections were dehydrated and cover slipped by using automated stainer Symphony (Roche, France), choosing program *Special Stains Coverslip*.

MANUAL SCORING OF THE SLIDES

Gonadal status evaluation & spawning progression

Slides colored with H&E were used for evaluation of gonadal status and spawning stage. Tissues were examined according to the scheme proposed by Seed (Seed, 1969) as described in Table 4 and Table 5. The visual assessment was performed blind by another operator not engaged in sample collection or preparation (Pampanin et al., 2019).

Abnormal mucus secretion assessment

H&E colored slides were used in the assessment of abnormal mucus secretion severity. Slides were scored 0 – 4, with rising score indicating the increasing severity of the lesion.

DIGITAL IMAGE ANALYSIS SOFTWARE

Prior to the method development, few open-source image analysis software were evaluated and briefly tested for the analysis of the images used in the study. The comparison of candidate software is presented in Table 9.

Table 9 Comparison of selected available open-source image analysis software with respect to whole slide image (WSI) compatibility, presence of artificial intelligence (AI) functions, required programming skills and plugins. Based on (Zuraw).

Software	WSI compatibility	IA	IT skills	Plugins	Reference
QuPath	YES	YES	No	YES	(Bankhead et al., 2017)
Ilastik	NO	YES	No	NO	(Berg et al., 2019)
ImageJ/FIJI	Through plugin	YES	Minimal	YES	(Schindelin et al., 2012)
CellProfiler	Through plugin	YES	No	YES	(Lamprecht et al., 2007)
Orbit	YES	YES	No	YES	(Goldberg et al., 2005)
Icy	YES	YES	No	YES	(Chaumont et al., 2012)

Based on the high compatibility with WSI, presence of ready-to-use AI algorithms, relatively fast performance, and user-friendly interface, QuPath was chosen as the software for the method

development. At the beginning of the project, the QuPath v0.2.0-m8 was used, followed by QuPath v0.2.0-m9, (which was available from 14 February 2020), v0.2.0-m11 (April 2020) and v0.2.0-m12 (May 2020).

QUANTITATIVE ANALYSIS

The analysis of the results was performed using MS Excel, R Studio and IMB SPSS 26.0. Different statistical tests were applied, according to amount of groups tested and data distribution. Mann-Whitney test was used for data with 2 independent groups of not normal distribution, Kruskal-Wallis H test for data with more than 2 independent groups of not normal distribution and One-Way ANOVA for data with more than 2 independent groups with normal distribution. All p-values reported are 2-sided. Normal distribution was tested using Shapiro-Wilk test. The results of statistical tests were reported as p-value (significance level 0.05). Additionally, for Kruskal-Wallis H test, Mann-Whitney U test and One-Way ANOVA, test statistic (H, U or F, respectively) and degrees of freedom (df) were reported.

All groups in data analysis refer to original scoring by visual assessment performed as integral part of WCM 2017. In sections concerning genital tissue, groups refer to manual scoring of gonadal status (0-5) or spawning stage (1-4). In sections concerning gill tissue, groups were created according to “Abnormal mucus secretion” score (0-4).

HARDWARE

The work was performed using HP EliteBook 840 G6, with the following device specifications:

Processor	Intel®Core™ i5-8265U CPU @ 1.60 GHz 1.8 GHz, x64-based processor
Installed RAM	8.00 GB (7.81 GB usable)
System type	64-bit operating system

QUPATH ANALYSES

DETERMIANTION OF THE MUSSEL’S GENDER USING DIGITAL IMAGE PROPERTIES

In 189 histology images available for the gonad analysis, gender could be evaluated based on the visual assessment. To automate the process of the gender recognition, an attempt was taken to link the digital image properties to the gender.

The image analysis protocol was created in QuPath (v0.2.0-m8) and the measurements were analyzed using R Studio. The protocol was applied to all 189 digital histology images of gonads. The method was based on the following outline and QuPath script.

1. Detection of the whole tissue region using a simple threshold approach on RGB channels. Exemplary outcome is illustrated in Figure 6;
2. Quantification of the digital image parameters: optical density, haematoxylin, eosin, red, green, blue, hue, saturation, brightness. For all the parameters, mean, standard deviation, minimum, maximum and median values were extracted.

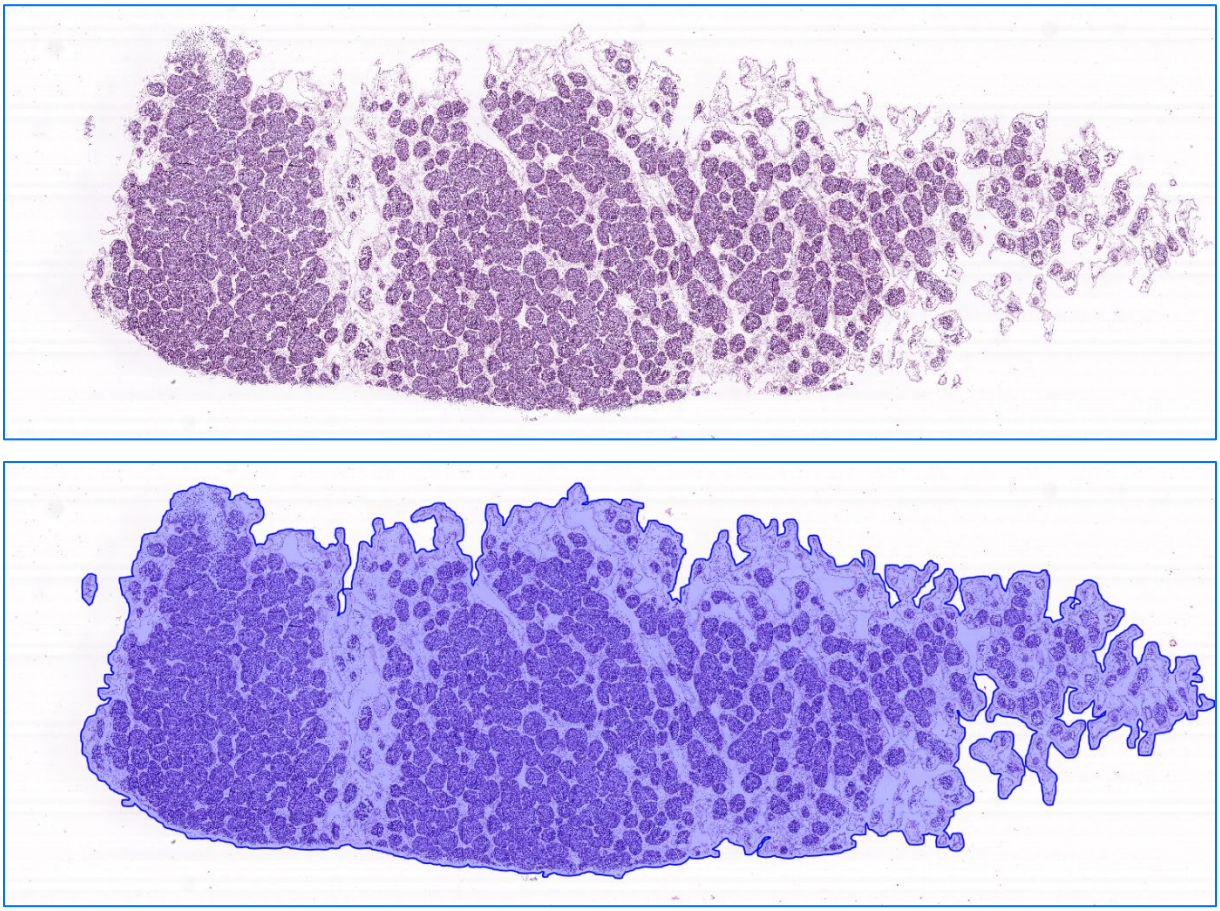

Figure 6 Example of tissue detection outcome in the image of gonad. Top: raw image of *Mytulis edulis* genital tissue. Bottom: simple tissue detection outcome. Detected tissue area outlined by the blue boundary and filled with transparent blue colour. Analysis performed in QuPath (v0.2.0-m9).

```
#setting image type as Hematoxylin & Eosin to define the stains vectors and
creating a full image annotation;

setImageType('BRIGHTFIELD_H_E');

setColorDeconvolutionStains('{"Name" : "H&E default", "Stain 1" :
"Hematoxylin", "Values 1" : "0.65111 0.70119 0.29049 ", "Stain 2" :
"Eosin", "Values 2" : "0.2159 0.8012 0.5581 ", "Background" : " 255 255 255
"}');
createSelectAllObject(true);

#tissue detection;

runPlugin('qupath.imagej.detect.tissue.SimpleTissueDetection2',
'{"threshold": 215,  "requestedPixelSizeMicrons": 20.0,  "minAreaMicrons":
```

```
50000.0, "maxHoleAreaMicrons": 1000000.0, "darkBackground": false,
"smoothImage": true, "medianCleanup": true, "dilateBoundaries": true,
"smoothCoordinates": true, "excludeOnBoundary": false,
"singleAnnotation": true}');

#quantification of digital image parameters;

selectAnnotations();
runPlugin('qupath.lib.algorithms.IntensityFeaturesPlugin',
'{"pixelSizeMicrons": 2.0, "region": "ROI", "tileSizeMicrons": 25.0,
"colorOD": true, "colorStain1": true, "colorStain2": true,
"colorStain3": false, "colorRed": true, "colorGreen": true, "colorBlue":
true, "colorHue": true, "colorSaturation": true, "colorBrightness":
true, "doMean": true, "doStdDev": true, "doMinMax": true, "doMedian":
true, "doHaralick": false, "haralickDistance": 1, "haralickBins": 32}');

saveAnnotationMeasurements('/C:/Users/OneDrive - Universitetet i
Stavanger/Environmental Engineering Master Thesis/Measurements
QuPath/Gonad_images_measurements_male/')
```

ASSESSMENT OF MALE GONAD DEVELOPMENTAL STAGE

The distribution of the developmental stage is an important characteristic of a population in this thesis. Herein, a method was developed to automate developmental stage assessment, based on the following outline and QuPath (v0.2.0-m8) script:

1. Due to large image size and thus high computational demand, in the following analysis steps, a region of approximately 40,000 μm² was selected on each image, which substantially reduced the processing time. The image region was chosen with the following criteria:
    a) The whole region must be located within the tissue area,
    b) The proportion of gonads within region had to be as close as possible to the average across the whole tissue by visual assessment. Gonads detection was used to better visualize their distribution in the tissue, making the region appointment more informed,
    c) The edge of tissue area was avoided due to possible displacement of the tissue and gonads.
2. Detection of gonads in the tissue using simple tissue detection;
3. Detection of cells within gonads in the appointed region using watershed cell detection;
4. Adding smoothed features, which supplement the current cell measurements by taking a weighted average of the corresponding measurements of neighboring cells within 15 μm radius (measurement maps illustrating raw and smoothed measurements in the detected cells are presented in Figure 8);

Classification of detected cells as spermatogonium I, spermatogonium II or spermatozoa using object classifier according to cell dimension measurements, stain intensity and nearby objects characteristics. Exemplary classification outcomes are presented in

5. Figure 7. Object classifier specifications:
    a) Classifier type: Random Trees (with default parameters);
    b) Cell features for classification: Area, Perimeter, Minimum caliper, Hematoxylin OD mean, Hematoxylin standard deviation, Hematoxylin OD minimum, Smoothed: 15 µm: Area, Smoothed: 15 µm: Perimeter, Smoothed: 15 µm: Min caliper, Smoothed: 15 µm: Hematoxylin OD mean, Smoothed: 15 µm: Hematoxylin OD standard deviation, Smoothed: 15 µm: Hematoxylin OD max, Smoothed: 15 µm: Hematoxylin OD min, Smoothed: 15 µm: Nearby detection count.
6. Export of measurements of the whole tissue region, cumulative area of gonads within the tissue and number of each cell type in the gonads within the region of choice.
7. Comparison of the results with manually assigned developmental stages.

The following QuPath script was created for already selected image regions of approximately 40,000 µm^2.

```
#setting the image type to define the vectors of the stains;
setImageType('BRIGHTFIELD_H_E');
setColorDeconvolutionStains('{"Name" : "H&E default", "Stain 1" : "Hematoxylin", "Values 1" : "0.65111 0.70119 0.29049 ", "Stain 2" : "Eosin", "Values 2" : "0.2159 0.8012 0.5581 ", "Background" : " 255 255 255 "}');
#creating full image annotation and saving the image region area measurements;
createSelectAllObject(true);
saveAnnotationMeasurements('/C:/Users/OneDrive - Universitetet i Stavanger/Environmental Engineering Master Thesis/Measurements QuPath/Male_segment_area/')
#detection of gonads within the image region;
runPlugin('qupath.imagej.detect.tissue.SimpleTissueDetection2', '{"threshold": 170, "requestedPixelSizeMicrons": 5.0, "minAreaMicrons": 5000.0, "maxHoleAreaMicrons": 2500.0, "darkBackground": false, "smoothImage": true, "medianCleanup": true, "dilateBoundaries": false, "smoothCoordinates": true, "excludeOnBoundary": false, "singleAnnotation": true}');
#cell detection;
selectAnnotations();
runPlugin('qupath.imagej.detect.cells.WatershedCellDetection', '{"detectionImageBrightfield": "Hematoxylin OD", "requestedPixelSizeMicrons": 0.25, "backgroundRadiusMicrons": 0.0, "medianRadiusMicrons": 0.5, "sigmaMicrons": 0.5, "minAreaMicrons": 1.0, "maxAreaMicrons": 1000.0, "threshold": 0.35, "maxBackground": 2.0, "watershedPostProcess": true, "cellExpansionMicrons": 0.0, "includeNuclei": true, "smoothBoundaries": true, "makeMeasurements": true}');
#adding smoothened features for the detected cells;
```

```
runPlugin('qupath.lib.plugins.objects.SmoothFeaturesPlugin',
'{"fwhmMicrons": 15.0, "smoothWithinClasses": false, "useLegacyNames":
false}');

#cell classification into three groups: spermatogonia I, spermatogonia II and
spermatozoa;

runClassifier('C:/Users/OneDrive - Universitetet i Stavanger/Environmental
Engineering Master Thesis/QuPath Projects/ Gonad_segments_0.2.0_2/
classifiers/Spermatogonia_I_II_spermatozoa.qpclassifier');

saveAnnotationMeasurements('/C:/Users/OneDrive - Universitetet i Stavanger/
Environmental Engineering Master Thesis/Measurements QuPath/
Male_gonad_cell_classification/')
```

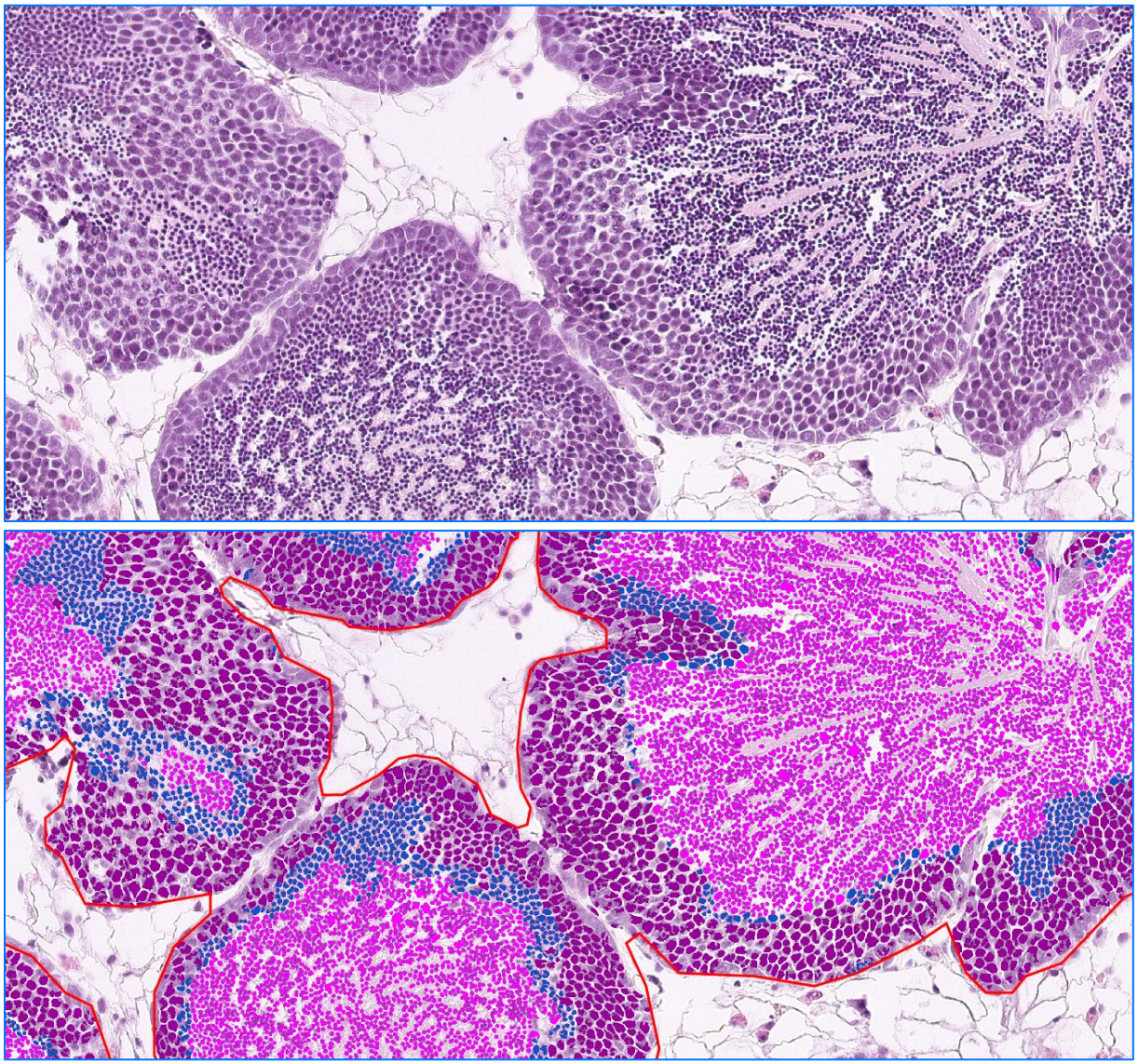

Figure 7 Top: Exemplary snapshot of Haematoxylin & Eosin coloured image of gonad tissue in male *M. edulis* individual. Bottom: The same image with object classifier outcomes. Dark purple indicates spermatogonia I, blue – spermatogonia II and magenta – spermatozoa. Red line indicates detected gonads. Details of the analysis described in text.

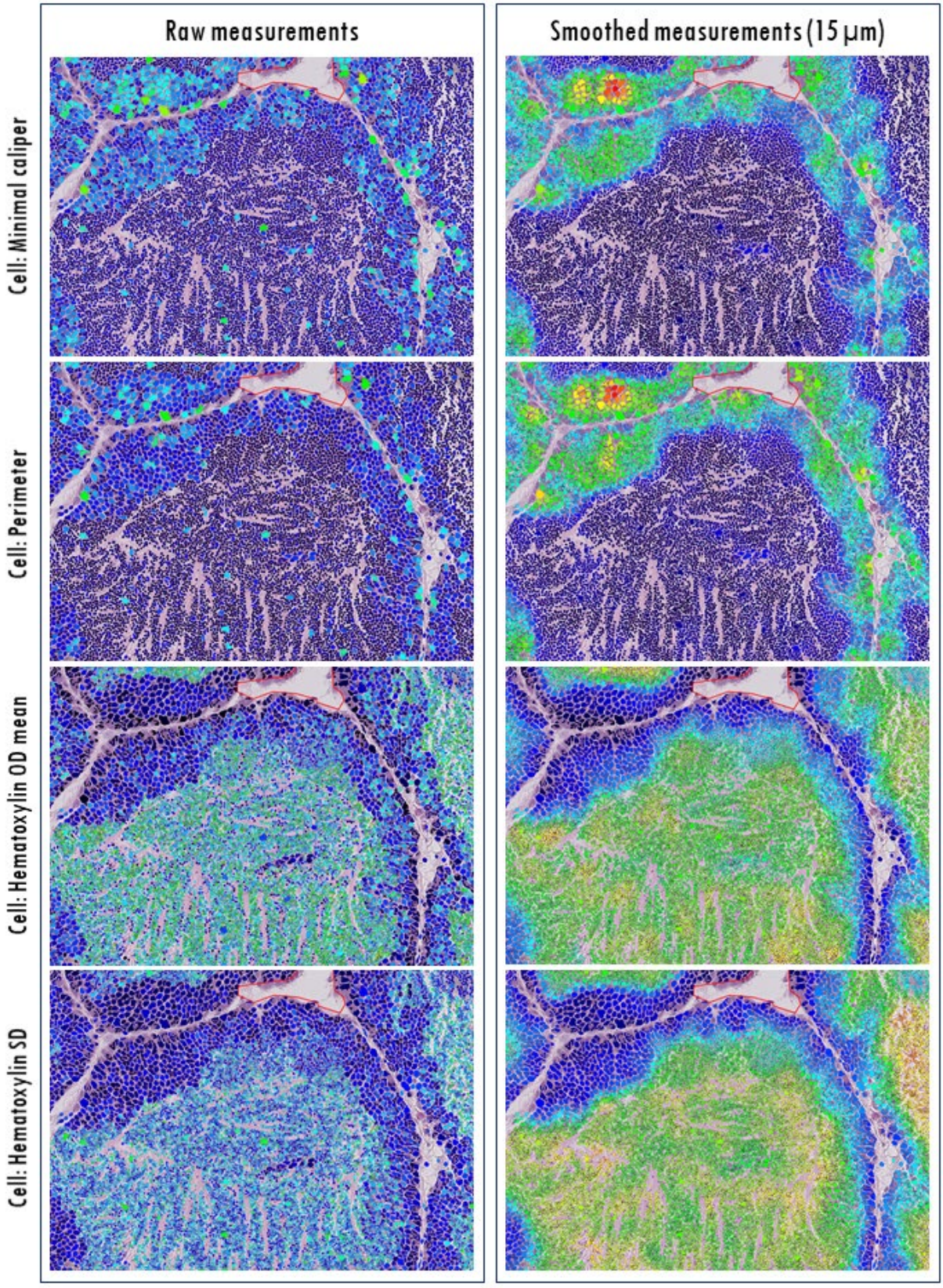

Figure 8 Measurement maps for the cells detected in male gonads of *M. edulis*, illustrating comparison between raw and smoothed measurements. Smoothed features command performed in QuPath (v0.2.0-m8) for 15 µm radius. Presented in the figure are cell measurements of minimal calliper, perimeter, mean optical density (OD) of haematoxylin and haematoxylin standard deviation (SD).

MUCUS CELL QUANTIFICATION IN GILL TISSUE

The abnormal mucus secretion in gill tissue could represent a response of the tissue to contaminants. Gill tissue slides were colored with AB-PAS staining, which allows for the visualization of acidic and neutral mucins, as well as distinction between those two (Myers, 2018). The method for detection of AB-PAS stained cells was developed in QuPath (version v0.2.0-m11), according to the following outline and QuPath script:

1. Detection of the gill tissue using a simple threshold approach on RGB channels;
2. Detection of cells within the tissue using the hematoxylin optical density threshold in the watershed cell detection method.
3. Classification of the detected cells based on the mean optical density threshold of the DAB channel. Cells were classified on dark blue (stained for acidic mucins) and dark purple (containing both acidic and neutral mucins).
4. Export and analysis of the measurements: total area of detected tissue, number of cells detected in the tissue area.

```
#setting the image type to define the stain vectors;

setImageType('BRIGHTFIELD_OTHER');

setColorDeconvolutionStains('{"Name" : "H-DAB default", "Stain 1" :
"Hematoxylin", "Values 1" : "0.65111 0.70119 0.29049 ", "Stain 2" : "DAB",
"Values 2" : "0.26917 0.56824 0.77759 ", "Background" : " 255 255 255 "}');

#simple tissue detection;

createSelectAllObject(true);

runPlugin('qupath.imagej.detect.tissue.SimpleTissueDetection2',
'{"threshold": 210,  "requestedPixelSizeMicrons": 15.0,  "minAreaMicrons":
10000.0,  "maxHoleAreaMicrons": 1000000.0,  "darkBackground": false,
"smoothImage": true,  "medianCleanup": true,  "dilateBoundaries": true,
"smoothCoordinates": true,  "excludeOnBoundary": false,
"singleAnnotation": true}');

saveAnnotationMeasurements('/C:/Users/OneDrive - Universitetet i
Stavanger/Environmental Engineering Master
Thesis/MeasurementsQuPathm11/Gill_tissue_size_ABPASD');

#cell detection within the tissue;

selectAnnotations();

runPlugin('qupath.imagej.detect.cells.WatershedCellDetection',
'{"detectionImageBrightfield": "Hematoxylin OD",
"requestedPixelSizeMicrons": 0.5,  "backgroundRadiusMicrons": 0.0,
"medianRadiusMicrons": 0.0,  "sigmaMicrons": 1.5,  "minAreaMicrons": 5.0,
"maxAreaMicrons": 4000.0,  "threshold": 0.75,  "maxBackground": 0.0,
"watershedPostProcess": true,  "excludeDAB": false,
```

```
"cellExpansionMicrons": 0.0,  "includeNuclei": false,  "smoothBoundaries":
true,  "makeMeasurements": true}');

#cell classification;

runObjectClassifier("Mucus cell classifier");

saveDetectionMeasurements('/C:/Users/OneDrive - Universitetet i
Stavanger/Environmental Engineering Master
Thesis/MeasurementsQuPathm11/Gill_mucus_cells_ABPAS/')
```

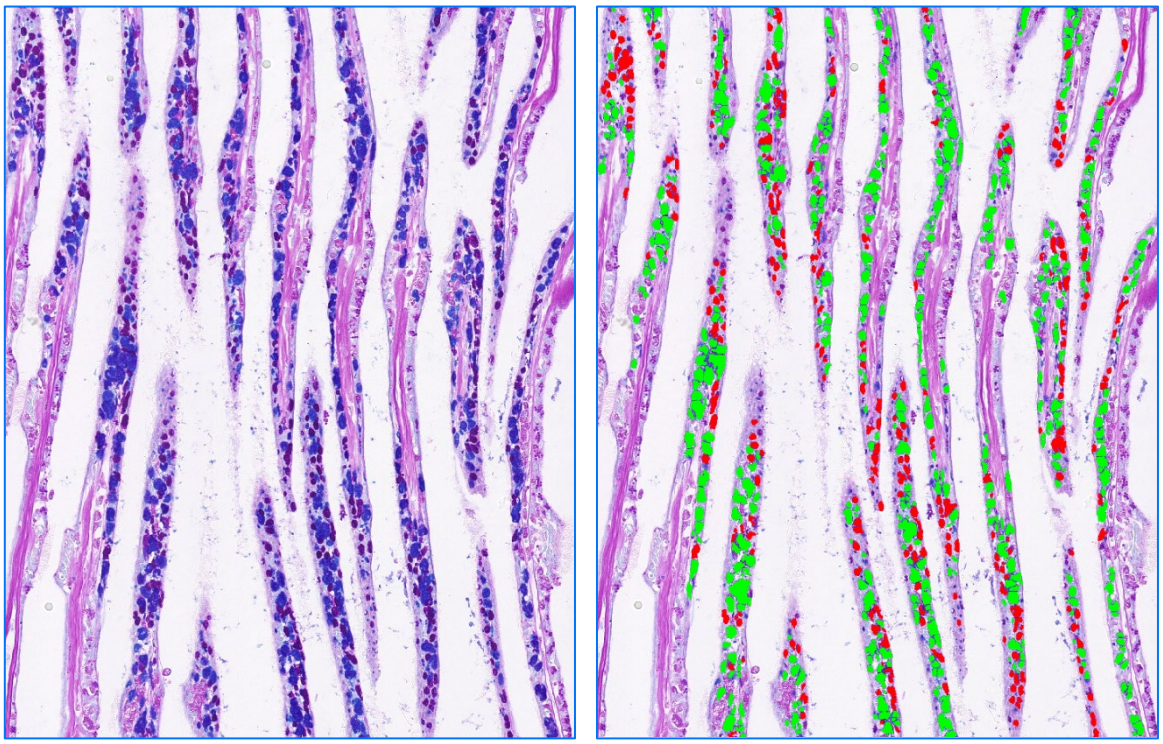

Figure 9 Example of outcome of object classification based on simple threshold in QuPath (v0.2.0-m12). Left: raw image of *M. edulis* gill tissue stained with Alcian Blue – Periodic Acid-Schiff (AB-PAS). Right: the same image with classification of detected cells. Colours indicate acidic mucus cells (green classification) and neutral mucus cells (red classification).

RESULTS

DETERMINATION OF THE MUSSELS'S GENDER USING DIGITAL IMAGE PROPERTIES

Simple threshold approach

In total, 5 basic statistics (mean, standard deviation, minimum, maximum and median) for the total of 9 intensity features (optical density sum, hematoxylin, eosin, red, green, blue, hue, saturation and brightness) were analyzed. After the visual assessment of the box plots, 6 measurements were considered as good candidates for the gender classification, based on the little overlap between male and female groups. All candidate features (optical density sum, hematoxylin, brightness, red, green and blue) concerned the standard deviation of the parameter. The other features exhibited a higher extent of overlapping values in both groups (data not shown). The distribution of values for 6 candidate parameters across the female and male groups is illustrated in box plots in Figure 10 and Figure 11. Data distribution within groups was evaluated using Shapiro-Wilk test for normality. Due to non-normal distribution in the majority of groups, non-parametric Mann-Whitney test was used for group comparison. Statistical tests were performed in IMB SPSS 26.0. The results are shown in Table 10.

The results of Mann-Whitney test show that differences between female and male groups are significant for all 6 candidate parameters.

Table 10 Mann-Whitney test results for gender-defined groups: female (F) and male (M). N – Sample size; SD – standard deviation.

Parameter	Group (N)	Mean (±SD)	Normality (p-value)	Mann-Whitney U	p-value (2-sided)
Optical density SD	F (90)	0.2278 (± 0.0265)	< 0.05	8907	< 0.0001
	M (99)	0.4474 (± 0.0522)	0.623		
Hematoxylin SD	F (90)	0.0804 (± 0.0087)	< 0.05	8910	< 0.0001
	M (99)	0.1755 (± 0.0202)	0.336		
Brightness SD	F (90)	0.1053 (± 0.0091)	<0.05	8906	< 0.0001
	M (99)	0.1681 (± 0.0129)	0.096		
Red SD	F (90)	28.42 (± 2.45)	< 0.05	8908	< 0.0001
	M (99)	47.18 (± 3.60)	< 0.05		
Green SD	F (90)	37.19 (± 3.49)	< 0.05	8887	< 0.0001
	M (99)	54.17 (± 3.88)	0.824		
Blue SD	F (90)	26.71 (± 2.37)	< 0.05	8906	< 0.0001
	M (99)	42.62 (± 3.26)	0.105		

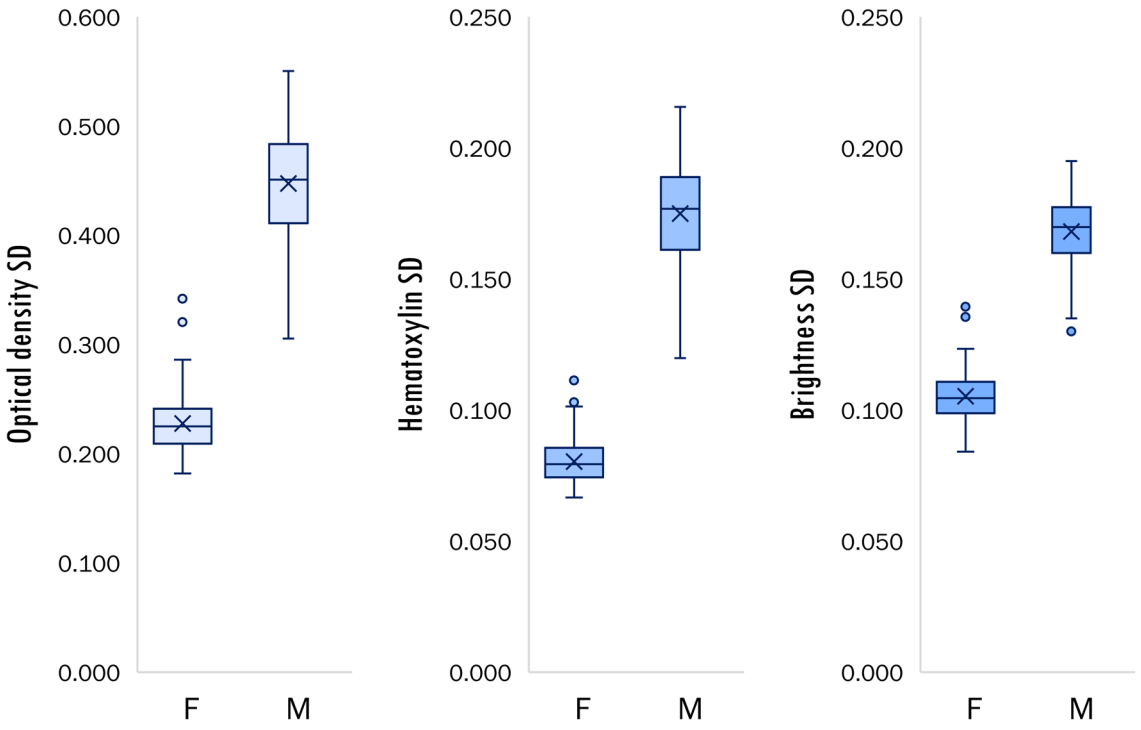

Figure 10 Between-gender comparison of standard deviation (SD) of image intensity values: Optical density (left), haematoxylin (centre) and Brightness (right). Box plots based on the analysis of all 189 images.

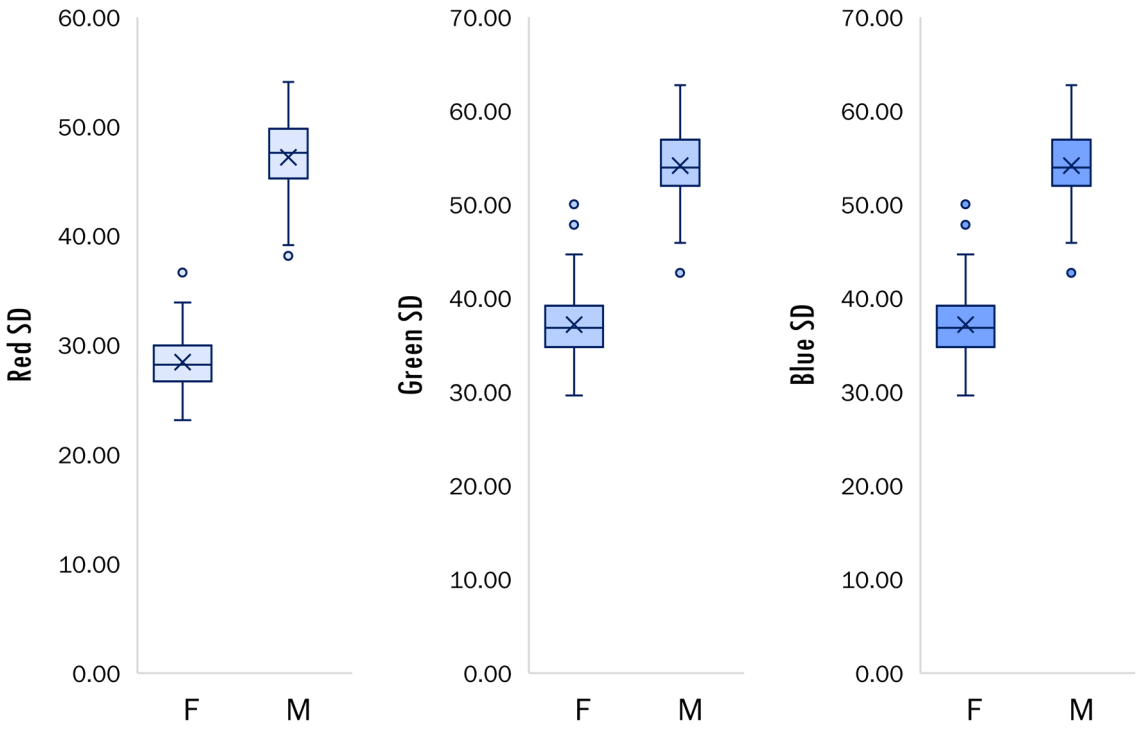

Figure 11 Between-gender comparison of standard deviation (SD) of the image intensity values: Red (left), Green (centre) and Blue (right). Box plots based on the analysis of all 189 images.

In order to apply simple threshold to divide the images between male and female individuals, thresholds for each intensity feature were defined as the mean value of the minimal and the maximum values in the male and the female group, accordingly. Thresholds and classification outcomes are presented in Table 11.

Table 11 Threshold-based gender classification, intensity features, thresholds used in the classification and the classification outcomes. SD – standard deviation.

Intensity feature (SD)	Threshold	Correctly classified images	Classification accuracy
Hematoxylin	0.11695	189/189	100.00%
Optical density sum	0.32385	187/189	98.94%
Brightness	0.13475	186/189	98.41%
Red	36.7534	187/189	98.94%
Green	46.3673	185/189	97.88%
Blue	34.2197	186/189	98.41%

The accuracy of the threshold-based classification was high for all intensity features. The highest (100%) accuracy was observed for the hematoxylin standard deviation. Other features were >97% accurate in classifying the images, and the slightly lower accuracy was due to the outliers with overlapping intensity features values, as seen in the box plots above. The detailed classification outcomes are not shown.

K-means clustering

Using the same measurements, a classification method using K-means clustering was tested. K-means clustering was chosen as an alternative to the simple threshold-based classification due to its capacity to account for multiple variables and higher flexibility, concerning the intensity feature variance. The shape of the potential data clusters is suitable for the k-means algorithm. K-means clustering was performed in R Studio.

Before the analysis, data were scaled in order to make the variables comparable. K-means clustering requires the specified cluster number prior to the analysis. The number of clusters of interest in the analysis was 2 (male and female). In order to determine the optimal amount of clusters for the data and assess the quality of clustering, the elbow method and the average silhouette method were used. In average silhouette, the higher the average silhouette width, the better the quality of clustering. Both methods confirmed the suitability of 2 clusters for the data (Figure 12) which were applied in the K-means algorithm. Clustering results are illustrated in Figure 13. The comparison of the gender classification based on the visual assessment and the K-means clustering algorithm is included in Table 16 (Appendix 1: Determination of the mussel's gender using digital image properties). Out of the whole dataset of 189 images, only 2 were not correctly

classified. The overall accuracy of the algorithm was 98.94%. Two principal components rendered by the k-means algorithm explained 99.78% of the variance between the groups.

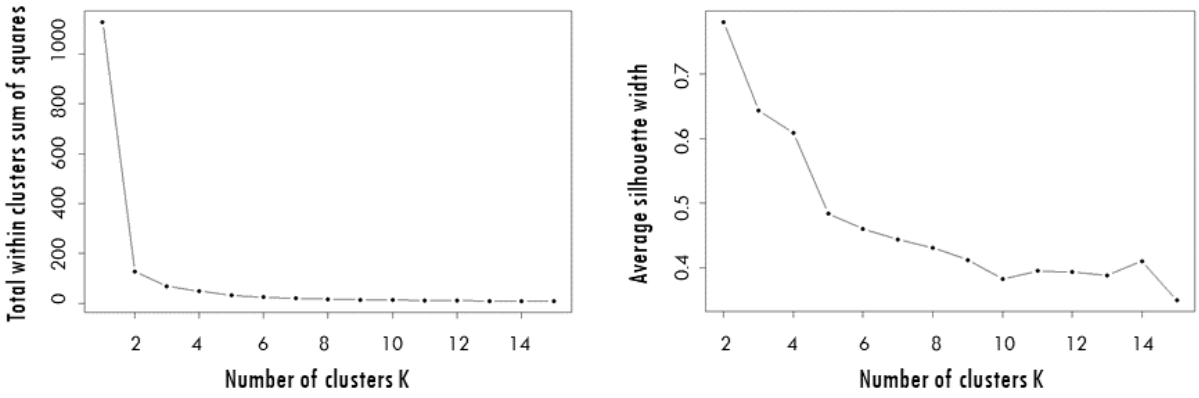

Figure 12 Determination of optimal number of clusters for K-means clustering using Elbow Method (left) and Average Silhouette Method. Analysis performed in R. Average silhouette analysis according to purrr library.

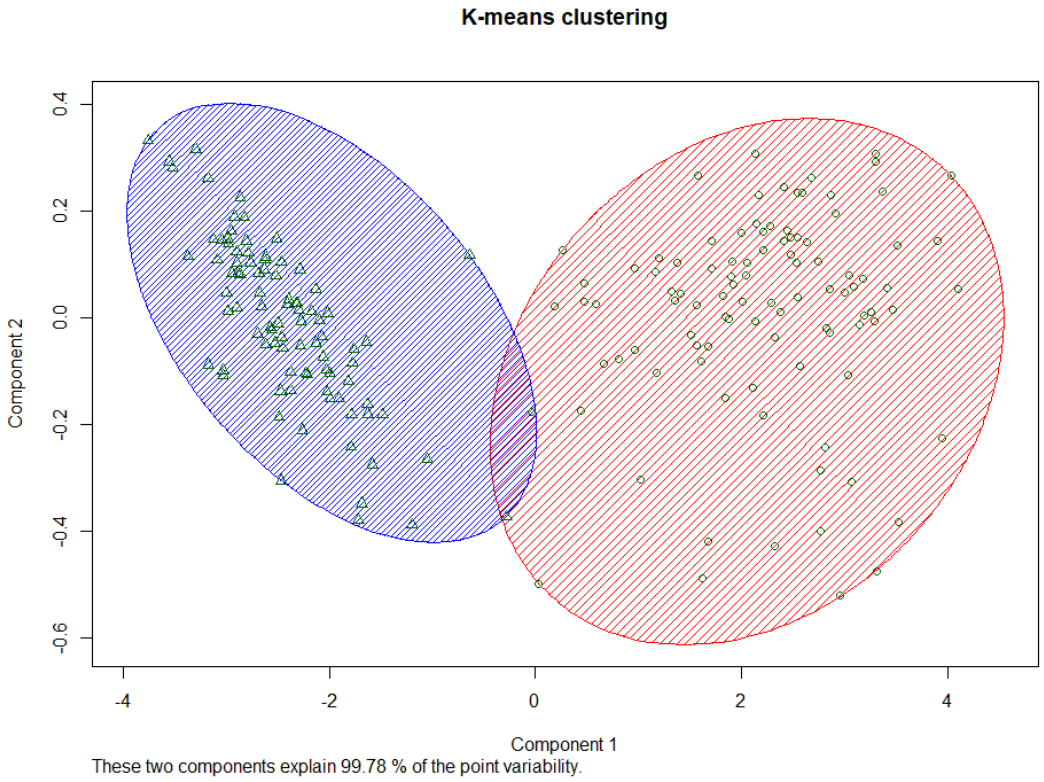

Figure 13 Principal component analysis. Representation of the K-means clustering outcomes.

DETERMINATION OF GONADAL STATUS IN MALE MUSSELS

Gonadal status

In the original dataset, the distribution of the visually assigned (according to Table 4) gonadal status was between stage 3 and stage 5, with 77 images classified as stage 3, 17 images as stage 4 and

4 images as stage 5. Due to small size of "Stage 5" group, it was merged with "Stage 4" groups into one bigger group, further referred to as "4 and 5". Between-group differences for the following measurements were analysed:

- Genital tissue content,
- Percentage of spermatogonia count per total cell count,
- Percentage of spermatogonia I count per total spermatogonia count.

After the visual assessment of measurements distribution illustrated in Figure 14, an observed overlap is seen between groups, especially for the content of different cell types. None of the measurements gave basis for the direct classification of groups. However, a tendency of increasing gonad content in the tissue was observed for advancing gonadal status. The tendency was also observed when plotting data according to original groups (Stage 3, Stage 4 and Stage 5), as shown in (Figure 15). Statistical significance between groups "Stage 3" and "Stage 4 and 5" was tested using Mann-Whitney test for two independent samples (Table 12). Significant differences were observed for the genital tissue percentage and spermatogonia percentage. Due to small size of originally scored "Stage 5" group, the statistical analysis between those original groups was impossible.

Table 12 Statistical test results: differences between groups derived from visually assessed gonadal status score. Groups with score 4 and 5 were merged into one group "4 and 5". N – Group size; SD – standard deviation.

Parameter	Groups (N)	Mean (±SD)	Shapiro-Wilk (p-value)	Mann-Whitney U	p-value (2-sided)
Genital tissue percentage	3 (77)	71.48 (± 16.18)	< 0.05	245	< 0.0001
	4 and 5 (21)	86.42 (± 7.39)	< 0.05		
Spermatogonia percentage	3 (77)	23.86 (± 6.86)	< 0.05	437	0.001
	4 and 5 (21)	19.37 (± 3.94)	0.556		
Spermatogonia I percentage	3 (77)	61.54 (± 8.49)	< 0.05	963	0.181
	4 and 5 (21)	64.37 (± 5.17)	0.258		

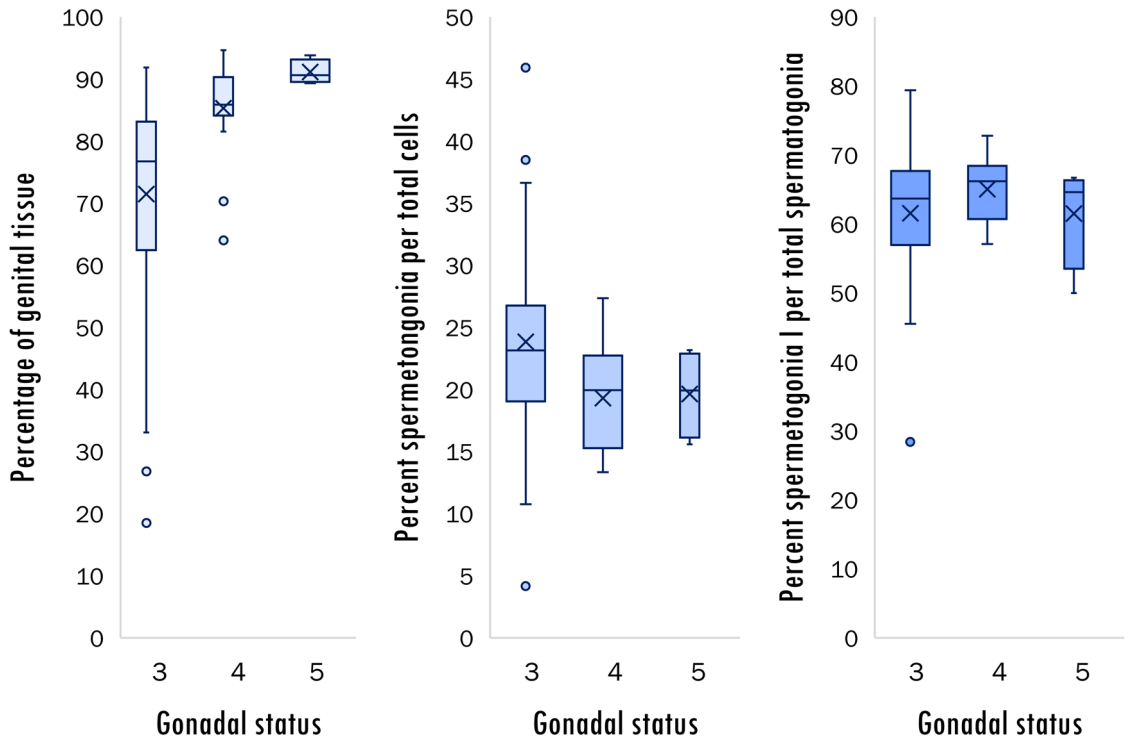

Figure 14 Differences between the gonadal status groups: "Stage 3" and "Stage 4 and 5". Left: percent of gonad area (genital tissue) per tissue area. Centre: count of spermatogonia per total cell count. Right: count of spermatogonia I per total spermatogonia (spermatogonia I and II) count.

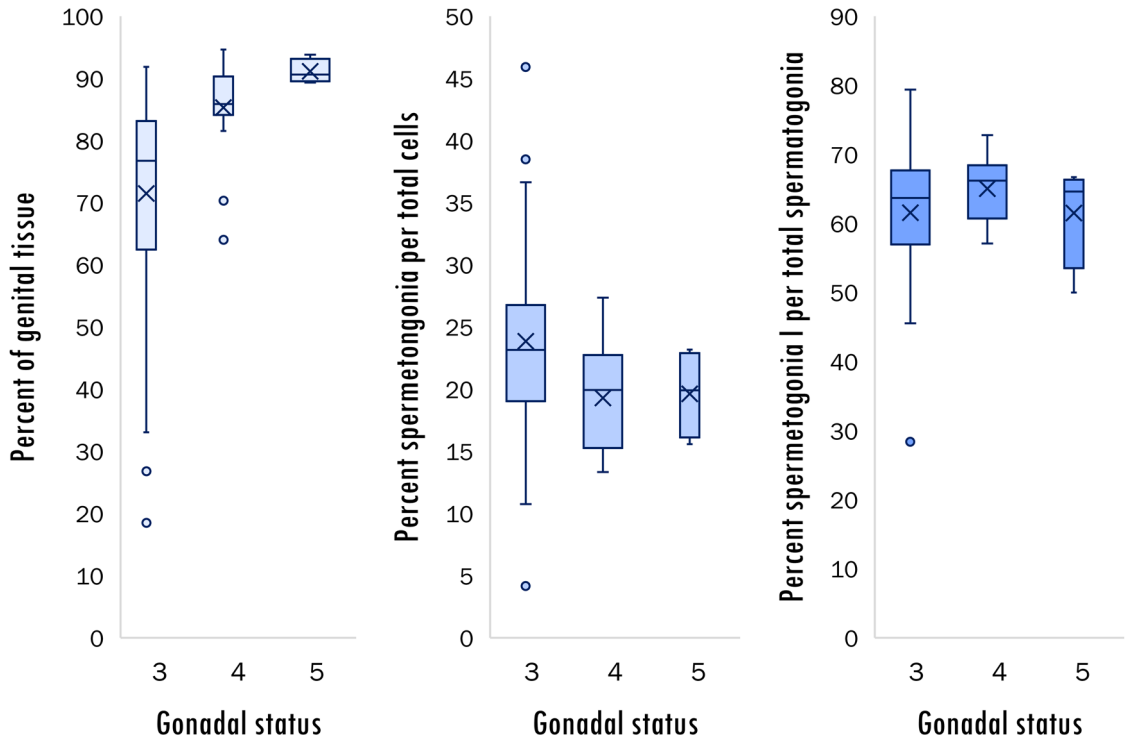

Figure 15 Differences between the gonadal status groups: "Stage 3", "Stage 4" and "Stage 5". Left: percent of gonad area (genital tissue) per tissue area. Centre: count of spermatogonia per total cell count. Right: count of spermatogonia I per total spermatogonia (spermatogonia I and II) count.

Spawning stage

The same parameters were analysed for between-group differences for spawning stages. The original grouping for spawning stages was kept, since each group has at least 7 samples: Stage 4 – 10 individuals, Stage 3 – 50 individuals, Stage 2 – 31 individuals and Stage 1 – 7 individuals. The distribution of measurements within and between the groups is illustrated in Figure 16. Similarly to gonadal status, a correlation was observed for the genital tissue content. With decreasing spawning stages, an increase in spermatogonia percentage of total cells was observed. Upon visual assessment, little differences between spawning stages were seen for the spermatogonia I percentage of total spermatogonia count.

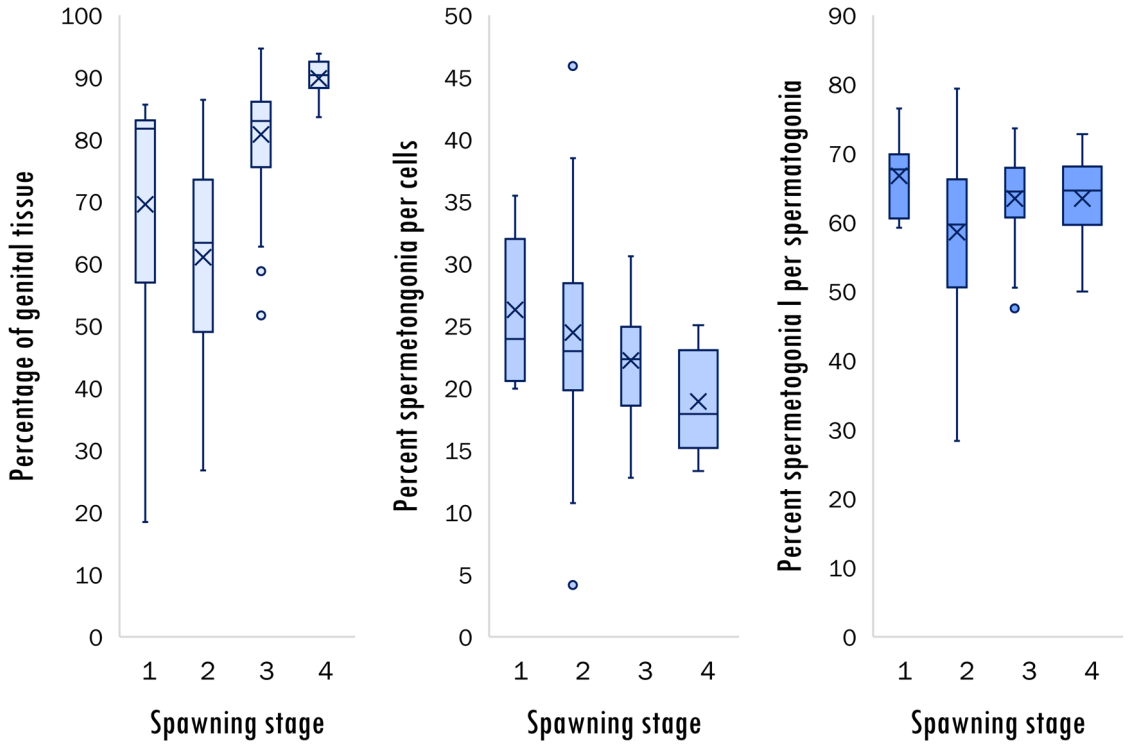

Figure 16 Differences between the spawning stage groups: "Stage 1", "Stage 2", "Stage 3" and "Stage 4". Left: percent of gonad area (genital tissue) per tissue area. Centre: count of spermatogonia per total cell count. Right: count of spermatogonia I per total spermatogonia (spermatogonia I and II) count.

Due to non-normal distributed data (Shapiro-Wilk test, Table 13), significance was evaluated using Kruskal – Wallis H test. Detailed results on comparison between each group is included in Appendix 2: Determination of gonadal status in male mussles**Feil! Fant ikke referansekilden.**. Significance was observed only for genital tissue content between groups 2-3, 2-4, 1-4 and 3-4. The remaining groups did not reach significance in pairwise comparison.

Table 13 Statistical tests for difference between groups derived from visually assigned spawning stage (1-4). N – Group size; SD – standard deviation. Kruskal-Wallis statistic reported as χ^2 (df) = H, where df = degrees of freedom.

Parameter	Groups (N)	Mean (± SD)	Shapiro-Wilk (p-value)	Kruskal-Wallis statistic	P-value (2-sided)
Genital tissue percentage	1 (7)	69.56 (± 24.50)	< 0.05	H(3) = 48.148	< 0.0001
	2 (31)	61.05 (± 15.31)	0.260		
	3 (50)	80.81 (± 8.82)	< 0.05		
	4 (10)	89.87 (± 3.23)	0.300		
Percentage spermatogonia	1 (7)	26.31 (± 6.41)	0.149	H(3) = 8.719	0.033
	2 (31)	24.47 (± 8.39)	0.691		
	3 (50)	22.23 (± 5.30)	< 0.05		
	4 (10)	18.94 (± 4.12)	0.432		
Percentage spermatogonia I	1 (7)	66.73 (± 5.98)	0.802	H(3) = 8.506	0.037
	2 (31)	58.56 (± 10.49)	0.325		
	3 (50)	63.47 (± 5.67)	< 0.05		
	4 (10)	63.47 (± 6.63)	0.848		

THE MUCUS CELL QUANTIFICATION IN GILL TISSUE

Out of 199 images, 32 were missing manually assessed degree of abnormal mucus secretion and were therefore excluded from some data processing steps. The analysis of image intensity values (measured for the detected tissue area) for the remaining 167 images revealed that in case of 17 images the median optical density value was 0, indicating improperly detected tissue area (white background area was erroneously detected as tissue, leading to low median value of the optical density). These 17 images were therefore excluded from the analysis, as the outcomes of the analysis were measured per tissue area.

The original manually assigned scores for “Abnormal mucus secretion” had values between 0 and 4, with increasing values indicating the gradually higher severity of the lesion. The distribution of the image count for each scoring group was unequal, with images per group: 0 – 44, 1 – 45, 2 – 42, 3 – 17 and 4 – 2 images. In order to facilitate the statistical analysis, groups scored as 3 and 4 were merged into one group with the score of 3.

The score-defined groups were analysed for between-group differences in regard to:

- AB-PAS stained cells count per gill tissue area,
- AB-PAS stained cells area per gill tissue area,
- Ratio of cells stained for acidic mucins (further referred to as “acidic mucus cells”) to the total stained cells.

The groups were also analysed in regard to intensity values of AB and PAS stains measured within the whole gill tissue area. All groups in the analysis refer to grouping based on original lesion score, according to visual assessment.

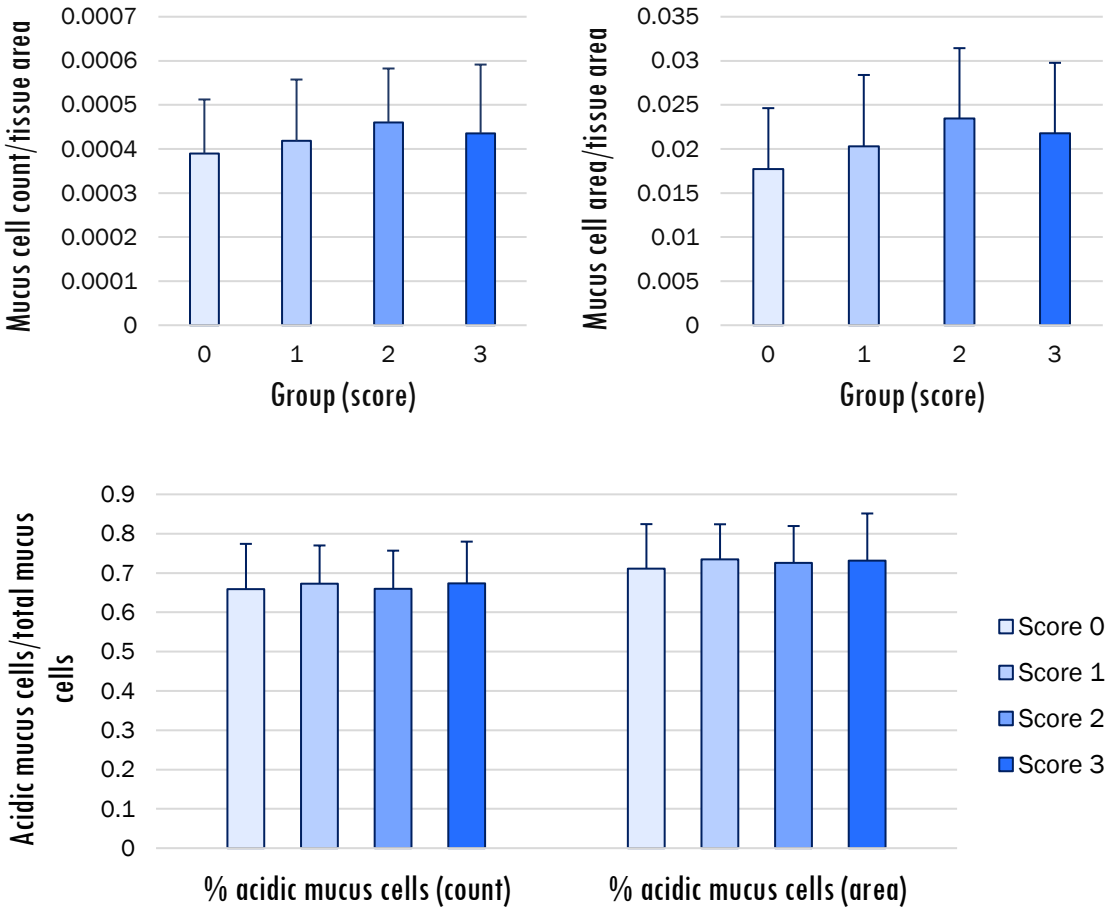

Figure 17 Characterization of the scoring groups in regard to mucus cell count per tissue area and total cell area per gill tissue area. Bars represent mean values across the groups, with standard deviation marked. Top left: mucus cell count per gill tissue area in each scoring group. Top right: total mucus cell area per gill tissue area. Bottom: ratio of cells stained for acidic mucins to the total stained cells. Ratio shown for the count of the cells and the area of the cells.

In Figure 17, weak trends between the scoring groups can be observed. Both mucus cell count and collective mucus cell area per gill tissue area had increasing mean value with higher severity of the abnormal mucus secretion score for groups 0, 1 and 2, but not for group with score "3", which exhibited a decreasing tendency for both parameters in regard to group with score "2". Statistical analysis indicated significant p-value only for analysis of mucus cell area per tissue area (Table 14**Feil! Ugyldig selvreferanse for bokmerke.**). Multiple comparisons between each group 1-4 for the mucus cell area indicated significant difference only between groups scored "0" and "2".

The ratio of cells stained for acidic mucins to the total detected cells was nearly the same for all groups, both for cell count and cell area based analysis, and no significance was reported (Table 14). Similar trends observed for both cell count and cell area-based analysis were expected, as there was a strong correlation between the cell count and the cumulative cell area (data not shown).

Table 14 Statistical analysis of groups derived from manual "Abnormal mucus secretion" score. Analysed parameters concerned the amount or cumulative area of mucus cells per gill tissue area or the ratio of acidic mucus cells to total mucus cells.

Parameter	Groups (N)	Mean (± SD)	Shaphiro-Wilk (p-value)	ANOVA/Kruskal-Wallis statistic	p-value (2-sided)
Mucus cell count per tissue area	0 (45)	0.00030 (±0.0001)	0.601	F(3) = 1.92	0.192
	1 (41)	0.00031 (±0.0001)	0.867		
	2 (42)	0.00036 (±0.0001)	0.968		
	3 (20)	0.00031 (±0.0002)	0.383		
Mucus cell area per tissue area	45	0.014 (±0.006)	0.081	**H(3) = 11.21**	**0.011**
	41	0.015 (±0.006)	0.016		
	42	0.018 (±0.007)	0.639		
	20	0.016 (±0.008)	0.847		
Acidic mucus cells count per total mucus cells	45	0.65 (±0.12)	0.278	F(3) = 0.554	0.646
	41	0.68 (±0.09)	0.159		
	42	0.66 (±0.10)	0.506		
	20	0.67 (±0.10)	0.994		
Acidic musuc cell area	45	0.69 (±0.12)	0.508	F(3) = 1.082	0.359
	41	0.74 (±0.09)	0.051		
	42	0.73 (±0.10)	0.530		
	20	0.72 (±0.12)	0.452		

The analysis of AB-PAS stain intensity values indicated only slight differences between the score-defined groups (Figure 18). The patterns for AB resembled the tendency observed for mucus cell count and mucus cell area (Figure 17). This could be explained by the fact that AB stain colours mucus cells and therefore with increased mucus cell amount in a tissue, AB intensity measured across the whole tissue would also increase. All measures of PAS channel intensity were alike across the groups. In order to verify the observations from the visual assessment of the plots (Figure 18), statistical analysis was run (data not shown). Between-group differences were compared using independent-samples Kruskal-Wallis test. Significant p-values were obtained for all AB parameters (p-values for AB mean, AB SD, AB maximum and AB median were 0.009, 0.005, 0.035 and 0.044, respectively). None of PAS parameters reached statistical significance. Further pairwise comparison of groups indicated significant differences only between groups 0 and 2 (AB mean, p = 0.009; AB SD, p = 0.003; AB maximum, p = 0.029).

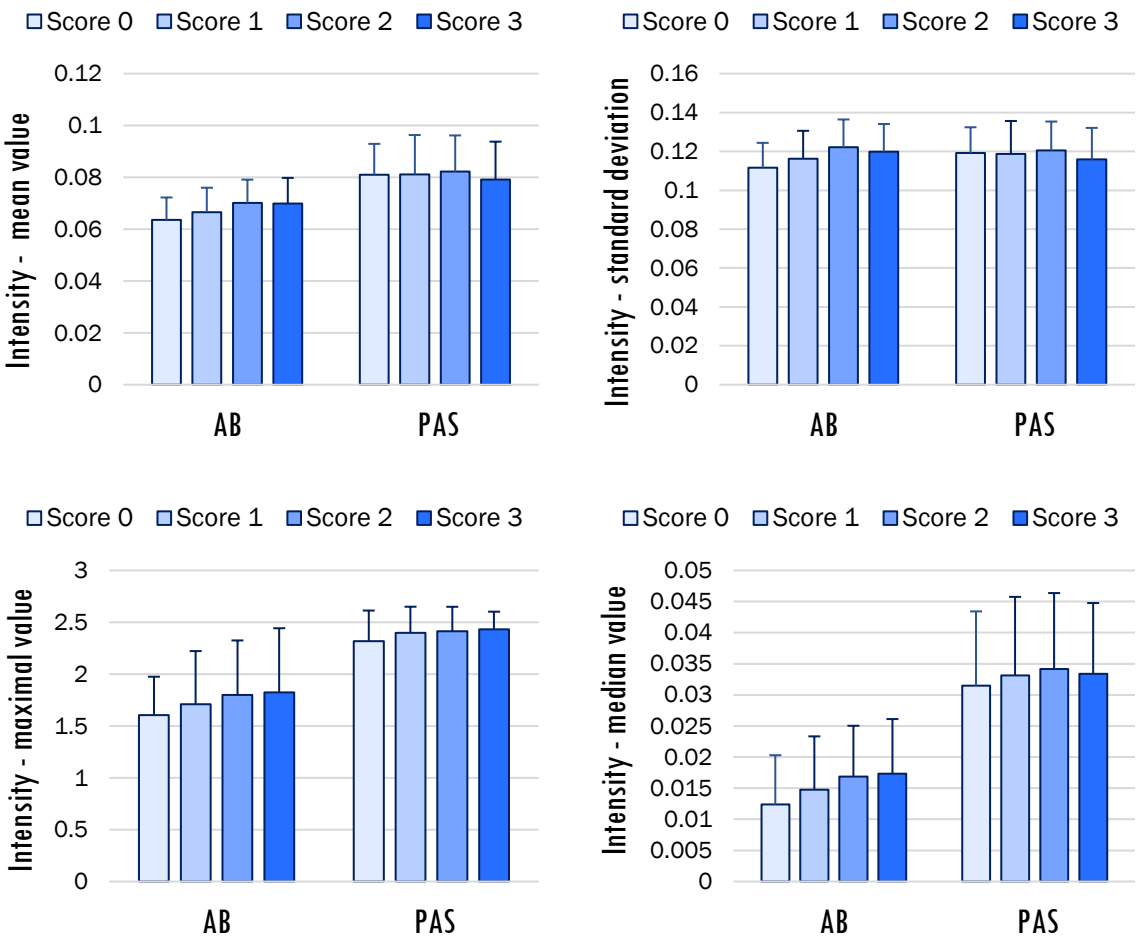

Figure 18 Characterization of the scoring groups in regard to Alcian Blue (AB) and Periodic Acid-Schiff (PAS) stains intensity values measured for the gill tissue area. Bars represent mean values across the groups, with standard deviation marked. Top left: mean stain values within the gill tissue area. Top right: standard deviation of stain values within the gill tissue area. Bottom left: maximal stain values within the gill tissue area. Bottom right: median stain values within the gill tissue area.

COMPARISON OF MUCUS CELL QUANTITY BETWEEN DIFFERENT MUSSEL RIG STATIONS

The analysis was performed to evaluate the differences in mucus cell abnormalities between the location-based groups. The locations were as described in Table 8.

The total of 199 images was filtered for the median value of optical density sum in order to remove images with improperly detected tissue area from further analysis. 21 images with the median value of 0 were removed from the dataset, leaving 178 images for the further analysis steps. Mean values for mucus cell count per tissue area, total mucus cell area per tissue area, percentage of acidic mucus cells count and area were compared between the groups. Results are shown in Table 21 in Appendix 4: Comparison of mucus cell quantity between different mussel rig stations. The distribution of values within and between groups is illustrated in box plots in Figure 19.

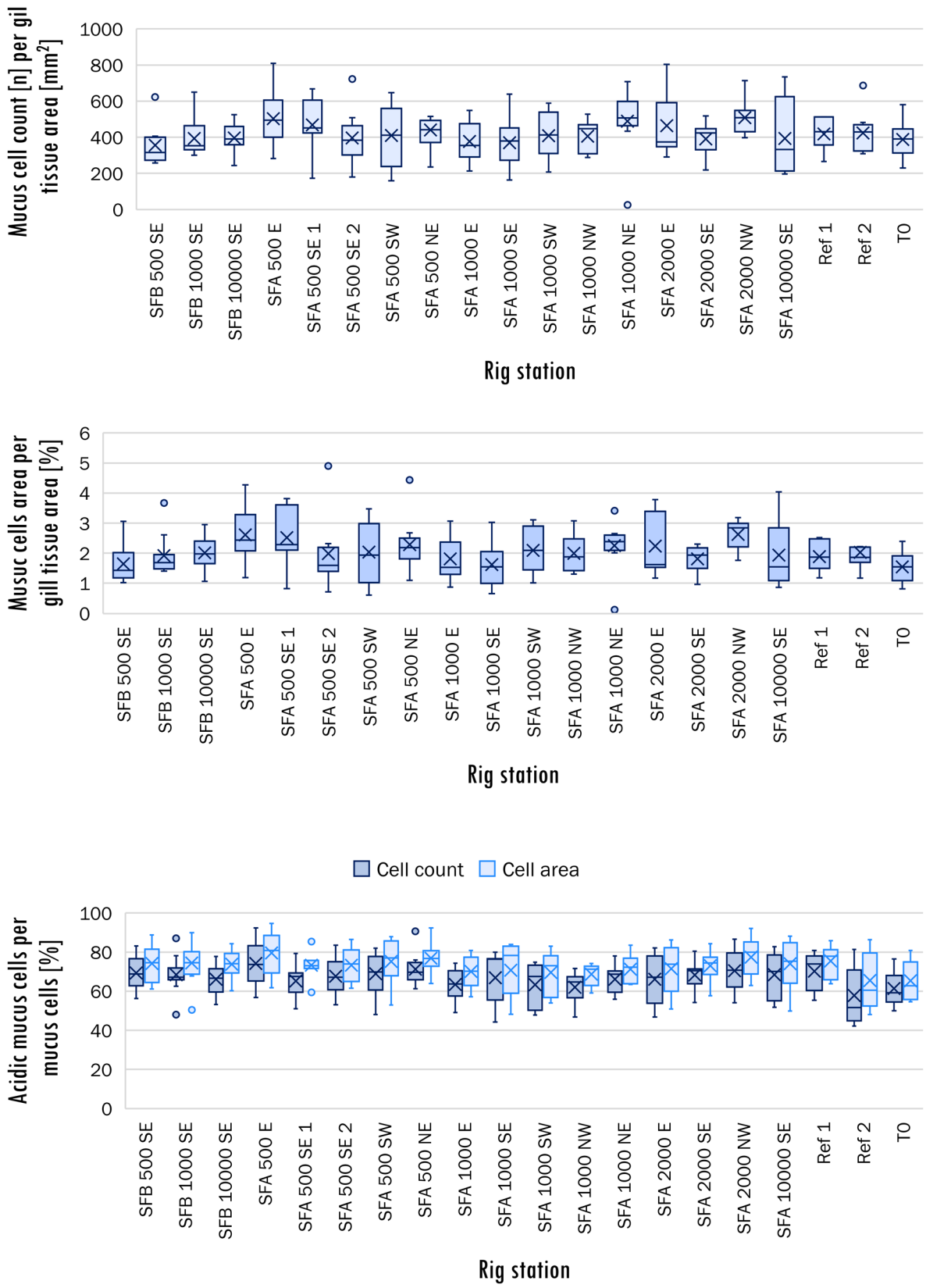

Figure 19 Comparison of mucus cell counts and cumulative mucus cell area between different mussel rig locations. Top: mucus cell count per gill tissue area [mm²]. Middle: cumulative area of all mucus cells per gill tissue area [%]. Bottom: ratio of acidic mucus cells (dark blue stained cells) to the total mucus cells [%] in regard to measures of cell count and cell area. T0 on the horizontal axis indicates the group of mussel individuals sampled prior to the exposure.

In order to illustrate the change of mucus cell parameters, a fold indication for to the T0 group (pre-exposure evaluation sample) was calculated. Results are reported in Figure 20. Analysis of statistical significance was also performed. The results are shown in

Table 22 in Appendix 4: Comparison of mucus cell quantity between different mussel rig stations, only for the mucus cell count per tissue area (results for other parameters not shown due to lack of significance). Stations which were significantly different from the reference stations are marked on the map in Figure 21. Figure 21 indicates increased mucus cell count in individuals from the stations located in close proximity to the offshore platforms, especially in SE and E directions.

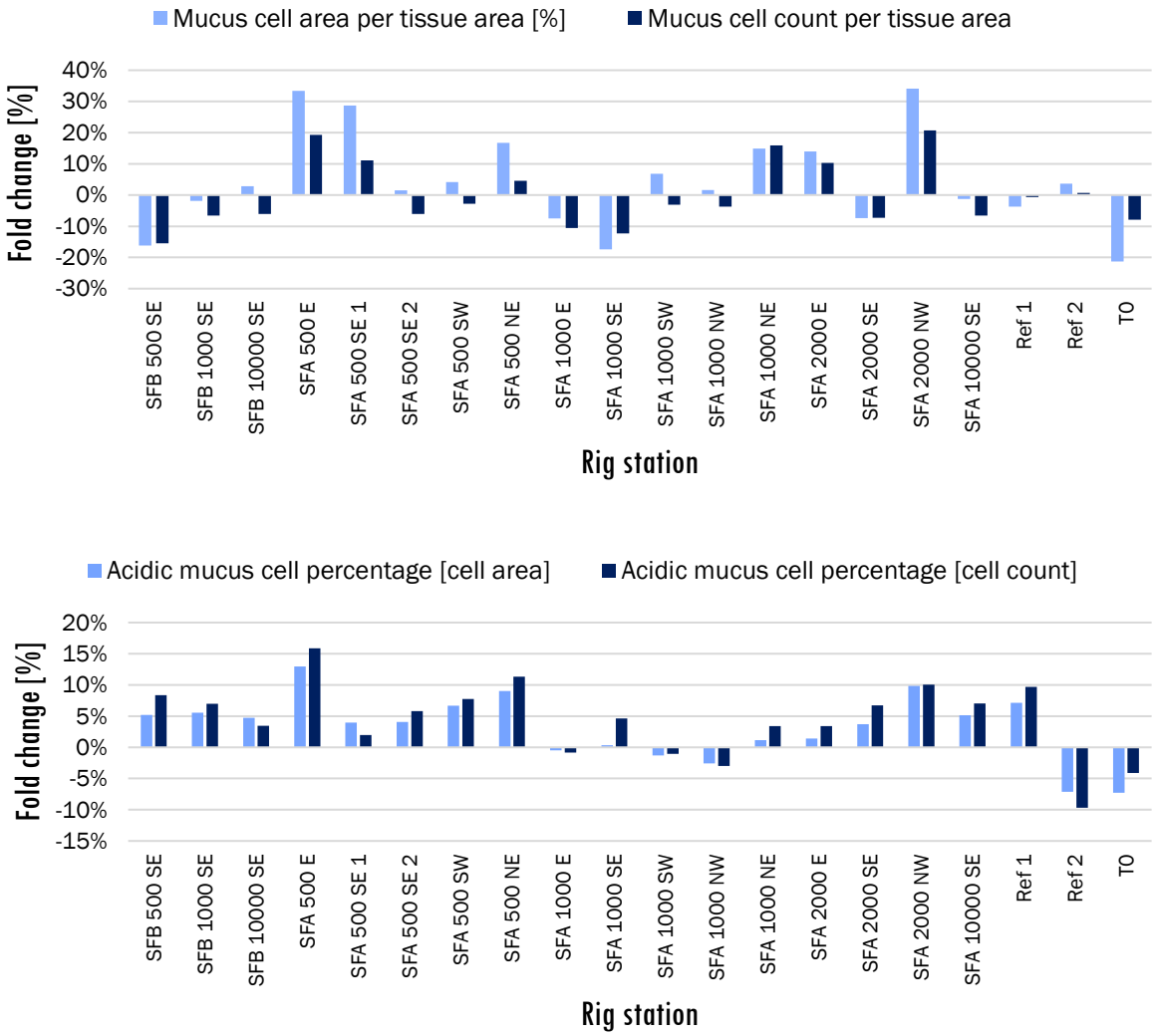

Figure 20 Fold change [%] of the mucus cell parameters in regard to mean of reference stations. Top: change in mucus cell count or area per gill tissue area. Bottom: change in acidic mucus cells ratio to the total mucus cells in respect to the cell count and the cumulative area. Standard deviation not shown.

Figure 20 shows that relative changes in mean mucus cell parameters differed across the mussel rig stations. The fold change in cumulative area of mucus cells per gill tissue area reached around 30% increase for stations SFA 500 E, SFA 500 SE1 and SFA 2000 NW, as compared to the reference stations. Similar trends were observed for mucus cell count per gill tissue area. Negative fold change of around 10% was seen for stations SFB 500 SE, SFA 1000 SE, SFA 1000 E, as well as for group T0 (pre-exposure group), which was expected. Trends illustrated using mucus cell area

and mucus cell count were alike, with more pronounced effects for the cell area-based measurements.

Differences between rig stations were also observed in the graph illustrating the percentage of acidic mucus cells of the total mucus cells. The general trend across the stations resembled to some extent the trend seen for the mucus cell content. Few stations (SFB 500 SE, SFA 2000 SE, SFA 10000 SE) had opposite fold change direction than the one seen for the mucus cell content. Interestingly, the T0 group was the only one that showed relatively high negative fold change for acidic mucus cell percentage.

Obtained results were compared with the concentration of contaminants measured in sampled individuals in the WCM 2017 program and a potential correlation was assessed. The correlation matrix analysis was carried out using MS Excel Data Analysis ToolPak.

Table 15 Heat map of correlations between the mucus cells-related parameters and the contaminants measured in *M. edulis*. Colours and intensities indicate the direction and the strength of the correlation: green – positive, red – negative. PAH – Polycyclic Aromatic Hydrocarbons; NPD – alkylated Naphthalenes, Phenanthrenes and Dibenzothiophenes; PAH EPA-16 – the list of 16 polycyclic aromatic hydrocarbons issued by the U.S. Environmental Protection Agency.

Chemical analysis parameter	Acidic mucus cell/total mucus cells [count]	Acidic mucus cells/total mucus cells [area]	Mucus cell count per tissue area	Mucus cell area per tissue area
Sum PAH	0.49	0.43	-0.16	0.00
PAH EPA-16	0.47	0.41	-0.21	-0.06
Sum NPD	0.49	0.43	-0.17	0.00
Total Naphthalenes	0.42	0.33	-0.32	-0.17
Total Alkylnaphtalenes (C1 - C3)	0.42	0.33	-0.32	-0.17
Ratio Alkylnaphtalenes / Naphthalene	0.20	0.15	-0.31	-0.18
Total Phenanthrenes/Anthracenes	0.52	0.50	0.07	0.24
Total Alkylphenanthrenes (C1 - C3)	0.62	0.59	0.07	0.22
Total Dibenzothiophenes	0.51	0.49	0.05	0.21
Ratio of Alkylphenanthrenes (C1 – C3) / Phenanthrene	0.56	0.59	0.22	0.42
Total Alkyldibenzothiophenes (C1 - C3)	0.52	0.50	0.05	0.21
Ratio Alkyldibenzothiophenes (C1 - C3)/ Dibenzothiophene	0.53	0.46	0.08	0.14

A modest correlations between the mucus cell quantification results and the concentration of contaminants were observed in correlation matrix in Table 15. All correlations between the ratios

of acidic mucus cell to total mucus cells were positive, regardless of the parameter. Slightly stronger positive correlations were estimated for chemical parameters concerning phenanthrenes, alkylphenanthrenes, dibenzothiophenes, alkyldibenzothiophenes and ratios of those. The strongest correlation (0.62 and 0.59) was estimated between the ratio of acidic mucus cells to total mucus cells and total alkylphenanthrenes (C1 – C3). The mucus cell count per tissue area had weak negative correlations with the parameters linked to PAH, NPD, naphthalene and alkylnaphtalenes. The strongest negative correlations were noted for total naphthalenes, total alkylnaphtalenes and the ratio of those, with the respective values of -0.32, -0.32 and -0.31. The negative correlation for these parameters was also observed for mucus cell area per tissue area. The remaining parameters had weak positive correlation with the mucus cell content.

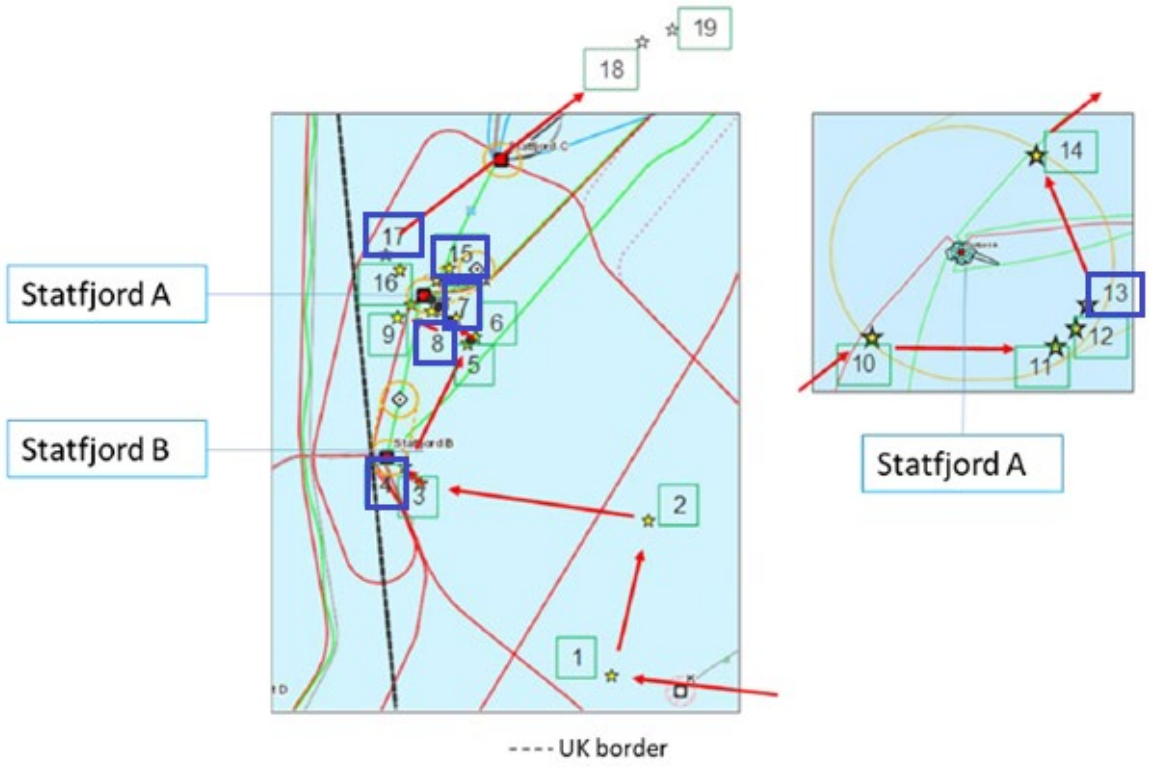

Figure 21 Map with rig stations of statistically significant difference in mucus cell count per tissue area as compared to reference groups. The stations marked with blue borders. Stations were compared to the mean of both reference stations.

DISCUSSION

GENDER CLASSIFICATION

The distribution of the gender is a basic parameter in characterizing the sample of organisms included in the environmental studies. In H&E colored histology slides of gonad tissue, differences between male and female gonads in *Mytilus edulis* are relatively easy to capture and therefore are a basis for traditional visual assessment. In the proposed method, the recognition of the gender based on measurements of image intensity values was proposed. The gender-stratified analysis of the intensity values indicated that the most suitable criterion for simple threshold method is the standard deviation of hematoxylin, which in our dataset was 100% efficient in dividing the images between male and female individuals. The standard deviation for other intensity values was also shown to be suitable, with an accuracy of classification higher than 97%. The principle of this method, based on a simple threshold, is easy and does not demand high computational power. Its simplicity, however, might implicate sensitivity to image quality and the variance in staining procedure.

In the second approach for data analysis, we used K-means clustering method. The classification of the gender rendered by the K-means algorithm was 98.94% correct. Even though few images were misclassified, the use of clustering algorithms, such as K-means seems to be a better alternative to simple threshold setting, being more prone to overfitting. Including multiple variables as classification basis can supply the method with more flexibility and better performance in a dataset with more variance within the groups.

A potential disadvantage of the developed method could be that the similar developmental stage of all gonads in the dataset (according to the visual assessment), which did not allow to verify whether the method would be suitable as universal for gender recognition.

An alternative method for gender recognition in *Mytilus edulis* could be based on learning algorithms: a pathologist would teach an algorithm the differences between female and male gonads by creating respective annotations. Based on those, the algorithm could be trained to recognize similar structures on a new set of images. The final gender assessment could involve a quantitative comparison of the detected structures of each class.

The automatized gender classification could be used as the first step in an integrated histopathological evaluation using digital image analysis. In this method, significant differences between male and female groups in regard to standard deviations of the 6 candidate features were observed. Although insignificant, partial stratification was observed also for other features. Therefore, analysis of histology images of male and female individuals could benefit from preceding

gender-classification, as separate methods could be developed and optimized for each gender to better recognize other structures, including lesions.

DETERMINATION OF GONADAL STATUS AND SPAWNING STAGES IN MALE MUSSELS

The assessment of gonadal status and spawning stages is important in characterizing of sampled individuals, as they might act as confounding factors during the analysis of other health parameters. Traditionally performed by the visual assessment, the determination of gonadal status and spawning stage is a laborious and time-consuming task. In this thesis work, the possible link between groups assigned by visual scoring of histology slides and the measurements obtained through digital analysis of corresponding images was investigated. Of all studied parameters: genital tissue percentage, spermatogonia percentage (of total cell count) and spermatogonia I percentage (of total spermatogonia cells count), the gonad content in a tissue seemed to be the most determinant (lowest p-values for comparison of groups across all studied parameters) of both gonadal status and spawning stage. Despite statistical significance in most of the group comparisons, the direct classification with simple threshold approach was impossible due to highly overlapping values between different groups in both gonadal status and spawning stage assessment.

The parameters suggested in the analysis concerned the genital tissue size and content of different cell types in the male gonad, which might have not fully addressed the basis required for gonadal status or spawning stage assessment through the traditional scoring system. Guidelines suggested by Seed (Seed, 1969; Seed and Brown, 1977) involve more detailed assessment of void space in the gonad, increasing during the spawning process, which was not quantified in our method. Quantification of void space within gonads could be one of the missing variables for a better classification. Another constraint of the presented analysis stems from the limited scope of gonadal status diversity in the dataset and unequal distribution across the groups: most of the individuals were in Stage 3 or Stage 4, and only a few in Stage 5 (according to manual scoring). More diversity could improve identifying correlation patterns across the groups and help identifying relevant parameters for the classification.

MUCUS CELL QUANTIFICATION IN THE GILL TISSUE

The potential of gill tissue histopathology in marine environmental pollution effects assessment stems from its direct contact with the seawater. Therefore, in polluted environment, epithelium of the gill tissue is a place where the adverse effects could manifest early after exposure. Mucus is thought to have important role in physiological processes in mussels (Bouallegui et al., 2017; David et al., 2008), especially in feeding and elimination of pseudofeces (Riisgård et al., 2011). Increase in both mucus cell number or size and the mucus secretion has previously been reported in other species when exposed to various types of contaminants, including PAH (David et al., 2008), oily

wastewater (Pirrone et al., 2018), preservative agents (Pagano et al., 2016). Upregulated mucus secretion in *Mytilus edulis* was shown to be induced by particulate matter (Jørgensen, 1981). Based on the literature, mucus cells are sensitive to environmental pollution and could be suitable for a routine assessment of quality of the environment. Long-known mucus-specific AB-PAS staining technique allows for mucus cell differentiation (Beninger and St-Jean, 1997) and thus, comprehensive analysis of changes in mucus cell population. In our study, a method to identify and classify the mucus cells in the gill tissue of *Mytilus edulis* was developed, with satisfactory outcomes, as illustrated in exemplary image in Figure 9. The results of mucus cell quantification were compared with groups derived from visual scoring for lesion "Abnormal mucus secretion". The differences were significant only between one pair of groups. Lack of significant associations was somewhat expected, since the visual scoring was performed using H&E stained slides, while digital analysis was done on AB-PAS colored images. The use of different stains was a major obstacle in evaluating the results of digital image analysis. However, it would be reasonable to expect that digital analysis method would be more accurate in detecting mucus cells, in addition to using images colored with AB-PAS, a stain that specifically targets mucins.

COMPARISON OF MUCUS CELL QUANTITY BETWEEN DIFFERENT MUSSEL RIG STATIONS

To further evaluate the mucus cell quantification, comparison of the results across the mussel rig stations from WCM 2017 program was performed. Across all mucus cell-related parameters, only the mucus cell count per gill tissue area reached significance levels in the comparison between rig stations and the reference stations. A trend in rig location and the significance was observed – vast majority of those stations were in close proximity to offshore platforms and SE/E direction. In order to further investigate potential correlations, comparison with concentration of contaminants in mussels was performed. 12 chemical parameters were included, concerning PAH, naphthalenes, phenanthrenes and dibenzothiophenes. The analysis indicated some degree of correlation between the mucus cell parameters and the contaminants. Interestingly, PAH, NPD and naphthalenes had negative correlation with mucus cell amount, while other contaminants did not (weak positive correlation was observed). The results of correlation analysis did not replicate previously reported increase in mucus cells in response to PAH exposure (David et al., 2008). The correlation analysis indicates that presence of different contaminants might have different, or even opposite effect on the mucus cell pathology (Table 15). Furthermore, positive correlation observed between all contaminants and the ratio of acidic mucus cells to all mucus cells demonstrates that exposure to pollution might induce secretion of specific types of mucins. Presence of other factors, such as particulate matter, which is known to elicit increased mucus secretion (Jørgensen, 1981; Riisgård et al., 2011), not analysed in this study, could act as a confounding factor in the analysis.

FINAL REMARKS AND FUTURE PERSPECTIVES

The use of digital pathology is increasingly gaining attention and applications in various fields and is aimed to improve the capacity, efficiency and accuracy of the diagnostic processes. In this thesis, the digital pathology was used in an attempt to develop methods that could be applied in the histopathological assessment of blue mussel, with emphasis on gonad and gill tissues. Chosen for the project was an open source image analysis software, QuPath, which was cited in 281 publications (as of 22.05.2020, according to QuPath GitHub website (Bankhead, 2020)), mostly in human histopathology. Up to date, no publications in marine species histopathology were reported to use QuPath in their methodologies, which makes this project unique. QuPath was first released at the end of 2016, and has been continuously updated through over 20 versions since, which makes it a dynamically developing environment for the digital image analysis. Its exceptional ability to handle WSI, the main factor for suitability for this project, makes it an exceptional choice among the available open source software. Taken together, QuPath would be a reasonable and cost-efficient option for the environmental monitoring-related histopathology analysis, especially in the early stage of method development, until a commercial-level software and quality assurance is desired.

Digital image analysis addresses the majority of obstacles encountered in traditional, visual assessment of histological biological markers and offers possibilities for automation of the image analysis process, standardization of the scoring systems and reduction of bias of human eye. The full potential of digital pathology can be explored by combining specialists from two main fields: pathology and digital image analysis. Most images in pathology are WSI: large images containing complex information. Analysis of images of such complexity might generate a large amount of data, and might therefore benefit from a data mining approach.

The image sample size in the project was around 200 images per tissue type. However, the unequal distribution of lesion severity across the sample made it difficult to perform comprehensive data analysis that ideally would include the whole spectrum of lesion severity.

In this study, digital image analysis was performed on both H&E stained images and images colored with more specific, AB-PAS stain. The advantage of widely used stains is that digital image analysis software often offers ready available settings or tools for their analysis. On the other hand, common stains might not be specific enough to allow for discrimination between types of cells or tissue pathologies. Digital pathology, due to accessibility to invisible data, might help in rediscovering the potential of common stains.

REFERENCES

Aarab, N., Godal, B. F., and Bechmann, R. K. (2011). “Seasonal variation of histopathological and histochemical markers of PAH exposure in blue mussel (Mytilus edulis L.),” *Marine environmental research*. Vol. 71, No. 3: pp. 213–217.

Aarab, N., Minier, C., Lemaire, S., Unruh, E., Hansen, P.-D., Larsen, B. K., Andersen, O.-K., and Narbonne, J.-F. (2004). “Biochemical and histological responses in mussel (Mytilus edulis) exposed to North Sea oil and to a mixture of North Sea oil and alkylphenols,” *Marine environmental research*. Vol. 58, 2-5: pp. 437–441.

Aarab, N., Pampanin, D. M., Naevdal, A., Oysaed, K. B., Gastaldi, L., and Bechmann, R. K. (2008). “Histopathology alterations and histochemistry measurements in mussel, Mytilus edulis collected offshore from an aluminium smelter industry (Norway),” *Marine pollution bulletin*. Vol. 57, 6-12: pp. 569–574.

Aeffner, F., Zarella, M. D., Buchbinder, N., Bui, M. M., Goodman, M. R., Hartman, D. J., Lujan, G. M., Molani, M. A., Parwani, A. V., Lillard, K., Turner, O. C., Vemuri, V. N. P., Yuil-Valdes, A. G., and Bowman, D. (2019). “Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association,” *Journal of pathology informatics*. Vol. 10, p. 9.

Au, D. W. T. (2004). “The application of histo-cytopathological biomarkers in marine pollution monitoring: a review,” *Marine pollution bulletin*. Vol. 48, 9-10: pp. 817–834.

Bankhead, P. (2020). *Citing QuPath*, https://github.com/qupath/qupath/wiki/Citing-QuPath. Accessed 24 May 2020.

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., and Hamilton, P. W. (2017). “QuPath: Open source software for digital pathology image analysis,” *Scientific reports*. Vol. 7, No. 1: p. 16878.

Ben Ameur, W., Lapuente, J. de, El Megdiche, Y., Barhoumi, B., Trabelsi, S., Camps, L., Serret, J., Ramos-López, D., Gonzalez-Linares, J., Driss, M. R., and Borràs, M. (2012). “Oxidative stress, genotoxicity and histopathology biomarker responses in mullet (Mugil cephalus) and sea bass (Dicentrarchus labrax) liver from Bizerte Lagoon (Tunisia),” *Marine pollution bulletin*. Vol. 64, No. 2: pp. 241–251.

Beninger, P. G., and St-Jean, S. D. (1997). “Particle processing on the labial palps of Mytilus edulis and Placopecten magellanicus (Mollusca: Bivalvia),” *Marine Ecology Progress Series*. Vol. 147, pp. 117–127.

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A., and Kreshuk, A. (2019). "ilastik: interactive machine learning for (bio)image analysis," *Nature methods*. Vol. 16, No. 12: pp. 1226–1232.

Beyer, J., Aarab, N., Tandberg, A. H., Ingvarsdottir, A., Bamber, S., Børseth, J. F., Camus, L., and Velvin, R. (2013). "Environmental harm assessment of a wastewater discharge from Hammerfest LNG: a study with biomarkers in mussels (Mytilus sp.) and Atlantic cod (Gadus morhua)," *Marine pollution bulletin*. Vol. 69, 1-2: pp. 28–37.

Beyer, J., Green, N. W., Brooks, S., Allan, I. J., Ruus, A., Gomes, T., Bråte, I. L. N., and Schøyen, M. (2017). "Blue mussels (Mytilus edulis spp.) as sentinel organisms in coastal pollution monitoring: A review," *Marine environmental research*. Vol. 130, pp. 338–365.

Bignell, J. P., Dodge, M. J., Feist, S. W., Lyons, B., Martin, P. D., Taylor, N. G.H., Stone, D., Travalent, L., and Stentiford, G. D. (2008). "Mussel histopathology: effects of season, disease and species," *Aquatic Biology*. Vol. 2, pp. 1–15.

Bignell, J. P., Stentiford, G. D., Taylor, N. G. H., and Lyons, B. P. (2011). "Histopathology of mussels (Mytilus sp.) from the Tamar estuary, UK," *Marine environmental research*. Vol. 72, 1-2: pp. 25–32.

Bouallegui, Y., Ben Younes, R., Bellamine, H., and Oueslati, R. (2017). "Histopathology and analyses of inflammation intensity in the gills of mussels exposed to silver nanoparticles: role of nanoparticle size, exposure time, and uptake pathways," *Toxicology mechanisms and methods*. Vol. 27, No. 8: pp. 582–591.

Chaumont, F. de, Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Meas-Yedid, V., Pankajakshan, P., Lecomte, T., Le Montagner, Y., Lagache, T., Dufour, A., and Olivo-Marin, J.-C. (2012). "Icy: an open bioimage informatics platform for extended reproducible research," *Nature methods*. Vol. 9, No. 7: pp. 690–696.

Cuevas, N., Zorita, I., Costa, P. M., Franco, J., and Larreta, J. (2015). "Development of histopathological indices in the digestive gland and gonad of mussels: integration with contamination levels and effects of confounding factors," *Aquatic toxicology (Amsterdam, Netherlands)*. Vol. 162, pp. 152–164.

Danuser, G. (2011). "Computer vision in cell biology," *Cell*. Vol. 147, No. 5: pp. 973–978.

David, J. A. d. O., Salaroli, R. B., and Fontanetti, C. S. (2008). "The significance of changes in Mytella falcata (Orbigny, 1842) gill filaments chronically exposed to polluted environments," *Micron (Oxford, England : 1993)*. Vol. 39, No. 8: pp. 1293–1299.

Duinker, A., Håland, L., Hovgaard, P., and Mortensen, S. (2008). "Gonad development and spawning in one and two year old mussels (Mytilus edulis) from Western Norway," *Journal of the Marine Biological Association of the United Kingdom*. Vol. 88, No. 7: pp. 1465–1473.

Gaitán-Espitia, J. D., Quintero-Galvis, J. F., Mesas, A., and D'Elía, G. (2016). "Mitogenomics of southern hemisphere blue mussels (Bivalvia: Pteriomorphia): Insights into the evolutionary characteristics of the Mytilus edulis complex," *Scientific reports*. Vol. 6, p. 26853.

Goldberg, I. G., Allan, C., Burel, J.-M., Creager, D., Falconi, A., Hochheiser, H., Johnston, J., Mellen, J., Sorger, P. K., and Swedlow, J. R. (2005). "The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging," *Genome biology*. Vol. 6, No. 5: R47.

Gosling, E. M. (2015). *Marine bivalve molluscs*, Chichester, West Sussex: Wiley Blackwell.

Hansson, N., Halling, M., and Norin, H. (2013). *Biomarkers for Environmental Monitoring*: *Suggestions for Norwegian monitoring programmes*, Trondheim.

Hylland, K., Burgeot, T., Martínez-Gómez, C., Lang, T., Robinson, C. D., Svavarsson, J., Thain, J. E., Vethaak, A. D., and Gubbins, M. J. (2017). "How can we quantify impacts of contaminants in marine ecosystems? The ICON project," *Marine environmental research*. Vol. 124, pp. 2–10.

Jørgensen, C. B. (1981). "Feeding and cleaning mechanisms in the suspension feeding bivalve (Mytilus edulis)," *Marine Biology*. Vol. 65, pp. 159–163.

Lamprecht, M. R., Sabatini, D. M., and Carpenter, A. E. (2007). "CellProfiler: free, versatile software for automated biological image analysis," *BioTechniques*. Vol. 42, No. 1: pp. 71–75.

López-Galindo, C., Vargas-Chacoff, L., Nebot, E., Casanueva, J. F., Rubio, D., Mancera, J. M., and Solé, M. (2010). "Sublethal responses of the common mussel (Mytilus galloprovincialis) exposed to sodium hypochlorite and Mexel432 used as antifoulants," *Ecotoxicology and environmental safety*. Vol. 73, No. 5: pp. 825–834.

Meijer, G. A., Beliën, J. A., van Diest, P. J., and Baak, J. P. (1997). "Origins of … image analysis in clinical pathology," *Journal of clinical pathology*. Vol. 50, No. 5: pp. 365–370.

Myers, R. PhD (2018). *Special Stain Techniques for the Evaluation of Mucins*, https://www.leicabiosystems.com/knowledge-pathway/special-stain-techniques-for-the-evaluation-of-mucins/. Accessed 20 April 2020.

Niazi, M. K. K., Parwani, A. V., and Gurcan, M. N. (2019). "Digital pathology and artificial intelligence," *The Lancet Oncology*. Vol. 20, No. 5: e253-e261.

Nogarol, L. R., Brossi-Garcia, A. L., Souza, R. B. de, and Fontanetti, C. S. (2016). “Histopathological effects of the herbicide atrazine on gills of the Brazilian endemic bivalve Diplodon expansus,” *International Journal of Environmental Analytical Chemistry*. Vol. 96, No. 4: pp. 387–403.

Pagano, M., Capillo, G., Sanfilippo, M., Palato, S., Trischitta, F., Manganaro, A., and Faggio, C. (2016). “Evaluation of Functionality and Biological Responses of Mytilus galloprovincialis after Exposure to Quaternium-15 (Methenamine 3-Chloroallylochloride),” *Molecules (Basel, Switzerland)*. Vol. 21, No. 2: p. 144.

Pampanin, D., Brooks, S., Grøsvik, B. E., and Sanni, S. (2019). *Water Column Monitoring 2017*: *Environmental monitoring of petroleum activities on the Norwegian continental shelf 2017*, Stavanger.

Pirrone, C., Rossi, F., Cappello, S., Borgese, M., Mancini, G., Bernardini, G., and Gornati, R. (2018). “Evaluation of biomarkers in Mytilus galloprovincialis as an integrated measure of biofilm-membrane bioreactor (BF-MBR) system efficiency in mitigating the impact of oily wastewater discharge to marine environment: a microcosm approach,” *Aquatic toxicology (Amsterdam, Netherlands)*. Vol. 198, pp. 49–62.

Riisgård, H. U., Egede, P. P., and Barreiro Saavedra, I. (2011). “Feeding Behaviour of the Mussel, Mytilus edulis : New Observations, with a Minireview of Current Knowledge,” *Journal of Marine Biology*. Vol. 2011, No. 3: pp. 1–13.

Rossi, F., Palombella, S., Pirrone, C., Mancini, G., Bernardini, G., and Gornati, R. (2016). “Evaluation of tissue morphology and gene expression as biomarkers of pollution in mussel Mytilus galloprovincialis caging experiment,” *Aquatic toxicology (Amsterdam, Netherlands)*. Vol. 181, pp. 57–66.

Santos, D. M. S., Melo, M. R. S., Mendes, D. C. S., Rocha, I. K. B. S., Silva, J. P. L., Cantanhêde, S. M., and Meletti, P. C. (2014). “Histological changes in gills of two fish species as indicators of water quality in Jansen Lagoon (São Luís, Maranhão State, Brazil),” *International journal of environmental research and public health*. Vol. 11, No. 12: pp. 12927–12937.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). “Fiji: an open-source platform for biological-image analysis,” *Nature methods*. Vol. 9, No. 7: pp. 676–682.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). “NIH Image to ImageJ: 25 years of image analysis,” *Nature methods*. Vol. 9, No. 7: pp. 671–675.

Seed, R. (1969). “The ecology of Mytilus edulis L. (Lamellibranchiata) on exposed rocky shores,” *Oecologia*. Vol. 3, 3-4: pp. 277–316.

Seed, R., and Brown, R. A. (1977). “A comparison of the reproductive cycles of Modiolus modiolus (L.), Cerastoderma (=Cardium) edule (L.), and Mytilus edulis L. in Strangford Lough, Northern Ireland,” *Oecologia*. Vol. 30, No. 2: pp. 173–188.

Snead, D. R. J., Tsang, Y.-W., Meskiri, A., Kimani, P. K., Crossman, R., Rajpoot, N. M., Blessing, E., Chen, K., Gopalakrishnan, K., Matthews, P., Momtahan, N., Read-Jones, S., Sah, S., Simmons, E., Sinha, B., Suortamo, S., Yeo, Y., El Daly, H., and Cree, I. A. (2016). “Validation of digital pathology imaging for primary histopathological diagnosis,” *Histopathology*. Vol. 68, No. 7: pp. 1063–1072.

Tizhoosh, H. R., and Pantanowitz, L. (2018). “Artificial Intelligence and Digital Pathology: Challenges and Opportunities,” *Journal of pathology informatics*. Vol. 9, p. 38.

Yancheva, V., Velcheva, I., Stoyanova, S., and Georgieva, E. (2016). “Histological biomarkers in fish as a tool in ecological risk assessment and monitoring programs: A review,” *Applied Ecology and Environmental Research*. Vol. 14, No. 1: pp. 47–75.

Zuraw, A. *6 free open source software programs for image analysis of pathology slides*: *6 most popular free open source image analysis programs which can be used in pathology*, https://digitalpathologyconsulting.com/6-free-open-source-software-programs-for-image-analysis-of-pathology-slides/. Accessed 2020, Apr 12.

APPENDICES

APPENDIX 1: DETERMINATION OF THE MUSSEL'S GENDER USING DIGITAL IMAGE PROPERTIES

R code used for K-means clustering analysis:

```
#data scaling/standarisation
data_all_standarized <- select(data_all, -"Image", -"Gender")
data_all_scaled <- scale(data_all_standarized)
#determining the optimal cluster no
k.max <- 15
wss <- sapply(1:k.max, function(k){kmeans(data_all_scaled, k,
nstart=10)$tot.withinss})
plot(1:k.max, wss, type="b", main="Determination of the optimal number of
clusters by Elbow Method", pch=19, xlab="Number of clusters K", ylab="Total
within clusters sum of squares")
#silhouette method
library(purrr)
avg_sil <- function(k){
  km.res <- kmeans(data_all_scaled, centers=k, nstart=10)
  ss <- silhouette(km.res$cluster, dist(data_all_scaled))
  mean(ss[, 3])
}
k.values <- 2:15
avg_sil_values <- map_dbl(k.values, avg_sil)
plot(k.values, avg_sil_values, type="b", pch=19, main="Determination of the
optimal number of clusters by the Average Silhouette Method", xlab="Number
of clusters K", ylab="Average silhouettes width")
#kmeans algorithm
cluster_scale <- kmeans(data_all_scaled, 2, nstart=25)
data_all_clusters <- mutate(data_all, Cluster = cluster_scale$cluster)
plot_clusters_scaled <- clusplot(data_all_scaled, cluster_scale$cluster,
main="K-means clustering", color=TRUE, shade=TRUE, labels=1)
fviz_cluster(cluster_scale, data = data_all_scaled)
```

Table 16 Dataset used in the gender classification and the results of the K-means clustering. Rows shaded blue indicate the images with incorrect classification. Original intensity features values shown. Gender - classification based on the visual assessment; Cluster – classification according to K-means clustering.

Image	OD_SD	Hem_SD	Bright_SD	Red_SD	Green_SD	Blue_SD	Gender	Cluster
G0005	0.256	0.089	0.118	31.49	41.09	29.85	F	2
G0006	0.212	0.076	0.101	27.24	34.70	25.35	F	2
G0008	0.241	0.084	0.108	28.80	36.75	27.30	F	2
G0014	0.234	0.081	0.109	29.03	38.62	27.51	F	2
G0019	0.218	0.078	0.104	27.78	35.29	26.29	F	2
G0021	0.262	0.088	0.118	31.11	42.22	30.13	F	2
G0031	0.259	0.092	0.116	31.57	41.39	29.37	F	2
G0032	0.238	0.086	0.111	30.10	38.53	28.02	F	2

G0034	0.206	0.074	0.099	26.66	34.27	25.17	F	2
G0038	0.220	0.079	0.102	27.79	35.38	25.87	F	2
G0040	0.206	0.074	0.098	26.78	34.68	25.00	F	2
G0043	0.211	0.075	0.099	26.69	34.80	24.98	F	2
G0044	0.221	0.078	0.101	27.49	36.84	25.45	F	2
G0049	0.201	0.074	0.094	25.79	32.21	23.75	F	2
G0052	0.261	0.087	0.113	30.26	42.79	28.99	F	2
G0059	0.279	0.095	0.122	32.79	44.68	30.99	F	2
G0065	0.247	0.087	0.111	29.99	39.44	28.08	F	2
G0071	0.206	0.073	0.099	26.57	34.39	25.07	F	2
G0072	0.254	0.090	0.113	30.85	40.89	28.65	F	2
G0073	0.202	0.072	0.095	25.80	33.81	24.03	F	2
G0077	0.239	0.082	0.111	29.33	37.70	28.04	F	2
G0079	0.222	0.078	0.104	27.84	37.03	26.28	F	2
G0086	0.230	0.080	0.105	28.35	38.09	26.72	F	2
G0087	0.221	0.078	0.103	27.83	36.73	26.03	F	2
G0089	0.245	0.087	0.112	30.13	38.35	28.33	F	2
G0091	0.236	0.082	0.111	29.47	38.33	28.29	F	2
G0093	0.209	0.076	0.101	27.18	33.94	25.51	F	2
G0094	0.204	0.073	0.098	26.32	33.74	24.74	F	2
G0097	0.190	0.068	0.091	24.64	32.75	22.77	F	2
G0099	0.216	0.078	0.102	27.65	36.03	25.61	F	2
G0100	0.228	0.083	0.107	29.26	37.17	27.13	F	2
G0102	0.221	0.080	0.106	28.48	35.37	26.67	F	2
G0103	0.251	0.088	0.115	30.78	40.15	29.06	F	2
G0104	0.256	0.103	0.112	32.81	40.57	28.27	F	2
G0107	0.202	0.072	0.094	25.78	35.81	23.75	F	2
G0108	0.226	0.082	0.106	28.73	36.87	26.69	F	2
G0112	0.231	0.081	0.108	28.91	37.89	27.45	F	2
G0115	0.231	0.080	0.105	28.21	38.37	26.61	F	2
G0117	0.232	0.081	0.106	28.73	39.40	27.06	F	2
G0122	0.204	0.072	0.093	25.58	35.90	23.68	F	2
G0125	0.196	0.068	0.092	24.97	35.14	23.43	F	2
G0126	0.227	0.072	0.103	27.03	39.34	26.90	F	2
G0127	0.207	0.075	0.098	26.55	33.79	24.67	F	2
G0131	0.214	0.075	0.103	27.39	35.62	26.23	F	2
G0132	0.238	0.084	0.112	29.63	38.11	28.43	F	2

G0134	0.226	0.082	0.105	28.58	36.04	26.55	F	2
G0135	0.209	0.073	0.098	26.26	35.19	24.92	F	2
G0138	0.207	0.075	0.099	26.58	33.65	25.07	F	2
G0140	0.243	0.084	0.112	29.73	39.18	28.53	F	2
G0145	0.222	0.079	0.102	28.04	38.44	25.76	F	2
G0149	0.228	0.080	0.109	28.91	37.04	27.52	F	2
G0150	0.217	0.075	0.101	26.93	36.27	25.87	F	2
G0156	0.187	0.068	0.089	24.21	30.83	22.46	F	2
G0159	0.183	0.068	0.088	24.25	30.59	22.26	F	2
G0166	0.201	0.070	0.098	25.89	33.49	24.80	F	2
G0173	0.204	0.072	0.098	26.18	33.79	24.98	F	2
G0175	0.210	0.075	0.101	26.85	33.45	25.47	F	2
G0178	0.218	0.076	0.104	27.64	35.43	26.26	F	2
G0179	0.225	0.077	0.105	27.88	36.79	26.53	F	2
G0181	0.226	0.076	0.104	27.66	37.96	26.48	F	2
G0187	0.210	0.074	0.102	26.92	34.31	25.70	F	2
G0189	0.226	0.081	0.103	28.37	37.48	26.15	F	2
G0191	0.233	0.082	0.108	29.10	37.58	27.32	F	2
G0192	0.213	0.074	0.102	27.06	34.85	25.76	F	2
G0193	0.214	0.075	0.103	27.35	35.27	26.28	F	2
G0196	0.199	0.073	0.094	25.89	33.36	23.79	F	2
G0197	0.221	0.079	0.102	27.90	36.71	25.92	F	2
G0205	0.191	0.069	0.095	25.29	31.15	23.87	F	2
G0209	0.222	0.081	0.104	28.37	36.02	26.29	F	2
G0212	0.182	0.067	0.084	23.14	29.63	21.26	F	2
G0213	0.225	0.081	0.106	28.86	36.79	27.00	F	2
G0216	0.236	0.083	0.107	29.10	38.64	27.18	F	2
G0217	0.238	0.086	0.111	29.97	37.60	27.92	F	2
G0219	0.253	0.089	0.116	31.00	39.98	29.34	F	2
G0221	0.248	0.087	0.113	30.32	40.11	28.44	F	2
G0222	0.207	0.072	0.096	25.92	34.96	24.33	F	2
G0229	0.321	0.111	0.136	36.64	47.83	34.42	F	2
G0230	0.223	0.078	0.105	28.19	37.13	26.67	F	2
G0237	0.244	0.086	0.110	29.84	39.62	27.98	F	2
G0239	0.256	0.087	0.114	30.33	41.29	29.04	F	2
G0244	0.342	0.114	0.139	37.25	50.02	35.44	F	1
G0247	0.240	0.087	0.110	30.11	38.97	27.78	F	2

G0249	0.233	0.082	0.108	28.99	37.32	27.20	F	2
G0251	0.256	0.089	0.114	31.04	42.66	28.84	F	2
G0255	0.245	0.085	0.112	29.96	39.81	28.14	F	2
G0262	0.210	0.074	0.097	26.15	34.49	24.55	F	2
G0263	0.263	0.090	0.120	31.95	41.73	30.82	F	2
G0265	0.205	0.073	0.095	26.14	34.66	24.21	F	2
G0268	0.286	0.101	0.123	33.89	44.23	31.43	F	2
G0272	0.253	0.090	0.115	31.24	39.85	29.26	F	2
G0003	0.478	0.180	0.175	48.92	58.74	44.22	M	1
G0009	0.445	0.168	0.167	46.44	54.98	42.46	M	1
G0010	0.459	0.184	0.176	49.01	53.57	44.59	M	1
G0011	0.425	0.167	0.164	46.24	53.05	41.74	M	1
G0012	0.467	0.181	0.174	48.38	55.23	44.09	M	1
G0013	0.460	0.178	0.171	47.62	54.88	43.40	M	1
G0016	0.494	0.194	0.184	51.43	58.03	46.73	M	1
G0017	0.508	0.196	0.179	50.16	57.29	45.24	M	1
G0018	0.427	0.165	0.162	45.71	54.26	41.27	M	1
G0024	0.456	0.176	0.165	46.09	53.05	41.85	M	1
G0026	0.455	0.179	0.175	48.51	54.19	44.46	M	1
G0027	0.361	0.139	0.146	40.55	47.21	36.99	M	1
G0030	0.435	0.176	0.170	47.61	52.58	43.10	M	1
G0033	0.355	0.134	0.144	40.03	49.07	36.72	M	1
G0035	0.451	0.180	0.177	49.08	54.10	44.92	M	1
G0036	0.431	0.171	0.165	46.40	52.77	41.64	M	1
G0039	0.464	0.183	0.173	48.24	54.10	43.97	M	1
G0042	0.443	0.176	0.170	47.60	53.23	43.09	M	1
G0046	0.467	0.181	0.173	48.45	56.43	43.72	M	1
G0047	0.405	0.162	0.157	44.60	51.05	39.73	M	1
G0056	0.425	0.167	0.162	45.47	51.73	40.96	M	1
G0057	0.470	0.186	0.178	49.83	56.71	45.13	M	1
G0058	0.495	0.187	0.180	49.86	58.23	45.55	M	1
G0062	0.348	0.137	0.143	40.06	45.93	36.22	M	1
G0066	0.346	0.132	0.143	39.16	46.47	36.19	M	1
G0067	0.431	0.167	0.164	45.78	52.93	41.49	M	1
G0069	0.337	0.129	0.135	38.18	47.53	34.36	M	1
G0074	0.416	0.161	0.160	45.00	52.77	40.51	M	1
G0075	0.432	0.168	0.166	46.02	52.44	42.10	M	1

G0078	0.395	0.155	0.155	43.44	49.68	39.30	M	1
G0080	0.503	0.191	0.182	50.36	57.89	46.07	M	1
G0083	0.449	0.173	0.172	47.65	54.89	43.55	M	1
G0084	0.447	0.169	0.169	46.98	55.72	42.57	M	1
G0090	0.522	0.193	0.181	49.79	57.36	45.99	M	1
G0101	0.410	0.159	0.155	44.90	56.40	39.34	M	1
G0106	0.362	0.141	0.148	41.17	49.19	37.74	M	1
G0110	0.428	0.173	0.165	46.61	52.42	41.62	M	1
G0111	0.395	0.160	0.154	43.88	49.75	38.94	M	1
G0113	0.496	0.195	0.182	51.07	57.99	46.01	M	1
G0114	0.413	0.160	0.163	45.30	52.64	41.25	M	1
G0121	0.484	0.189	0.181	50.40	56.77	45.84	M	1
G0129	0.410	0.161	0.159	44.87	52.01	40.06	M	1
G0136	0.374	0.148	0.155	42.79	48.70	39.19	M	1
G0143	0.445	0.177	0.170	47.19	51.41	43.24	M	1
G0144	0.393	0.156	0.158	44.27	50.46	40.09	M	1
G0146	0.427	0.165	0.164	45.67	52.43	41.77	M	1
G0147	0.439	0.169	0.167	46.35	53.26	42.42	M	1
G0152	0.371	0.144	0.149	42.14	49.68	37.83	M	1
G0154	0.402	0.158	0.162	45.25	52.18	40.92	M	1
G0155	0.360	0.142	0.147	41.47	47.98	37.10	M	1
G0158	0.550	0.216	0.195	53.99	58.98	49.29	M	1
G0160	0.419	0.169	0.162	45.83	51.31	41.09	M	1
G0161	0.442	0.171	0.167	46.64	52.99	42.27	M	1
G0163	0.451	0.177	0.170	47.40	53.01	42.82	M	1
G0165	0.453	0.176	0.175	48.19	53.68	44.37	M	1
G0168	0.466	0.181	0.173	49.40	59.82	43.80	M	1
G0170	0.410	0.160	0.157	45.22	55.92	39.91	M	1
G0171	0.532	0.207	0.191	53.59	62.76	48.37	M	1
G0174	0.470	0.183	0.179	49.12	53.65	45.51	M	1
G0177	0.405	0.159	0.158	44.25	50.31	39.94	M	1
G0185	0.443	0.171	0.167	47.67	58.34	42.26	M	1
G0190	0.390	0.152	0.154	43.18	51.31	39.15	M	1
G0194	0.451	0.180	0.174	48.61	53.69	44.23	M	1
G0195	0.391	0.149	0.149	41.86	52.57	37.85	M	1
G0199	0.411	0.168	0.161	45.51	49.80	40.76	M	1
G0202	0.357	0.139	0.146	40.46	47.53	37.11	M	1

G0204	0.472	0.188	0.179	50.13	56.26	45.27	M	1
G0206	0.464	0.185	0.176	49.23	54.66	44.60	M	1
G0207	0.453	0.178	0.169	46.99	52.28	42.91	M	1
G0210	0.447	0.178	0.170	47.98	54.42	42.90	M	1
G0214	0.386	0.152	0.150	42.45	50.31	37.72	M	1
G0218	0.441	0.174	0.167	47.17	54.11	42.32	M	1
G0225	0.408	0.161	0.155	44.06	50.97	39.37	M	1
G0226	0.498	0.194	0.178	51.41	62.61	45.07	M	1
G0228	0.468	0.190	0.174	49.12	53.97	43.90	M	1
G0233	0.470	0.187	0.174	49.79	58.76	44.03	M	1
G0234	0.509	0.201	0.181	51.18	58.65	45.76	M	1
G0235	0.499	0.199	0.177	50.38	57.10	44.72	M	1
G0238	0.471	0.183	0.167	46.95	53.41	42.42	M	1
G0243	0.541	0.208	0.183	51.37	58.37	46.36	M	1
G0246	0.517	0.204	0.184	51.10	55.98	46.66	M	1
G0250	0.550	0.214	0.191	53.24	59.65	48.08	M	1
G0252	0.473	0.184	0.173	48.26	54.73	43.69	M	1
G0254	0.495	0.194	0.177	49.56	55.04	44.85	M	1
G0257	0.494	0.193	0.178	49.59	55.42	45.23	M	1
G0258	0.512	0.198	0.182	52.09	62.61	46.11	M	1
G0261	0.485	0.192	0.176	50.74	60.29	44.60	M	1
G0267	0.523	0.203	0.183	51.37	58.53	46.33	M	1
G0269	0.490	0.193	0.174	49.55	57.26	43.95	M	1
G0270	0.477	0.187	0.172	50.21	61.71	43.66	M	1
G0271	0.536	0.205	0.183	50.70	56.95	46.35	M	1
G0273	0.483	0.187	0.176	49.08	55.49	44.67	M	1
G0274	0.306	0.120	0.130	36.26	42.71	33.00	M	2
G0277	0.436	0.177	0.165	46.96	52.34	41.73	M	1
G0278	0.551	0.213	0.194	54.10	60.99	49.35	M	1
G0279	0.411	0.164	0.159	45.27	51.81	40.26	M	1
G0280	0.513	0.202	0.185	51.16	56.03	46.86	M	1
G0282	0.457	0.185	0.172	48.44	53.15	43.62	M	1
G0284	0.516	0.202	0.185	51.88	59.02	46.76	M	1

APPENDIX 2: DETERMINATION OF GONADAL STATUS IN MALE MUSSLES

Table 17 Pairwise comparison of spawning stage groups for genital tissue percentage.

Spawning stage groups	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.
2-1	20.433	11.899	1.717	0.086	0.516
2-3	-35.110	6.500	-5.402	0.000	0.000
2-4	-62.890	10.341	-6.082	0.000	0.000
1-3	-14.677	11.475	-1.279	0.201	1.000
1-4	-42.457	14.012	-3.030	0.002	0.015
3-4	-27.780	9.850	-2.820	0.005	0.029

Table 18 Pairwise comparison of spawning stage groups for spermatogonia percentage of total cells.

Spawning stage groups	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.
4-3	17.380	9.850	1.764	0.078	0.466
4-2	26.555	10.341	2.568	0.010	0.061
4-1	34.057	14.012	2.430	0.015	0.090
3-2	9.175	6.500	1.412	0.158	0.949
3-1	16.677	11.475	1.453	0.146	0.877
2-1	7.502	11.899	0.631	0.528	1.000

Table 19 Pairwise comparison of spawning stage groups for spermatogonia I percentage of total spermatogonia.

Spawning stage groups	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.
2-4	-15.300	10.341	-1.480	0.139	0.834
2-3	-15.640	6.500	-2.406	0.016	0.097
2-1	27.429	11.899	2.305	0.021	0.127
4-3	0.340	9.850	0.035	0.972	1.000
4-1	12.129	14.012	0.866	0.387	1.000
3-1	11.789	11.475	1.027	0.304	1.000

APPENDIX 3: MUCUS CELL QUANTIFICATION IN GILL TISSUE

Table 20 Multiple comparisons between groups derived from manual "Abnormal mucus secretion" scoring. Analysed parameter: mucus cell area per tissue area. Results from post hoc test of One-Way ANOVA. Only p-value reported. Significant p-value in bold.

Groups	0			1			2			3		
	1	2	3	0	2	3	0	1	3	0	1	2
p-value	1.00	**0.01**	1.00	1.00	0.08	1.00	**0.01**	0.08	1.00	1.00	1.00	1.00

APPENDIX 4: COMPARISON OF MUCUS CELL QUANTITY BETWEEN DIFFERENT MUSSEL RIG STATIONS

Table 21 Comparison of mucus cell-related measurements between different locations of the mussel rigs. T0 – subset of mussels sampled to establish pre-exposure health parameters. Image count: amount of images in each group after the tissue detection quality control. Acidic mucus cell count calculated as ratio of dark blue stained cells to the total detected cells, multiplied by 100. Acidic mucus cell area calculated as ratio of dark blue stained cells area to the cumulative area of all detected cells, multiplied by 100.

Station	Image count	Mucus cell count [n] per tissue area [mm^2]		Mucus cell area per tissue area [%]		Acidic mucus cell count [%]		Acidic mucus cells area [%]	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
SFB 500 SE	8	356.1	121.2	1.64	0.67	69.36	8.68	74.14	9.48
SFB 1000 SE	11	393.7	103.5	1.92	0.67	68.47	9.61	74.39	10.24
SFB 10000 SE	10	395.8	78.6	2.01	0.52	66.20	7.21	73.81	7.16
SFA 500 E	9	502.6	155.3	2.61	0.92	74.16	11.13	79.60	10.97
SFA 500 SE 1	7	468.1	158.9	2.52	1.01	65.27	8.93	73.29	7.70
SFA 500 SE 2	9	396.0	152.8	1.99	1.19	67.73	9.40	73.33	8.53
SFA 500 SW	10	409.9	168.7	2.04	0.98	68.94	10.71	75.19	11.27
SFA 500 NE	10	440.6	135.1	2.29	0.88	71.24	8.11	76.85	7.60
SFA 1000 E	9	376.8	111.0	1.81	0.71	63.47	8.17	70.15	8.18
SFA 1000 SE	7	369.5	153.1	1.62	0.80	66.95	13.82	70.70	14.11
SFA 1000 SW	8	408.3	129.9	2.09	0.76	63.32	11.22	69.57	11.03
SFA 1000 NW	9	405.9	87.9	1.99	0.60	62.10	7.61	68.65	5.64
SFA 1000 NE	9	488.4	192.9	2.25	0.89	66.17	7.05	71.30	7.20
SFA 2000 E	8	464.6	177.6	2.23	1.04	66.16	13.45	71.49	12.63
SFA 2000 SE	9	390.7	88.4	1.81	0.44	68.31	7.25	73.10	7.50
SFA 2000 NW	9	508.6	94.2	2.63	0.50	70.44	10.52	77.41	9.72
SFA 10000 SE	10	393.8	206.6	1.93	1.17	68.49	11.75	74.10	12.74
Ref 1	7	418.6	88.1	1.89	0.50	70.20	9.63	75.49	8.08
Ref 2	9	424.3	117.1	2.03	0.66	57.80	14.42	65.46	14.59
T0	9	388.3	103.7	1.54	0.51	61.36	8.64	65.34	10.04

Table 22 Pairwise comparison between groups numbered according to mussel rig stations. Analysed parameter: mucus cell count per gill tissue area. ns – no significance.

Group 1	Group 2	Significance	p-value
Ref 1 + Ref 2	SFB 10000 SE	ns	0.2188
Ref 1 + Ref 2	SFA 500 SW	ns	0.5783
Ref 1 + Ref 2	SFA 500 SE 2	ns	0.2375
Ref 1 + Ref 2	SFA 500 SE 1	*	0.0463
Ref 1 + Ref 2	SFA 500 E	***	0.0002
Ref 1 + Ref 2	SFA 500 NE	ns	0.3583
Ref 1 + Ref 2	SFA 1000 NE	**	0.0019
Ref 1 + Ref 2	SFA 1000 NW	ns	0.4691
Ref 1 + Ref 2	SFA 2000 NW	****	<0.0001
Ref 1 + Ref 2	SFA 10000 SE	ns	0.1846
Ref 1 + Ref 2	SFB 1000 SE	ns	0.1709
Ref 1 + Ref 2	SFB 500 SE	**	0.0035
Ref 1 + Ref 2	SFA 2000 SE	ns	0.1526
Ref 1 + Ref 2	SFA 2000 E	ns	0.0536
Ref 1 + Ref 2	SFA 1000 E	*	0.0383
Ref 1 + Ref 2	SFA 1000 SE	*	0.0265
Ref 1 + Ref 2	SFA 1000 SW	ns	0.5565
Ref 1 + Ref 2	T0	ns	0.1233